Variations on the Standard Model of the Universe

Thesis by

Kris Sigurdson

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2005

(Defended May 26, 2005)

For the other wandering wonderers

Acknowledgements

If you told me not-too-many years ago that I would be so close to finishing a Ph.D. in physics at Caltech studying cosmology, I just wouldn't have believed you. But yet here I am and there are so many people to thank for sending me on this journey and helping me arrive at today.

Deep in my engineering studies at Simon Fraser University three people emboldened me with the courage to considered the other possibility. I would like to thank Howard Trottier for, through his wonderful lectures during the first year of my B.A.Sc and our many subsequent discussions in his office, inspiring me to imagine the Universe and I would like to thank Leigh Palmer and Anthony Arrott for their encouragement during that time.

I would like thank Douglas Scott at the University of British Columbia for being so persuasive on the phone that day in the engineering lab and convincing me to become a summer student in cosmology. While I don't know where I would be right now if I had opted to study string theory instead, I'm very lucky I opted to work with Douglas. He introduced me to the world of cosmology, the Cosmic Microwave Background, and helped send me off to CITA, the University of Cambridge, and Caltech amongst other things. But getting to Caltech isn't even half the story!

Firstly, I'd especially like to thank my advisor Marc Kamionkowski for his advice, encouragement, and wisdom during the past four years here at Caltech. I can remember many occasions where I would storm in to his always-open door and explain to him my latest (crazy?) idea. He would always listen and critique it with healthy skepticism. His comments on physics or otherwise were always helpful and always correct. I hope over the years I have assimilated some of his good judgment about which problems and ideas make for timely and important research.

I would like to thank Mark Wise for letting me be an unofficial member of his group, and for the many enlightening and thought-provoking discussions I had in his office about high energy physics.

I would like to thank my collaborators during my time here at Caltech, and especially those who contributed to the work presented in this thesis: Robert Caldwell, Asantha Cooray, Michael Doran, Steve Furlanetto, Andriy Kurylov, Stefano Profumo, and Piero Ullio.

I would like to thank the members of the TAPIR and Theory groups at Caltech, and especially everyone in Marc's cosmology group. Thanks to JoAnn Boyd, Carol Silberstein, and especially Shirley Hampton for their administrative support.

Last, but not least, I would like to thank all my friends here at Caltech for their support, encouragement, conversations, and diversions over the years: Anura Abeyesinghe, Lotty Ackerman, Igor Bargatin, Daniel Busby, Songye Chen, Ben Collins, Ilya Mandel, Tristan McLoughlin, Alison Farmer, Paul O'Gorman, Margaret Pan, Elizabeth Jones, Mike Kesden, Ilya Mandel, Karin Menendez-Delmestre, Mike Santos, Graeme Smith, Tristan Smith, Sherry Suyu, Ben Toner, Lisa Tracy, Anastasios (Tasos) Vayonakis, Nevin Weinberg, and everyone else I'm forgetting at the moment.

To my officemate, cosmological beer comrade-in-arms, and (soon to be) senior student Jonathan Pritchard: Don't wait for them to install the AC. Get out of 123 Bridge as soon as possible. Get a window man, you've done your time!

To my good friend Alejandro Jenkins: I've enjoyed all our conversations about physics and otherwise. "That reminds me of this episode of the Simpsons."

Most of all I would like to thank my two best friends here at Caltech, the first people I met, Ardis (Disa) Eliasdottir and Donal O'Connell. It would have been so different here without you; I don't think I could have navigated these waters without you keeping me afloat from time to time. Thank you so much.

To J, M, and C: Your love reminds me of what is important in this life, and what is just physics.

Dad, Mom, Ryan: I wouldn't be here without you, without your phone calls from afar, without your encouragement at every single step along the way, without your patience with me, without your visits to me at every far corner of the world, without you pointing out those lights in the sky to me all those years ago.

I would like to thank the Natural Sciences and Engineering Research Council of Canada for supporting me with PGS-A and PGS-D Postgraduate Scholarships, Caltech for giving me a Special Institute Fellowship during my first year, and the PMA Division for giving me Teaching and Research Assistantships. At various points in time the research presented in this thesis was supported at Caltech in part by DoE DE-FG03-92-ER40701 and DE-FG03-02-ER41215, NSF PHY00-71856, and NASA NAG5-9821.

Abstract

In the past decade, due to compelling measurements of the angular power spectrum of the cosmic microwave background (CMB) radiation, the large-scale matter distribution, the recent acceleration of the expansion rate of the Universe over cosmic time, and the current expansion rate (the Hubble constant), cosmology has culminated in a standard model of the Universe. By connecting this standard cosmological model with predictive theories of physics we can hope to look for signatures of these theories in the data. Along this line of inquiry we consider in this thesis: (i) the effects on CMB temperature and polarization anisotropies of spatial fluctuations of the fine-structure parameter α between causally disconnected regions of the Universe at the time of recombination, (ii) the suppression of the small-scale matter power spectrum due to the decay of charged matter to dark matter prior to recombination, (iii) the consequences of a neutral dark-matter particle with a nonzero electric and/or magnetic dipole moment, (iv) how charged-particles decaying in the early Universe can induce a scale-dependent or 'running' spectral index in the small-scale matter power spectrum and examples of this effect in minimal supersymmetric models in which the lightest neutralino is a viable cold-dark-matter candidate. With improved tests and cross-checks of standard-cosmological-model predictions we can search for anomalies that may reveal the character of the underlying physics. In this direction we propose in this thesis: (v) a new method for removing the effect of gravitational lensing from maps of CMB polarization anisotropies using observations of anisotropies or structures in the cosmic 21-cm radiation, (vi) that measurements of fluctuations in the absorption of CMB photons by hydrogen in the 21-cm line and deuterium in the 92-cm line during the cosmic dark ages could be used to determine the primordial deuterium abundance.

Contents

List of Figures

Chapter 1

Introduction

1.1 The Standard Model of the Universe

In the past decade cosmology has culminated in a standard model of the Universe. The emergence of this standard model is due to compelling new data from measurements of the angular power spectrum of the cosmic microwave background (CMB) radiation [1·1], galaxy surveys that have measured the large-scale matter distribution [1·2], observations of Type Ia supernovae [1·3], which have measured the relative expansion rate of the Universe over cosmic time, and the Hubble Space Telescope key project [1·4], which has performed an absolute measurement of the current expansion rate (the Hubble constant H_0). From these observations we now have overwhelmingly compelling evidence that ∼25% of the energy density of the Universe is in the form of dark matter [1·5], ∼70% is in a non-clustering negative-pressure component known as dark energy [1·6] (that acts like Einstein's cosmological constant [1·7] and is driving a contemporary period of accelerated expansion), and surprisingly only ∼5% is in the form of the familiar elements listed on the periodic table and described by the standard model of particle physics. The geometry of the Universe is now known to be flat [1·1], which is presumed to be due to a period of exponential expansion or inflation [1·8] that also, through quantum fluctuations in the inflaton field or other fields present during inflation, created a nearly scale-invariant spectrum of primordial Gaussian density perturbations [1·9] and gravitational waves [1·10]. These fluctuations have been observed in the angular power spectrum of the cosmic microwave background [1·1]. It is believed that, through gravitational instability, these very same fluctuations grew and seeded the formation of galaxies and the large-scale structure in the Universe (e.g. [1·11]). This model forms a consistent cosmological paradigm that has thusfar been consistent with all cosmological observations and can be used as a baseline to test refined or alternate theories of cosmology.

While the emergence of this standard cosmological model (SCM) is remarkable, and it is consistent in so far as it describes the observed cosmology, the physics underlying the key unknown ingredients of dark matter, dark energy, and inflation in the model are not known. We must search

for a physical understanding of these phenomena and equally importantly look for new ways to test and constrain this model. By connecting the SCM with predictive theories of physics we can hope to look for signatures of these theories in the data, and with refined tests and cross-checks of the SCM predictions we can hope to look for anomalies that point the way to the underlying physics. This thesis explores both of these lines of inquiry. In Chapters 2–5 physical theories that have novel predictions for cosmological, astrophysical, or physical observables are considered, while in Chapters 6–7 new techniques for probing the SCM are proposed. These chapters are self-contained, stand on their own, and have been presented in chronological order. In the sections below I provide a brief synopsis of the chapter of the same title including the major results of that chapter, and a description of my primary contributions to the work.

1.2 Spatial Variation of the Fine-Structure Parameter and the Cosmic Microwave Background

In Chapter 2, we consider the effects on CMB temperature and polarization anisotropies of spatial fluctuations of the fine-structure parameter α between causally disconnected regions of the Universe at the time of recombination. Such fluctuations in α might be expected in theories where the gauge coupling constants are set by the vacuum-expectation-value (VEV) of a light scalar field.

We first review previous work [1·12, 1·13] that shows how the recombination history depends on a homogeneous shift in α and then discuss how it affects the CMB power spectrum. We then discuss how spatial variations of α, in which the mean value of α is unaltered, affect the CMB temperature and polarization power spectra. Interestingly, we find these effects are analogous to those of weak gravitational lensing on the CMB, but differ in detail. Like weak-gravitational lensing, spatial variation of α induces a spatially-varying angular power spectrum across the sky. We show how this has the effect of inducing a curl-mode (B-mode) component in the CMB polarization pattern. It also induces "non-Gaussian" signatures in the CMB, in the form of locally anisotropic correlation functions, that cannot be described by the power spectrum alone. We calculate the CMB bispectrum and trispectrum induced by spatial α variation. Effects like those we investigate here may also arise if there are other spatial variations in recombination physics that do not involve significant density/pressure perturbations. Our calculations are thus illustrative theoretically, apart from the specific application on which we focus.

In the final section of Chapter 2, we discuss the properties of and constraints to a toy field-theory model for spatial variation of α that produces the CMB effects outlined above without inducing significant density perturbations.

This chapter was originally published as "Spatial Variation of the Fine-Structure Parameter and the Cosmic Microwave Background," Kris Sigurdson, Andriy Kurylov and Marc Kamionkowski,

Physical Review D **68**, 103509 (2003). I was primarily responsible for deriving the effects of spatial variation of α on the CMB and writing the technical sections related to that work in the paper. Kurylov and I worked closely together on the last section of the paper which discusses the toy field-theory model of α variation. Kamionkowski and I discussed the CMB calculation often during the course of the project, and jointly edited the manuscript.

1.3 Charged-Particle Decay and Suppression of Primordial Power on Small Scales

While the standard cosmological model is in remarkable agreement with observation, on subgalactic scales there are possible problems that warrant further investigation. Namely, the model overpredicts the number of subgalactic halos by an order of magnitude compared to the eleven observed dwarf satellite galaxies of the Milky Way [1·14]. Several possible resolutions have been proposed to this apparent discrepancy, ranging from astrophysical mechanisms that suppress dwarf-galaxy formation in subgalactic halos (see, for example, Ref. [1·15]) to features in the inflaton potential that suppress small-scale power and thus reduce the predicted number of subgalactic halos [1·16].

In Chapter 3 we study the suppression of the small-scale matter power spectrum due to the decay of charged matter to dark matter prior to recombination. In the model discussed, prior to decay, the charged particles couple electromagnetically to the primordial plasma and participate in its acoustic oscillations. After decay, the photon-baryon fluid is coupled only gravitationally to the neutral dark matter. We show how this generically leads to suppression of power for scales that enter the horizon prior to decay. This suppression, reduces the amount of halo substructure on galactic scales while preserving the successes of the standard hierarchical-clustering paradigm on larger scales. For decay times of $\sim$3.5 years this leads to suppression of power on subgalactic scales, bringing the observed number of Galactic substructures in line with observation. We discuss how decay times of a few years may be possible if the dark matter is purely gravitationally interacting, such as the gravitino in supersymmetric models or a massive Kaluza-Klein graviton in models with universal extra dimensions.

This chapter was originally published as "Charged-Particle Decay and Suppression of Primordial Power on Small Scales," Kris Sigurdson and Marc Kamionkowski, *Physical Review Letters* **92**, 171302 (2004). I was primarily responsible for deriving the modified perturbation equations which include the effects of the charged-decay process, and for writing the initial draft of the manuscript. Kamionkowski and I discussed the calculation often during the course of the project, and jointly edited the manuscript.

1.4 Dark-Matter Electric and Magnetic Dipole Moments

A wealth of observational evidence indicates the existence of considerably more mass in galaxies and clusters of galaxies than we see in stars and gas. The source of the missing mass has been a problem since Zwicky's 1933 measurement of the masses of extragalactic systems [1·5]. Given the evidence from galaxy clusters, galaxy dynamics and structure formation, big-bang nucleosynthesis, and the cosmic microwave background that baryons can only account for $\sim 1/6$ of this matter, most of it must be nonbaryonic. Although we know it likely exists, we do not know what the underlying theory of dark matter is or what the detailed particle properties of it are. While theorists have identified promising candidates for the dark matter, such as the neutralino (the supersymmetric partner of the photon, Z^0 boson, and/or Higgs boson) [1·17] or axion [1·18], there is currently no direct evidence that these particles constitute the dark matter. Other candidates are certainly possible.

In Chapter 4, we ask the question, "How dark is 'dark'?" In other words, how weak must the coupling of the dark-matter particle to the photon be in order to be consistent with laboratory and astrophysical constraints? In particular we consider the consequences of a neutral dark-matter particle with a nonzero electric and/or magnetic dipole moment. Theoretical constraints, as well as constraints from direct searches, precision tests of the standard model, the cosmic microwave background and matter power spectra, and cosmic gamma rays, are included. Interestingly, we find that a relatively light particle with mass between an MeV and a few GeV and an electric or magnetic dipole as large as $\sim 3 \times 10^{-16} e$ cm (roughly $1.6 \times 10^{-5}\,\mu_B$) satisfies all experimental and observational bounds and remains a phenomenologically viable candidate for dark matter. Some of the remaining parameter space may be probed with forthcoming, more sensitive, direct searches and with the Gamma-Ray Large Area Space Telescope.

This chapter was originally published as "Dark-Matter Electric and Magnetic Dipole Moments," Kris Sigurdson, Michael Doran, Andriy Kurylov, Robert R. Caldwell and Marc Kamionkowski, *Physical Review D* **70**, 083501 (2004). I computed the dark-matter scattering and annihilation cross sections used throughout the paper and wrote the introductory section on the dipole Lagrangian; the cross sections were independently verified by Kurylov and Doran. I derived the perturbation evolution equations for dipolar-dark-matter, implemented them in a Boltzmann code, and wrote the corresponding section of the paper; these perturbation equations were independently verified by Doran and implemented in a separate Boltzmann code for comparison. Kamionkowski and Caldwell completed the relic abundance, direct detection, and galactic annihilation calculations. Kurylov computed the limits from precision measurements and particle physics. Caldwell and Kamionkowski were primarily responsible for the overall editing of the manuscript.

1.5 A Running Spectral Index in Supersymmetric Dark-Matter Models with Quasi-Stable Charged Particles

In Chapter 5 we show that charged-particles decaying in the early Universe can induce a scale-dependent or 'running' spectral index in the small-scale linear and nonlinear matter power spectrum and discuss examples of this effect in minimal supersymmetric models in which the lightest neutralino is a viable cold-dark-matter candidate. If all of the present-day dark matter is produced through the late decay of charged next-to-lightest dark-matter particles (NLDPs), then, as discussed in Ref. [1·19], the effect is to essentially cut off the matter power spectrum on scales that enter the horizon before the NLDP decays. However, if only a fraction f_ϕ of the present-day dark matter is produced through the late decay of charged NLDPs, the matter power spectrum is suppressed on small scales only by a factor $(1-f_\phi)^2$. This induces a scale-dependent spectral index for wavenumbers that enter the horizon when the age of the Universe is equal to the lifetime of the charged particles. What we show in this chapter is that, for certain combinations of f_ϕ and of the lifetime of the charged particle τ, this suppression modifies the nonlinear power spectrum in a way similar (but different in detail) to the effect of a constant $\alpha_s \equiv dn_s/d\ln k \neq 0$. Although these effects are different, constraints based on observations that probe the nonlinear power spectrum at redshifts of 2 to 4, such as measurements of the Lyman-α forest, might confuse a running index with the effect we describe here even if parametrized in terms of a constant α_s. This has significant implications for the interpretation of the detection of a large running of the spectral index as a constraint on simple single-field inflationary models. The detection of a unexpectedly large spectral running in future observations could instead be revealing properties of the dark-matter particle spectrum in conjunction with a more conventional model of inflation.

While, even with future Lyman-α data, it may be difficult to discriminate the effect of a constant running of the spectral index from a scale-dependent spectral index due to a charged NLDP, other observations may nevertheless discriminate between the two scenarios. Future measurements of the power spectrum of neutral hydrogen through the 21cm-line might probe the linear matter power spectrum in exquisite detail over the redshift range $z \approx 30 - 200$ at comoving scales less than 1 Mpc and perhaps as small as 0.01 Mpc [1·20]; such a measurement could distinguish between the charged-particle decay scenario we describe here and other modifications to the primordial power spectrum. If, as in some models we discuss in this chapter, the mass of these particles is in reach of future particle colliders the signature of this scenario would be spectacular and unmistakable—the production of very long-lived charged particles that slowly decay to stable dark matter.

Although we describe the cosmological side of our calculations in a model-independent manner, remarkably, there are configurations in the minimal supersymmetric extension of the standard model (MSSM) with the right properties for the effect we discuss here. In particular, we find that if the

lightest supersymmetric particle (LSP) is a neutralino quasi-degenerate in mass with the lightest stau, we can naturally obtain, at the same time, LDPs providing the correct dark matter abundance $\Omega_\chi h^2 = 0.113$ [1·1] and NLDPs with the long lifetimes and the sizable densities in the early Universe needed in the proposed scenario. Such configurations arise even in minimal supersymmetric schemes, such as the minimal supergravity (mSUGRA) scenario [1·21] and the minimal anomaly-mediated supersymmetry-breaking (mAMSB) model [1·22]. This implies that a detailed study of the (τ,f_ϕ) parameter space using current and future cosmological data may constrain regions of the MSSM parameter space that are otherwise viable. We discuss the expected signatures of this scenario at future particle colliders, such as the large hadron collider (LHC), and prospects for detection in experiments searching for WIMP dark matter.

This chapter was originally published as "A Running Spectral Index in Supersymmetric Dark-Matter Models with Quasi-Stable Charged Particles," Stefano Profumo, Kris Sigurdson, Piero Ullio, and Marc Kamionkowski, *Physical Review D* **71**, 023518 (2005). I computed the modified perturbation evolution equations, calculated the modified matter power spectra, and wrote the corresponding sections of the paper. Ullio and Profumo worked out how to track the abundances of multiple coannihilating species, and Profumo was responsible for finding example supersymmetric models that produced the desired effect. Kamionkowski and I were responsible for editing the manuscript.

1.6 Cosmic 21-cm Delensing of Microwave Background Polarization and the Minimum Detectable Energy Scale of Inflation

The curl (B) modes of cosmic microwave background (CMB) polarization anisotropies are a unique probe of the primordial background of cosmological gravitational waves [1·23]. At these long wavelengths, inflation [1·8] is the only known mechanism to causally generate such a background of gravitational waves [1·10]. Since the amplitude of these inflationary gravitational waves (IGWs) is proportional to $\mathcal{V}$, the value of the inflaton potential $V(\varphi)$ during inflation, the amplitude of gravitational-wave induced B-mode polarization anisotropies directly constrains the energy scale of inflation $\mathcal{V}^{1/4}$ (see, for example, Ref. [1·24]). While the experimental sensitivity to B-mode polarization can be improved, the expected signal is contaminated by foreground effects [1·25]. The main confusion to the detection of B-mode polarization anisotropies generated by IGWs at recombination is the mixing of gradient-mode (E-mode) and B-mode anisotropies via gravitational lensing [1·26].

In Chapter 6, we propose a new method for removing the effect of gravitational lensing from maps of cosmic microwave background (CMB) polarization anisotropies. Using observations of anisotropies or structures in the cosmic 21-cm radiation, which was emitted or absorbed by neutral hydrogen

atoms that underwent a spin-flip transition at redshifts 10 to 200, the CMB can be delensed. We find that this method could allow CMB experiments to have increased sensitivity to a primordial background of inflationary gravitational waves (IGWs) compared to methods which rely on CMB observations alone — reducing the minimum detectable energy scale of inflation below 10^{15} GeV. While the detection of cosmic 21-cm anisotropies at high resolution is a challenging endeavor, the detection of these fluctuations is already being pursued as a probe of the Universe at or before the epoch of reionization. A combined study with a relatively low-resolution (but high-sensitivity) CMB polarization experiment may constrain alternative models of inflation which were heretofore considered to have undetectable IGW amplitudes. The ultimate theoretical limit to the detectable inflationary energy scale via this method may be as low as 3×10^{14} GeV.

This chapter was originally available online as "Cosmic 21-cm Delensing of Microwave Background Polarization and the Minimum Detectable Energy Scale of Inflation," Kris Sigurdson and Asantha Cooray, arXiv:astro-ph/0502549 and has been submitted to *Physical Review Letters*. I had the initial idea of using cosmic 21-cm radiation to infer the projected potential. Cooray calculated the curves shown in the figures. Cooray and I jointly wrote the text of the paper, and I was the primary editor.

1.7 Measuring the Primordial Deuterium Abundance During the Cosmic Dark Ages

In Chapter 7 we discuss how measurements of fluctuations in the absorption of cosmic microwave background (CMB) photons by neutral gas during the cosmic dark ages, at redshifts $z \approx 7$–200, could reveal the primordial deuterium abundance of the Universe.

After the cosmic microwave background (CMB) radiation decoupled from the baryons at a redshift $z \approx 1100$, most CMB photons propagated unfettered through the neutral primordial medium. This has allowed exquisite measurements of the temperature fluctuations in the primordial plasma at the surface of last scattering, and the statistical properties of these fluctuations have recently been used, in conjunction with other observations, to determine the cosmology of our Universe [1·1]. After the photons kinetically decoupled from the gas at $z \sim 200$, the latter cooled adiabatically with $T_{\rm g} \propto (1+z)^2$, faster than the $T_\gamma \propto (1+z)$ cooling of the CMB. This epoch, with most of the baryons in the form of relatively cold neutral atoms and before the first stars formed, is known as the cosmic dark ages.

The reason most CMB photons propagate unimpeded through the neutral primordial gas is elementary quantum mechanics — atoms absorb non-ionizing radiation only at the discrete wavelengths determined by the differences of their atomic energy levels. One interesting example is the well-known 21-cm spin-flip transition [1·27], due to the hyperfine splitting of the ground state of the

hydrogen (H) atom. At any given z, CMB photons with wavelength $\lambda_{21} = 21.1$ cm can resonantly excite this transition. By measuring brightness-temperature fluctuations due to density fluctuations in the neutral gas [1·28], radio telescopes observing at $\lambda = (1+z)\lambda_{21}$ can probe the matter power spectrum at $z \approx 30$–200 [1·20].

In this chapter we discuss another application of these measurements. Less well-known than the 21-cm transition of neutral H is the spin-flip transition of neutral deuterium (D) at $\lambda_{92} = 91.6$ cm [1·29,1·30]. We show that the strength of the cross-correlation of brightness-temperature fluctuations at a wavelength $\lambda_{\rm H} = (1+z)\lambda_{21}$ due to resonant absorption of CMB photons in the 21-cm line of neutral hydrogen with those at a wavelength $\lambda_{\rm D} = (1+z)\lambda_{92}$ due to resonant absorption of CMB photons in the 92-cm line of neutral deuterium is proportional to the fossil deuterium to hydrogen ratio [D/H] fixed during big bang nucleosynthesis (BBN). In principle, a sufficiently large future experiment could constrain the primordial value of $[{\rm D/H}] \equiv n_{\rm D}/n_{\rm H}$ to better than 1% without any systematics. While there is no physical obstacle to such a measurement, it would certainly be technically challenging; simply detecting neutral D during the cosmic dark ages would be a significantly easier goal and may be possible with next-generation cosmic 21-cm experiments.

We emphasize that, although technically challenging, this measurement could provide the cleanest possible determination of [D/H], free from contamination by structure formation processes at lower redshifts, and has the potential to improve BBN constraints to the baryon density of the Universe $\Omega_{\rm b}h^2$. In this chapter we also present our results for the thermal spin-change cross-section for deuterium-hydrogen scattering, which may be useful in a more general context than we describe here.

This chapter was originally available online as "Measuring the Primordial Deuterium Abundance During the Cosmic Dark Ages," Kris Sigurdson and Steven R. Furlanetto, arXiv:astro-ph/0505173 and has been submitted to *Physical Review Letters*. I had the initial idea of using cosmic 92-cm radiation to measure the primordial deuterium abundance through cross-correlation with the hydrogen signal. I calculated the deuterium-hydrogen spin-change cross section, the spin temperature evolution, the brightness temperature evolution, and the brightness temperature fluctuations. Furlanetto made the initial estimate of the signal-to-noise and I wrote a code to make a more realistic estimate. Furlanetto and I had many discussions during the course of this project and he contributed his extensive knowledge of cosmic 21-cm fluctuations. I wrote the bulk of the text of the paper, and Furlanetto and I jointly edited it.

Bibliography

[1·1] P. de Bernardis *et al.* [Boomerang Collaboration], Nature **404**, 955 (2000); S. Hanany *et al.*, Astrophys. J. **545**, L5 (2000); N. W. Halverson *et al.*, Astrophys. J. **568**, 38 (2002); B. S. Mason

et al., Astrophys. J. **591**, 540 (2003); A. Benoit *et al.* [the Archeops Collaboration], Astron. Astrophys. **399**, L25 (2003); D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148**, 175 (2003);

[1·2] J. A. Peacock *et al.*, Nature **410**, 169 (2001); W. J. Percival, Mon. Not. Roy. Astron. Soc. **327**, 1313 (2001); M. Tegmark *et al.* [SDSS Collaboration], Astrophys. J. **606**, 702 (2004).

[1·3] J. L. Tonry *et al.* [Supernova Search Team Collaboration], Astrophys. J. **594**, 1 (2003); R. A. Knop *et al.* [The Supernova Cosmology Project Collaboration], Astrophys. J. **598**, 102 (2003); B. J. Barris *et al.*, Astrophys. J. **602**, 571 (2004); A. G. Riess *et al.* [Supernova Search Team Collaboration], Astrophys. J. **607**, 665 (2004).

[1·4] W. L. Freedman *et al.*, Astrophys. J. **553**, 47 (2001).

[1·5] F. Zwicky, Helv. Phys. Acta **6**, 110 (1933).

[1·6] M. S. Turner and M. J. White, Phys. Rev. D **56**, 4439 (1997); R. R. Caldwell, R. Dave and P. J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998).

[1·7] A. Einstein, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1917**, 142 (1917).

[1·8] A. H. Guth, Phys. Rev. D **23**, 347 (1981); A. D. Linde, Phys. Lett. B **108**, 389 (1982); A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).

[1·9] J. M. Bardeen, P. J. Steinhardt and M. S. Turner, Phys. Rev. D **28**, 679 (1983).

[1·10] A. A. Starobinsky, JETP Lett. **30**, 682 (1979) [Pisma Zh. Eksp. Teor. Fiz. **30**, 719 (1979)].

[1·11] P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton University Press, Princeton, 1993).

[1·12] S. Hannestad, Phys. Rev. D **60**, 023515 (1999).

[1·13] M. Kaplinghat, R. J. Scherrer and M. S. Turner, Phys. Rev. D **60**, 023516 (1999).

[1·14] G. Kauffmann, S. D. M. White and B. Guiderdoni, Mon. Not. Roy. Astron. Soc. **264**, 201 (1993); A. A. Klypin, A. V. Kravtsov, O. Valenzuela and F. Prada, Astrophys. J. **522**, 82 (1999); B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel and P. Tozzi, Astrophys. J. **524**, L19 (1999).

[1·15] A. J. Benson, C. S. Frenk, C. G. Lacey, C. M. Baugh and S. Cole, Mon. Not. Roy. Astron. Soc. **333**, 177 (2002); R. S. Somerville, Astrophys. J. **572**, L23 (2002); L. Verde, S. P. Oh and R. Jimenez, Mon. Not. Roy. Astron. Soc. **336**, 541 (2002).

[1·16] M. Kamionkowski and A. R. Liddle, Phys. Rev. Lett. **84**, 4525 (2000).

[1·17] G. Jungman, M. Kamionkowski and K. Griest, Phys. Rept. **267**, 195 (1996).

[1·18] M. S. Turner, Phys. Rept. **197**, 67 (1990); G. G. Raffelt, Phys. Rept. **198**, 1 (1990); L. J. Rosenberg and K. A. van Bibber, Phys. Rept. **325**, 1 (2000).

[1·19] K. Sigurdson and M. Kamionkowski, Phys. Rev. Lett. **92**, 171302 (2004).

[1·20] A. Loeb and M. Zaldarriaga, Phys. Rev. Lett. **92**, 211301 (2004).

[1·21] A. H. Chamseddine, R. Arnowitt and P. Nath, Phys. Rev. Lett. **49**, 970 (1982); R. Barbieri, S. Ferrara and C. A. Savoy, Phys. Lett. B **119**, 343 (1982); L. J. Hall, J. Lykken and S. Weinberg; Phys. Rev. D **27**, 2359 (1983); P. Nath, R. Arnowitt and A. H. Chamseddine, Nucl. Phys. B **227**, 121 (1983).

[1·22] L. Randall and R. Sundrum, Nucl. Phys. B **557**, 79 (1999); G. F. Giudice, M. A. Luty, H. Murayama and R. Rattazzi, JHEP **9812**, 027 (1998); T. Gherghetta, G. F. Giudice and J. D. Wells, Nucl. Phys. B **559**, 27 (1999).

[1·23] U. Seljak and M. Zaldarriaga, Phys. Rev. Lett. **78**, 2054 (1997); M. Kamionkowski, A. Kosowsky and A. Stebbins, Phys. Rev. Lett. **78**, 2058 (1997).

[1·24] M. Kamionkowski and A. Kosowsky, Ann. Rev. Nucl. Part. Sci. **49**, 77 (1999).

[1·25] M. Tucci, E. Martinez-Gonzalez, P. Vielva and J. Delabrouille, arXiv:astro-ph/0411567.

[1·26] M. Zaldarriaga and U. Seljak, Phys. Rev. D **58**, 023003 (1998); W. Hu, Phys. Rev. D **62**, 043007 (2000).

[1·27] E. Fermi, Zeits. f. Physik **60**, 320 (1930).

[1·28] C. J. Hogan and M. J. Rees, Mon. Not. Roy. Astron. Soc. **188**, 791 (1979); D. Scott and M. J. Rees, Mon. Not. Roy. Astron. Soc. **247**, 510 (1990).

[1·29] J. E. Nafe, E. B. Nelson, and I. I. Rabi, Phys. Rev. **71** 914 (1947).

[1·30] S. Weinreb, Nature **195** 367 (1962); J. M. Pasachoff and D. A. Cesarsky, Astrophys. J. **193**, 65 (1974); K. R. Anantharamaiah and V. Radhakrishnan, Astron. Astrophys. **79**, L9 (1979); L. Blitz and C. Heiles, Astrophys. J. Lett. **313**, L95 (1987); J. N. Chengalur, R. Braun and W. B. Burton, Astron. Astrophys. **318**, L35 (1997).

Chapter 2

Spatial Variation of the Fine-Structure Parameter and the Cosmic Microwave Background

We study the effects on cosmic microwave background (CMB) temperature and polarization anisotropies of spatial fluctuations of the fine-structure parameter between causally disconnected regions of the Universe at the time of recombination. Analogous to weak gravitational lensing, in addition to modifying the mean power spectra and inducing a curl component (B mode) to the polarization, spatial fluctuations of the fine-structure parameter induce higher-order (non-Gaussian) temperature and polarization correlations in the CMB. We calculate these effects for the general case of arbitrary correlation between temperature fluctuations and fine-structure parameter fluctuations, and show the results for a model where these two types of fluctuations are uncorrelated. The formalism we present here may also be applied to other modifications of recombination physics that do not significantly alter the evolution of the dominant density perturbations. We discuss the constraints on the effective Lagrangian for variable fine-structure parameter necessary to realize this scenario.

Originally published as K. Sigurdson, A. Kurylov and M. Kamionkowski, *Phys. Rev. D* **68**, 103509 (2003).

2.1 Introduction

The possibility that the fine-structure parameter α may vary in time has long been entertained [2·1–2·8], and has received renewed interest with recent evidence from quasar spectra that may support a variation of less than one part in 10^4 over a time scale of ~ 10 Gyr [2·9, 2·10]. Although the results may still be controversial, the observational work has inspired theoretical work that investigates models with variable α [2·11–2·18], as well as other work that investigates possible connections with dark energy and new long-range forces [2·19–2·21]. It has also stimulated a more careful re-investigation of the constraints placed on variable α from big-bang nucleosynthesis [2·22].

Of course, if a relativistic theory allows for temporal variation of α then it must also allow for spatial variations of α between regions not in causal contact. In this chapter we study cosmological probes of spatial variations in α, focusing in particular on the cosmic microwave background (CMB), which is rapidly becoming an increasingly precise probe of cosmological models [2·23–2·25], as well as the physics that underlies them [2·26]. As we will see, spatial variation in α induces a spatially-varying power spectrum. This induces "non-Gaussian" signatures in the CMB, in the form of locally anisotropic correlation functions, that cannot be described by the power spectrum alone (although strictly speaking the joint probability distribution of temperatures at n points remains a multivariate Gaussian). Quite interestingly, these effects are analogous to those of weak gravitational lensing ("cosmic shear") on the CMB, but differ in detail. So, for example, spatial variation of α can alter slightly the CMB power spectrum, induce a curl component (B mode) in the polarization, and induce higher-order temperature and polarization correlations in the CMB. Effects like those we investigate here may also arise if there are other spatial variations in recombination physics that do not involve significant density/pressure perturbations. Our calculations are thus illustrative theoretically, apart from the specific application on which we focus.

Below, we first review previous work [2·27, 2·28] that shows how the recombination history depends on a homogeneous shift in α and then discuss how it affects the CMB power spectrum. We then show how spatial variations of α, in which the mean value of α is unaltered, affect the CMB temperature and polarization power spectra, and in so doing show that a curl component is induced in the CMB polarization. We then calculate the CMB bispectrum and trispectrum induced by spatial α variation. Throughout, we compare with the analogous calculations for weak lensing, and show how the effects of weak lensing and spatial α variations differ. We then discuss the properties of and constraints on a toy field-theory model for spatial α variation that produces the CMB effects we investigate here, without inducing significant density perturbations.

Before proceeding further, we clarify that here we investigate spatial variations in the fine-structure parameter $\alpha = e^2/\hbar c$ that arise *only* from spatial variations in the electromagnetic gauge coupling e; we do not tinker with relativity nor quantum mechanics.

2.2 Recombination and α

Recombination depends on the value of α because the visibility function, the probability distribution of when a photon last scattered, is dependent on α. The visibility function is defined as

$$g(t) = e^{-\tau} \frac{d\tau}{dt}\,, \tag{2.1}$$

where

$$\frac{d\tau}{dt} = x_e n_p c \sigma_T \,, \tag{2.2}$$

is the differential optical depth of photons due to Thomson scattering. Here,

$$\sigma_T = \frac{8\pi\hbar^2\alpha^2}{3m_e^2c^2} \,, \tag{2.3}$$

is the Thomson cross section, n_p is the total number density of protons (both free and bound), and x_e is the fraction of free electrons. The strongest effect of variations of α on this quantity occur due to the alteration of the ionization history $x_e(t)$.

The recombination of hydrogen cannot proceed through direct recombination to the ground state because the emitted photon will immediately ionize a neighboring atom with high probability. Instead, the ionized fraction decreases mainly through the two-photon process $2s \to 1s$, or via the cosmological redshifting of $2p \to 1s$ Lyman-α photons out of the Lyman-α line. These processes are described by a single differential equation [2·29],

$$\frac{dx_e}{dt} = \mathcal{C}\left[\beta e^{-\frac{B_1-B_2}{k_BT}}(1-x_e) - \mathcal{R}n_p x_e^2\right] \,, \tag{2.4}$$

where β is the ionization coefficient, $\mathcal{R}$ is the recombination coefficient, $\mathcal{C}$ is the Peebles efficiency factor (discussed below), and

$$B_n = \frac{m_e c^2 \alpha^2}{2n^2} \,, \tag{2.5}$$

is the binding energy of the level with principle quantum number n.

Through the Einstein relations we relate β to $\mathcal{R}$ through

$$\beta = \left(\frac{m_e k_B T}{2\pi\hbar^2}\right)^{\frac{3}{2}} e^{-\frac{B_2}{k_BT}} \mathcal{R} \,. \tag{2.6}$$

The recombination coefficient can be written

$$\mathcal{R} = \sum_{n=2}^{\infty} \sum_{l=0}^{n-1} \alpha_{nl} w_n \,, \tag{2.7}$$

where

$$\alpha_{nl} = \frac{8\pi(2l+1)}{(2\pi m_e k_B T)^{\frac{3}{2}} c^2} e^{\frac{B_n}{k_BT}} \int_{B_n}^{\infty} d(h\nu) \sigma_{nl}^{bf} \frac{(h\nu)^2}{e^{\frac{h\nu}{k_BT}} - 1} \,, \tag{2.8}$$

is the rate at which atoms recombine to the n, l energy level and w_n is the efficiency for an n level to survive in a plasma [2·30]. The details of w_n are not important for the present discussion, other than to note that at the densities of interest it is unity for $n < n_{max} \sim 500$. Since the dominant contributions to $\mathcal{R}$ come from $n \lesssim 50$, $\mathcal{R}$ is insensitive to the weak α dependence of n_{max}. Above,

α_{nl} is written in terms of the ionization cross section σ_{nl}^{bf}, which can be expressed in the form [2·31]

$$\sigma_{nl}^{bf} = \alpha^{-1} f_n \left(\frac{h\nu}{B_1}\right) . \tag{2.9}$$

Thus, we can write

$$\alpha_{nl} = \frac{8\pi(2l+1)}{\alpha c^2} \left(\frac{k_B T}{2\pi m_e}\right)^{\frac{3}{2}} e^{\xi/n^2} \int_{\xi/n^2}^{\infty} dy f_n \left(\frac{y}{\xi}\right) \frac{y^2}{e^y - 1}, \tag{2.10}$$

where

$$\xi = \frac{B_1}{k_B T} = \frac{m_e c^2 \alpha^2}{2 k_B T} , \tag{2.11}$$

and it immediately follows that the α dependence of $\mathcal{R}$ is of the form,

$$\mathcal{R} = \alpha^{-1} T^{\frac{3}{2}} F(\xi) = \alpha^2 G(\xi) . \tag{2.12}$$

As shown in Ref. [2·32], for the temperatures of interest,

$$\mathcal{R} = \frac{64}{3} \frac{\hbar^2 \alpha^2}{m_e^2 c} \sqrt{\frac{\pi B_1}{3 k_B T}} \phi_2 , \tag{2.13}$$

where

$$\phi_2 \simeq \frac{13\sqrt{3}}{16\pi} \ln\left(\frac{B_1}{k_B T}\right) . \tag{2.14}$$

Given the scaling in Eq. (2.12), we can then read off the α dependence. Explicitly, the recombination coefficient is

$$\mathcal{R} = \frac{52}{3} \frac{\hbar^2 \alpha^3}{\sqrt{2\pi m_e^3 k_B T}} \ln\left(\frac{m_e c^2 \alpha^2}{2 k_B T}\right) . \tag{2.15}$$

The rate of recombination is inhibited by ionizing photons which can disrupt atoms in the $n = 2$ state before they can decay to the ground state. The efficiency of recombination from the $n = 2$ state is described by the Peebles efficiency factor,

$$\mathcal{C} = \frac{\Lambda_H + \Lambda_{2s\to 1s}}{\Lambda_H + \Lambda_{2s\to 1s} + \beta} , \tag{2.16}$$

which is just the ratio of the recombination rates to the sum of the recombination and ionization rates from the $n = 2$ level.

In Eq. (2.16),

$$\Lambda_H = \frac{8\pi H}{(\lambda_{2p\to 1s})^3 n_{1s}} , \tag{2.17}$$

is the rate at which a recombination is successful because the emitted Lyman-α photon is redshifted out of the Lyman-α line before ionizing a hydrogen atom. Here, $H = (1/a)(da/dt)$ is the Hubble

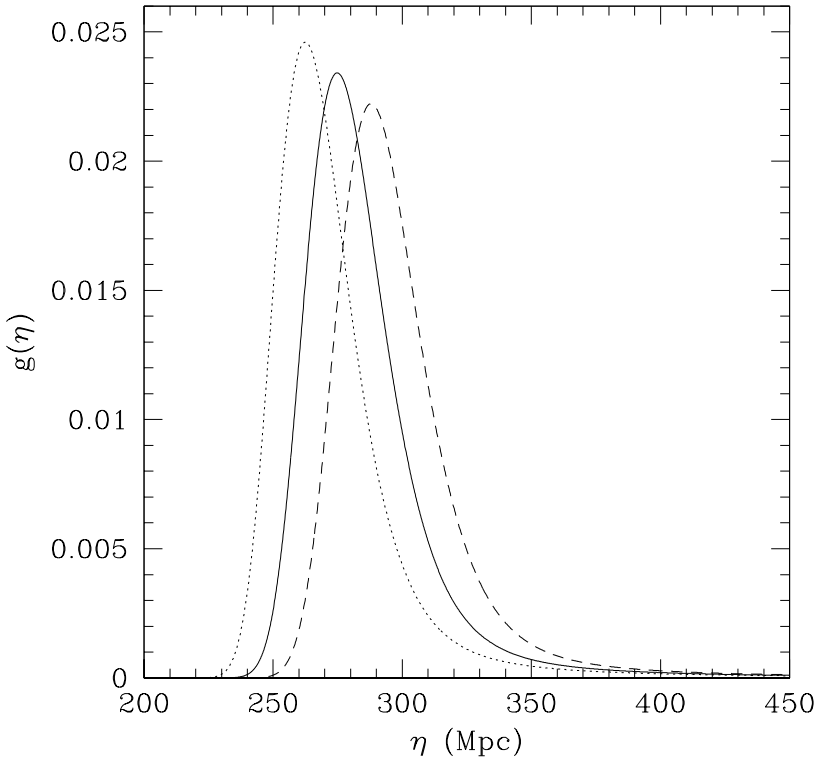

Figure 2.1: The probability distribution for the last scattering of a photon, the visibility function, as a function of conformal time η in the ΛCDM model for $\varphi = (\alpha - \alpha_0)/\alpha_0 = 0.03$ (dotted), $\varphi = 0$ (solid), and $\varphi = -0.03$ (dashed).

expansion rate, $n_{1s} \simeq (1 - x_e) n_p$ is the number density of atoms in the $1s$ state (almost all hydrogen atoms are in the $1s$ state), and

$$\lambda_{2p\to 1s} = \frac{16\pi\hbar}{3 m_e c \alpha^2}\,, \tag{2.18}$$

is the Lyman-α rest wavelength.

The two-photon process $2s \to 1s$ proceeds through virtual atomic states at a rate $\Lambda_{2s\to 1s} = 8.22458\ \mathrm{s}^{-1}$ [2·33], and scales as [2·34]

$$\Lambda_{2s\to 1s} \propto \alpha^8\,. \tag{2.19}$$

Eqs. (2.4)–(2.19) account for the α dependence of $x_e(t)$. Along with the α dependence of σ_T, this completely determines how $g(t)$ varies with α in a given cosmology. While we expect a more complete calculation of recombination, such as that in Ref. [2·35], may yield further refinements to the α dependence, modifying hydrogen recombination as described above is adequate for our purposes

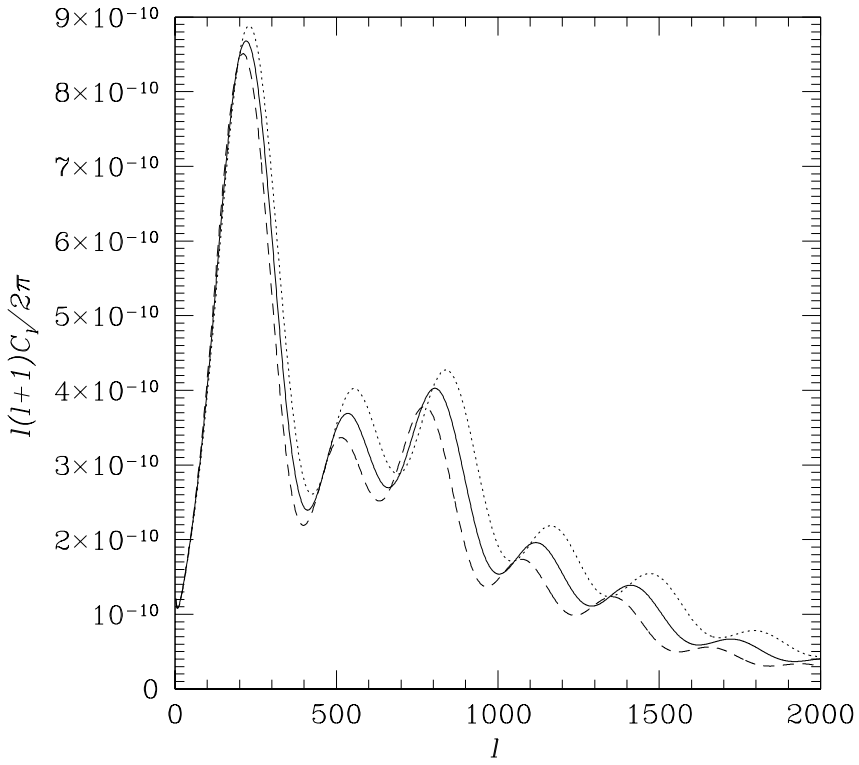

Figure 2.2: The angular power spectrum of CMB anisotropies for a spatially uniform $\varphi = 0.03$ (dotted), $\varphi = 0$ (solid), and $\varphi = -0.03$ (dashed).

because we are primarily concerned with the α dependence of the visibility function at very high redshift where the simple calculation accurately captures the physics. Also, it was determined in Ref. [2·28] that other effects, such as modifications to the details of helium recombination, or the cooling of baryons, are small compared to the effect of variations of α on hydrogen recombination.

In this chapter we work within the flat geometry ΛCDM cosmology with baryon and matter densities $\Omega_b = 0.05$ and $\Omega_m = 0.30$, Hubble parameter $h = 0.72$, and spectral index $n = 1$. Fig. 2.1 shows the visibility function $g(\eta) = \exp(-\tau) d\tau/d\eta$ plotted versus the conformal time $\eta = \int dt/a$ for three different values of $\varphi = (\alpha - \alpha_0)/\alpha_0$ where $\alpha_0 = 0.00729735 \simeq 1/137$ is the value of the electromagnetic fine structure parameter [2·36]. For positive values of φ the visibility function is narrower and peaks earlier, while for negative values of φ the visibility function is broader and peaks later. These effects impact directly the CMB angular power spectrum because the peak of the visibility function determines the physical distance to the last scattering surface, while the width of the visibility function determines the thickness of the last scattering surface.

Fig. 4.13 shows the angular power spectrum of CMB anisotropies calculated assuming a spatially homogeneous value of φ. The dependence of the spectrum on φ is easy to understand qualitatively.

For positive φ the angular-diameter distance is larger, so the features are scaled systematically to higher values of l. The last-scattering surface is also narrower so that small-scale (high-l) features are less damped due to photon diffusion. For negative φ the opposite holds, the smaller angular-diameter distance scales features to lower values of l while the broader last-scattering surface leads to more damping of power on small scales. In the next two Sections we derive the changes in the angular power spectrum due to spatial fluctuations in φ between causally disconnected regions of the Universe.

2.3 Power Spectra

2.3.1 CMB Power Spectra Fundamentals

The CMB radiation is observed to be a nearly isotropic background of blackbody radiation at a temperature of $T_{\rm CMB} = 2.728 \pm 0.004\,{\rm K}$ [2·37]. Anisotropies in the temperature are observed with a fractional amplitude of $\sim 10^{-5}$ [2·38], and in the polarization with a fractional amplitude of $\sim 10^{-6}$ [2·39]. For a review of the physics of CMB anisotropies see, e.g., Ref. [2·40].

The fundamental CMB anisotropy observables are the Stokes parameters (Θ, Q, U, V), which can be expressed in the Pauli basis as (see, for example, Ref. [2·41])

$$\mathbf{P}(\hat{\mathbf{n}}) = \Theta(\hat{\mathbf{n}})\mathbf{1} + Q(\hat{\mathbf{n}})\sigma_{\mathbf{3}} + U(\hat{\mathbf{n}})\sigma_{\mathbf{1}} + V(\hat{\mathbf{n}})\sigma_{\mathbf{2}}\,. \tag{2.20}$$

In this expression,

$$\Theta(\hat{\mathbf{n}}) = \frac{T(\hat{\mathbf{n}}) - T_{\rm CMB}}{T_{\rm CMB}}\,, \tag{2.21}$$

denotes the fractional temperature anisotropy in a direction $\hat{\mathbf{n}}$, and the remaining Stokes parameters are normalized to this quantity. Here Q and U describe independent linear polarization states, while V describes circular polarization. Because circular polarization cannot be generated via Thomson scattering, $V = 0$ for the CMB.

It is convenient to introduce the quantities,

$${}_{\pm}A(\hat{\mathbf{n}}) = Q(\hat{\mathbf{n}}) \pm iU(\hat{\mathbf{n}})\,, \tag{2.22}$$

which have the spin-2 transformation properties,

$${}_{\pm}A(\hat{\mathbf{n}}) \rightarrow e^{\mp 2i\phi}{}_{\pm}A(\hat{\mathbf{n}})\,, \tag{2.23}$$

under a counterclockwise rotation of the coordinate axis by an angle ϕ.

We can expand $\Theta(\hat{\mathbf{n}})$ and ${}_{\pm}A(\hat{\mathbf{n}})$ in normal modes as [2·42],

$$\Theta(\hat{\mathbf{n}}) = \sum_{l=1}^{\infty} \sum_{m=-l}^{m=l} (-i)^l \sqrt{\frac{4\pi}{2l+1}} \Theta_{lm} Y_l^m(\hat{\mathbf{n}}) \,, \tag{2.24}$$

and

$${}_{\pm}A(\hat{\mathbf{n}}) = \sum_{l=1}^{\infty} \sum_{m=-l}^{m=l} (-i)^l \sqrt{\frac{4\pi}{2l+1}} \left({}_{\pm}A_{lm}\right) \left({}_{\pm 2}Y_l^m(\hat{\mathbf{n}})\right) \,, \tag{2.25}$$

where

$$\Theta_{lm} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \Theta_l^{(m)}(k) e^{i\mathbf{k}\cdot\mathbf{x}} \,, \tag{2.26}$$

and

$${}_{\pm}A_{lm} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left({}_{\pm}A_l^{(m)}(k)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \,. \tag{2.27}$$

Here ${}_sY_l^m$ are the spin-s weighted spherical harmonics [2·43], with $Y_l^m = {}_0Y_l^m$, and $X_l^{(m)}(k)$ is the contribution to the angular mode X_{lm} from wave vectors of the primordial density field of magnitude k.

It is conventional to write the polarization in terms of the moments of the curl-free (scalar) configurations E_{lm}, and the moments of the divergence-free (pseudo-scalar) configurations B_{lm}, where

$${}_{\pm}A_{lm} = E_{lm} \pm iB_{lm} \,. \tag{2.28}$$

We can then provide a complete description of an arbitrary CMB anisotropy field using the moments Θ_{lm}, E_{lm}, and B_{lm}.

The basic observables of the random fields $X(\hat{\mathbf{n}})$ are the power spectra $C_l^{X\tilde{X}}$, defined by,

$$\langle X_{lm}^* \widetilde{X}_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l^{X\tilde{X}} \,, \tag{2.29}$$

where $X, \widetilde{X} \in \{\Theta, E, B\}$ and the angle brackets denote an average over all realizations. Here,

$$C_l^{X\tilde{X}} = \frac{2}{\pi(2l+1)^2} \int \frac{dk}{k} \sum_{m=-2}^{2} k^3 X_l^{(m)*}(k) \widetilde{X}_l^{(m)}(k) \,. \tag{2.30}$$

A set of Gaussian random fields—and we expect $\{\Theta, E, B\}$ to be Gaussian—are completely characterized by their power spectra and cross-power spectra. Because the pseudo-scalar B has opposite parity to the scalars Θ and E, the only non-vanishing power spectra are $C_l^{\Theta\Theta}$, $C_l^{\Theta E}$, C_l^{EE}, and C_l^{BB}.

For small patches of sky it is an excellent approximation to treat the sky as flat and expand the field $X(\hat{\mathbf{n}})$ in spin-weighted Fourier modes rather than spin-weighted spherical harmonics. Thus we

have (for example, Ref. [2·44])

$$\Theta(\hat{\mathbf{n}}) = \int \frac{d^2\mathbf{l}}{(2\pi)^2}\Theta(\mathbf{l})e^{i\mathbf{l}\cdot\hat{\mathbf{n}}}\,, \tag{2.31}$$

$${}_{\pm}A(\hat{\mathbf{n}}) = -\int \frac{d^2\mathbf{l}}{(2\pi)^2}{}_{\pm}A(\mathbf{l})e^{\pm 2i(\phi_l-\phi)}e^{i\mathbf{l}\cdot\hat{\mathbf{n}}}\,, \tag{2.32}$$

and we again define E and B through

$${}_{\pm}A(\mathbf{l}) = E(\mathbf{l}) \pm iB(\mathbf{l})\,. \tag{2.33}$$

In this notation the power spectra are defined by

$$\langle X(\mathbf{l})\widetilde{X}(\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')C_l^{X\tilde{X}}\,. \tag{2.34}$$

Unless otherwise noted, we work within this flat-sky approximation for the remainder of the chapter.

2.3.2 Derivative Power Spectra

How do the expressions for the power spectra change if we allow for spatial fluctuations of α? As a warmup we first consider a spatially uniform variation, $\alpha = \alpha_0(1+\varphi)$, where $\varphi \ll 1$.

For a given primordial density field $\delta(\mathbf{x})$, the temperature and polarization patterns, $\Theta(\hat{\mathbf{n}})$ and ${}_{\pm}A(\hat{\mathbf{n}})$, can be calculated by solving the combined Einstein equations and radiative-transfer equations, as well as the equations for the recombination history. As discussed above, this recombination history depends on α. Thus, the temperature and polarization fields are implicitly functions of $\varphi = (\alpha-\alpha_0)/\alpha_0$. We can expand $\Theta(\hat{\mathbf{n}}) = \Theta(\hat{\mathbf{n}};\varphi)$ and ${}_{\pm}A(\hat{\mathbf{n}}) = {}_{\pm}A(\hat{\mathbf{n}};\varphi)$ in Taylor series about $\varphi = 0$,

$$\Theta(\hat{\mathbf{n}}) = \Theta_0(\hat{\mathbf{n}}) + \partial_\varphi\Theta_0(\hat{\mathbf{n}})\varphi + \frac{1}{2}\partial_\varphi^2\Theta_0(\hat{\mathbf{n}})\varphi^2 + \cdots\,, \tag{2.35}$$

$${}_{\pm}A(\hat{\mathbf{n}}) = {}_{\pm}A_0(\hat{\mathbf{n}}) + \partial_\varphi\left({}_{\pm}A_0(\hat{\mathbf{n}})\right)\varphi + \frac{1}{2}\partial_\varphi^2\left({}_{\pm}A_0(\hat{\mathbf{n}})\right)\varphi^2\cdots\,. \tag{2.36}$$

Inverting Eqs. (2.31) and (2.32), we find that

$$\Theta(\mathbf{l}) = \int d^2\hat{\mathbf{n}}\,\Theta(\hat{\mathbf{n}})e^{-i\mathbf{l}\cdot\hat{\mathbf{n}}}\,, \tag{2.37}$$

and

$$ {}_{\pm}A(\mathbf{l}) = -\int d^2\hat{\mathbf{n}}\,({}_{\pm}A(\hat{\mathbf{n}}))\,e^{\pm 2i(\phi-\phi_l)}e^{-i\mathbf{l}\cdot\hat{\mathbf{n}}}\,. \tag{2.38} $$

Thus, the φ expansions can be written in $\mathbf{l}$ space as

$$ \Theta(\mathbf{l}) = \Theta_0(\mathbf{l}) + \partial_\varphi\Theta_0(\mathbf{l})\varphi + \frac{1}{2}\partial_\varphi^2\Theta_0(\mathbf{l})\varphi^2 + \cdots\,, \tag{2.39} $$

$$ {}_{\pm}A(\mathbf{l}) = {}_{\pm}A_0(\mathbf{l}) + \partial_\varphi\left({}_{\pm}A_0(\mathbf{l})\right)\varphi + \frac{1}{2}\partial_\varphi^2\left({}_{\pm}A_0(\mathbf{l})\right)\varphi^2 + \cdots\,, \tag{2.40} $$

and in fact for any field $X \in \{\Theta, E, B\}$ we may write

$$ X(\mathbf{l}) = X_0(\mathbf{l}) + \partial_\varphi X_0(\mathbf{l})\varphi + \frac{1}{2}\partial_\varphi^2 X_0(\mathbf{l})\varphi^2 + \cdots\,. \tag{2.41} $$

To $O(\varphi^2)$ we then have,

$$ \begin{aligned} \langle X(\mathbf{l})\widetilde{X}(\mathbf{l}')\rangle = {} & \langle X_0(\mathbf{l})\widetilde{X}_0(\mathbf{l}')\rangle + \langle X_0(\mathbf{l})\partial_\varphi\widetilde{X}_0(\mathbf{l}')\rangle\varphi + \langle\partial_\varphi X_0(\mathbf{l})\widetilde{X}_0(\mathbf{l}')\rangle\varphi \\ & + \frac{1}{2}\langle X_0(\mathbf{l})\partial_\varphi^2\widetilde{X}_0(\mathbf{l}')\rangle\varphi^2 + \frac{1}{2}\langle\partial_\varphi^2 X_0(\mathbf{l})\widetilde{X}_0(\mathbf{l}')\rangle\varphi^2 + \langle\partial_\varphi X_0(\mathbf{l})\partial_\varphi\widetilde{X}_0(\mathbf{l}')\rangle\varphi^2, \end{aligned} \tag{2.42} $$

and in terms of power spectra this becomes

$$ C_l^{X\tilde{X}} = C_l^{X_0\tilde{X}_0} + \left(C_l^{X_0\partial\tilde{X}_0} + C_l^{\partial X_0\tilde{X}_0}\right)\varphi + \left(\frac{1}{2}C_l^{X_0\partial^2\tilde{X}_0} + \frac{1}{2}C_l^{\partial^2 X_0\tilde{X}_0} + C_l^{\partial X_0\partial\tilde{X}_0}\right)\varphi^2. \tag{2.43} $$

Since we are for the time being assuming a spatially uniform value of φ we may also write

$$ C_l^{X\tilde{X}} = C_l^{X\tilde{X}}\Big|_0 + \partial_\varphi C_l^{X\tilde{X}}\Big|_0\varphi + \frac{1}{2}\,\partial_\varphi^2 C_l^{X\tilde{X}}\Big|_0\varphi^2. \tag{2.44} $$

This allows us to make the identifications

$$ \begin{aligned} C_l^{X\tilde{X}}\Big|_0 &= C_l^{X_0\tilde{X}_0}\,, \\ \partial_\varphi C_l^{X\tilde{X}}\Big|_0 &= C_l^{X_0\partial\tilde{X}_0} + C_l^{\partial X_0\tilde{X}_0}\,, \\ \partial_\varphi^2 C_l^{X\tilde{X}}\Big|_0 &= C_l^{X_0\partial^2\tilde{X}_0} + C_l^{\partial^2 X_0\tilde{X}_0} + 2C_l^{\partial X_0\partial\tilde{X}_0}\,. \end{aligned} \tag{2.45} $$

These identifications make it clear how to calculate the individual 'derivative' power spectra. We just differentiate Eq. (2.30) with respect to φ, evaluate the expression at $\varphi = 0$, and pick off the terms with the requisite structure.

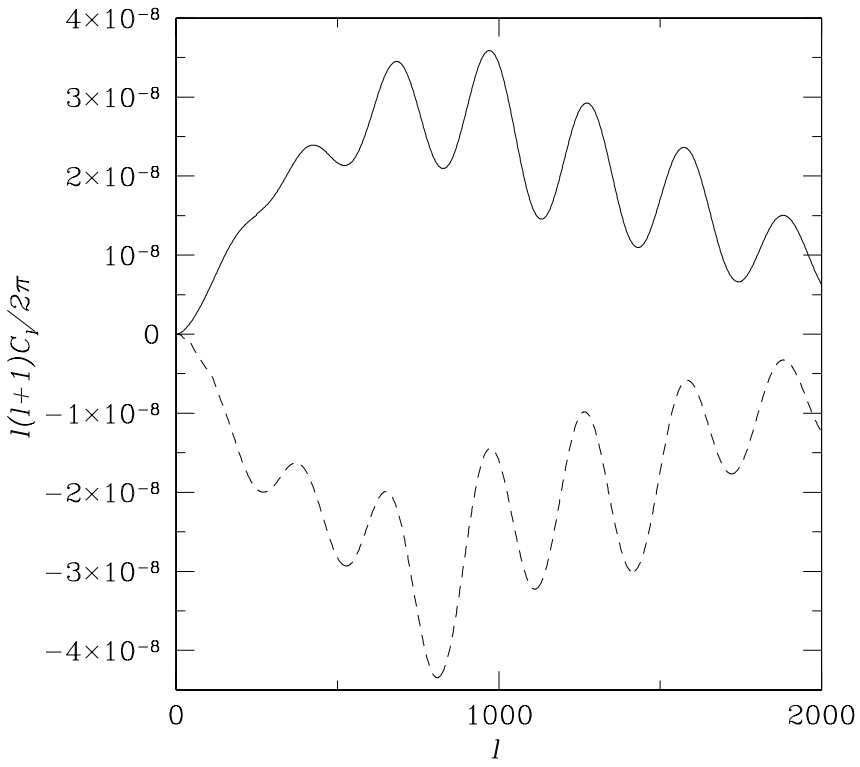

Figure 2.3: The $C_l^{\partial\Theta_0\partial\Theta_0}$ (solid) and $C_l^{\Theta_0\partial^2\Theta_0}$ (dashed) derivative power spectra.

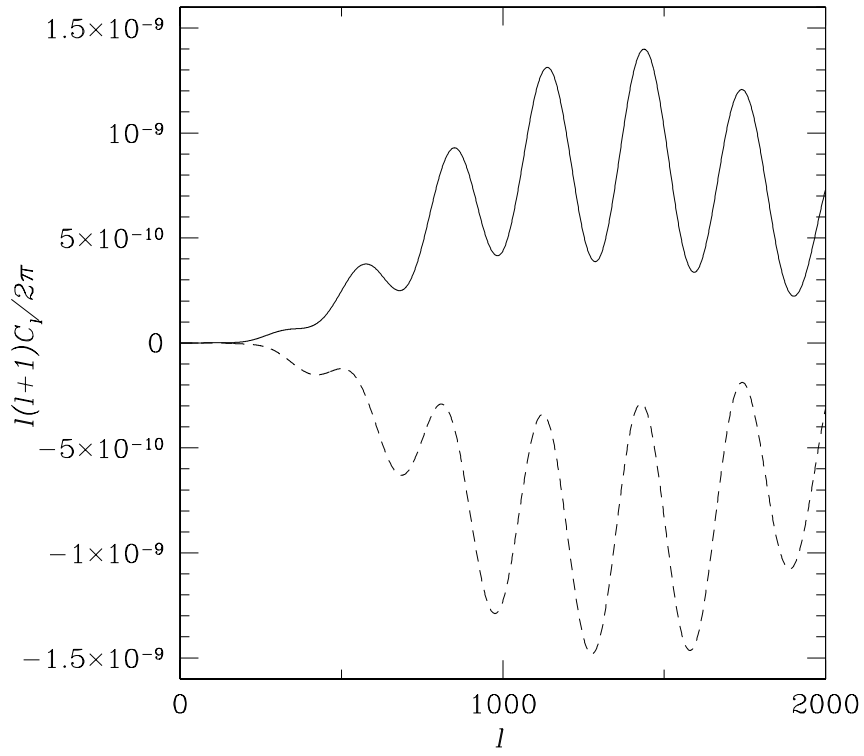

Figure 2.4: The $C_l^{\partial E_0 \partial E_0}$ (solid) and $C_l^{E_0 \partial^2 E_0}$ (dashed) derivative power spectra.

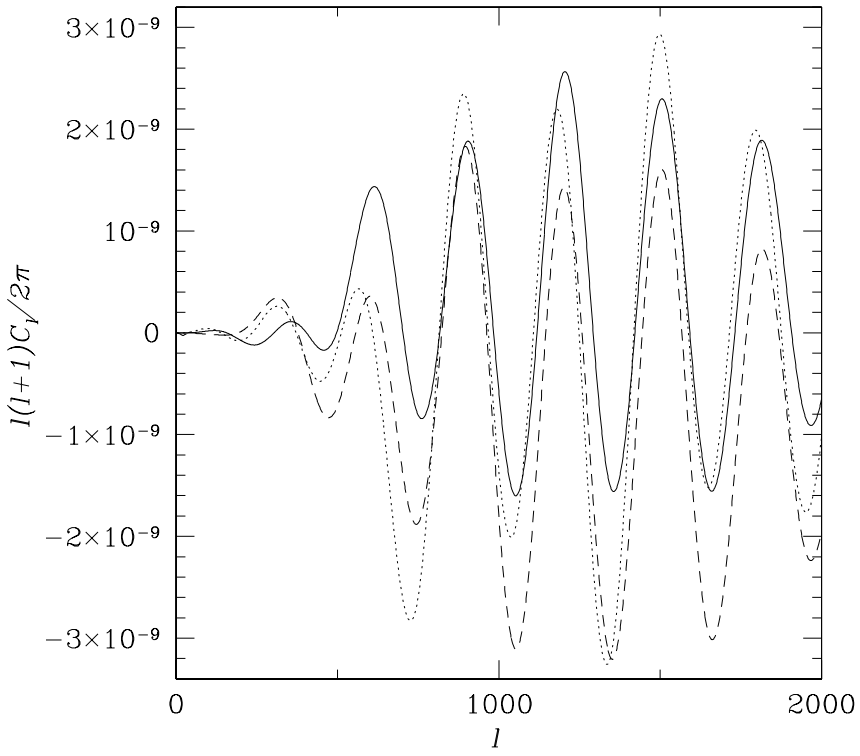

Figure 2.5: The $C_l^{\partial\Theta_0\partial E_0}$ (solid), $C_l^{\Theta_0\partial^2 E_0}$ (dashed) and $C_l^{E_0\partial^2\Theta_0}$ (dotted) derivative power spectra.

For example,

$$C_l^{\partial X_0 \partial \tilde{X}_0} = \frac{2}{\pi(2l+1)^2} \int \frac{dk}{k} \sum_{m=-2}^{2} k^3 \partial_\varphi X_l^{(m)*}(k) \partial_\varphi \widetilde{X}_l^{(m)}(k) \Big|_{\varphi=0} , \tag{2.46}$$

where factors like $\partial_\varphi \widetilde{X}_l^{(m)}(k)$ can be calculated numerically or directly from first principles using the expressions for the $\widetilde{X}_l^{(m)}(k)$ derived in, for instance, Ref. [2·45]. Because they are used in subsequent calculations, we have created a modified version of the code `CMBFAST` [2·46] that can compute these derivative power spectra. (See Figs. 2.3–2.5).

2.4 Spatial Variations of α

We now consider the effects of spatial variations of α (parametrized by φ) between different causally disconnected regions of the Universe.

First, suppose that there were *no* density fluctuations, but spatial variations of α. In that case, photons from different points on the sky would be scattered last at different cosmological times, but they would all still have the same frequency when observed by us. However, if there are density fluctuations, the manner in which they are imprinted on the CMB depends on the value of α, as discussed above. Thus, if there are spatial variations in α, the power spectra (or two-point correlation functions) will vary from one place on the sky to another. This implies that the stochasticity of the spatial variations in α induce non-Gaussianity in the CMB quantified by non-zero (connected) higher order correlation functions (trispectra and perhaps bispectra). It also implies a correction to the mean power spectrum (i.e., that measured by mapping regions of the sky that contain many coherence regions of α), as well as the introduction of a non-zero curl in the polarization. All of these effects are analogous to similar effects induced by weak lensing of the CMB. The only difference is that in our case, the temperature and polarization patterns are modulated by a variable α, rather than lensing by an intervening density field along the line of sight.

In this section, we first calculate the modified power spectra $C_l^{\Theta\Theta}$, $C_l^{\Theta E}$, C_l^{EE}, and C_l^{BB}. We then determine the form of the higher order correlations (bispectra and trispectra) in the next two sections.

2.4.1 Observable Modes in the Presence of φ Fluctuations

We assume that at a given position $\hat{\mathbf{n}}$ at the surface of last scatter, the value of α is $\alpha(\hat{\mathbf{n}}) = \alpha_0[1 + \varphi(\hat{\mathbf{n}})]$. Here we treat $\varphi(\hat{\mathbf{n}})$ as a random field with angular power spectrum $\langle \varphi(\mathbf{l})\varphi(\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')C_l^{\varphi\varphi}$ in the flat-sky approximation.

We assume that the surface of last scatter is much thinner than the spatial correlation length of

φ, and that in a given direction α is constant throughout recombination. We also assume that the dynamics responsible for the variations of φ have a negligible effect on the perturbation evolution, so that the sole effect of variations of φ are a modification of the microphysics. We will discuss the validity of these assumptions later.

Again we expand our fields $\Theta(\hat{\mathbf{n}}) = \Theta(\hat{\mathbf{n}};\varphi)$ and ${}_{\pm}A(\hat{\mathbf{n}}) = {}_{\pm}A(\hat{\mathbf{n}};\varphi)$ in a Taylor series about $\varphi = 0$ as

$$\Theta(\hat{\mathbf{n}}) = \Theta_0(\hat{\mathbf{n}}) + \partial_\varphi\Theta_0(\hat{\mathbf{n}})\varphi(\hat{\mathbf{n}}) + \frac{1}{2}\partial_\varphi^2\Theta_0(\hat{\mathbf{n}})\varphi^2(\hat{\mathbf{n}}), \tag{2.47}$$

$${}_{\pm}A(\hat{\mathbf{n}}) = {}_{\pm}A_0(\hat{\mathbf{n}}) + \partial_\varphi\left({}_{\pm}A_0(\hat{\mathbf{n}})\right)\varphi(\hat{\mathbf{n}}) + \frac{1}{2}\partial_\varphi^2\left({}_{\pm}A_0(\hat{\mathbf{n}})\right)\varphi^2(\hat{\mathbf{n}}), \tag{2.48}$$

where $\varphi(\hat{\mathbf{n}})$ is now a function of position.

By taking the Fourier transform of $\Theta(\hat{\mathbf{n}})$ we find

$$\Theta(\mathbf{l}) = \Theta_0(\mathbf{l}) + \left[(\partial_\varphi\Theta_0)\star\varphi\right](\mathbf{l}) + \frac{1}{2}\left[(\partial_\varphi^2\Theta_0)\star\varphi\star\varphi\right](\mathbf{l}), \tag{2.49}$$

where

$$\left[X\star\psi\right](\mathbf{l}) = \int\frac{d^2\mathbf{l}'}{(2\pi)^2}X(\mathbf{l}')\psi(\mathbf{l}-\mathbf{l}') \tag{2.50}$$

is the convolution of two fields X and ψ, and

$$\left[X\star\psi\star\lambda\right](\mathbf{l}) = \int\frac{d^2\mathbf{l}'}{(2\pi)^2}\left[X\star\psi\right](\mathbf{l}')\lambda(\mathbf{l}-\mathbf{l}'), \tag{2.51}$$

is the double convolution of three fields X, ψ, and λ.

Similarly, taking linear combinations of the Fourier transforms of ${}_{\pm}A(\hat{\mathbf{n}})$ we find that

$$\begin{aligned}E(\mathbf{l}) = E_0(\mathbf{l}) &+ \left[(\partial_\varphi E_0)\star_{\mathbf{c}}\varphi\right](\mathbf{l}) - \left[(\partial_\varphi B_0)\star_{\mathbf{s}}\varphi\right](\mathbf{l})\\ &+ \frac{1}{2}\left[(\partial_\varphi^2 E_0)\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l}) - \frac{1}{2}\left[(\partial_\varphi^2 B_0)\star_{\mathbf{s}}\varphi\star\varphi\right](\mathbf{l}),\end{aligned} \tag{2.52}$$

and

$$\begin{aligned}B(\mathbf{l}) = B_0(\mathbf{l}) &+ \left[(\partial_\varphi B_0)\star_{\mathbf{c}}\varphi\right](\mathbf{l}) + \left[(\partial_\varphi E_0)\star_{\mathbf{s}}\varphi\right](\mathbf{l})\\ &+ \frac{1}{2}\left[(\partial_\varphi^2 B_0)\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l}) + \frac{1}{2}\left[(\partial_\varphi^2 E_0)\star_{\mathbf{s}}\varphi\star\varphi\right](\mathbf{l}),\end{aligned} \tag{2.53}$$

where

$$\left[X\star_{\mathbf{c}}\psi\right](\mathbf{l}) = \int\frac{d^2\mathbf{l}'}{(2\pi)^2}\cos\left(2\phi_{l'}\right)X(\mathbf{l}')\psi(\mathbf{l}-\mathbf{l}'), \tag{2.54}$$

and

$$[X\star_{\mathbf{s}}\psi]\,(\mathbf{l}) = \int \frac{d^2\mathbf{l}'}{(2\pi)^2}\sin\left(2\phi_{l'}\right)X(\mathbf{l}')\psi(\mathbf{l}-\mathbf{l}')\,, \tag{2.55}$$

are the even- and odd-parity spin-2 weighted convolutions of X and ψ respectively.

Examining these expressions we find that a given mode $X(\mathbf{l})$ receives corrections due to the combination of modes $\{\partial_\varphi^n X_0(\mathbf{l}_0), \varphi(\mathbf{l}_1), \varphi(\mathbf{l}_2), \ldots, \varphi(\mathbf{l}_n)\}$ such that $\sum_{i=0}^n \mathbf{l}_i = \mathbf{l}$. Furthermore, the E and B modes mix so that, for example, the mode $B(\mathbf{l})$ can be induced by the combinations of modes $\{\partial_\varphi^n E_0(\mathbf{l}_0), \varphi(\mathbf{l}_1), \varphi(\mathbf{l}_2), \ldots, \varphi(\mathbf{l}_n)\}$ such that $\sum_{i=0}^n \mathbf{l}_i = \mathbf{l}$. These effects modify the angular power spectra of CMB anisotropies, and introduce higher-order connected (non-Gaussian) correlation functions.

2.4.2 The $\Theta\Theta$ Power Spectrum

Using Eq. (2.49) we find that the expansion for the two point correlation function (in Fourier space) is

$$\begin{aligned}\langle\Theta(\mathbf{l})\Theta(\mathbf{l}')\rangle &= \langle\Theta_0(\mathbf{l})\Theta_0(\mathbf{l}')\rangle + \langle\Theta_0(\mathbf{l})\left[\partial_\varphi\Theta_0\star\varphi\right](\mathbf{l}')\rangle + \langle\left[\partial_\varphi\Theta_0\star\varphi\right](\mathbf{l})\Theta_0(\mathbf{l}')\rangle \\ &+ \frac{1}{2}\langle\Theta_0(\mathbf{l})\left[\partial_\varphi^2\Theta_0\star\varphi\star\varphi\right](\mathbf{l}')\rangle + \frac{1}{2}\langle\left[\partial_\varphi^2\Theta_0\star\varphi\star\varphi\right](\mathbf{l})\Theta_0(\mathbf{l}')\rangle + \langle\left[\partial_\varphi\Theta_0\star\varphi\right](\mathbf{l})\left[\partial_\varphi\Theta_0\star\varphi\right](\mathbf{l}')\rangle\end{aligned} \tag{2.56}$$

to $O(\varphi^2)$. In the above expression and what follows we adopt the convention that the differential operators ∂_φ act only the the field immediately following them.

We assume that Θ_0 and φ are zero-mean Gaussian random fields without higher-order connected correlators. By writing out the convolutions and Wick expanding the correlators it is easy to verify that correlators involving an odd number of fields vanish, and so there are no corrections to first order in φ. It is also straightforward to verify that

$$\begin{aligned}\langle\Theta_0(\mathbf{l})[\partial_\varphi^2\Theta_0\star\varphi\star\varphi](\mathbf{l}')\rangle &= \langle[\partial_\varphi^2\Theta_0\star\varphi\star\varphi](\mathbf{l})\Theta_0(\mathbf{l}')\rangle \\ &= (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\left[\sigma^{(\varphi\varphi)}C_l^{\Theta_0\partial^2\Theta_0} + 2\sigma^{(\varphi\partial^2\Theta_0)}C_l^{\Theta_0\varphi}\right]\,,\end{aligned} \tag{2.57}$$

where

$$\sigma^{(\psi\lambda)} = \int\frac{d^2\mathbf{l}}{(2\pi)^2}C_l^{\psi\lambda} \tag{2.58}$$

is the covariance between two fields ψ and λ. Similarly we can show that

$$\langle\left[\partial_\varphi\Theta_0\star\varphi\right](\mathbf{l})\left[\partial_\varphi\Theta_0\star\varphi\right](\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\{\left[C^{\varphi\varphi}\star C^{\partial\Theta_0\partial\Theta_0}\right]_l + \left[C^{\partial\Theta_0\varphi}\star C^{\partial\Theta_0\varphi}\right]_l\}\,, \tag{2.59}$$

where we have dropped terms that contribute only when $\mathbf{l}=0$.

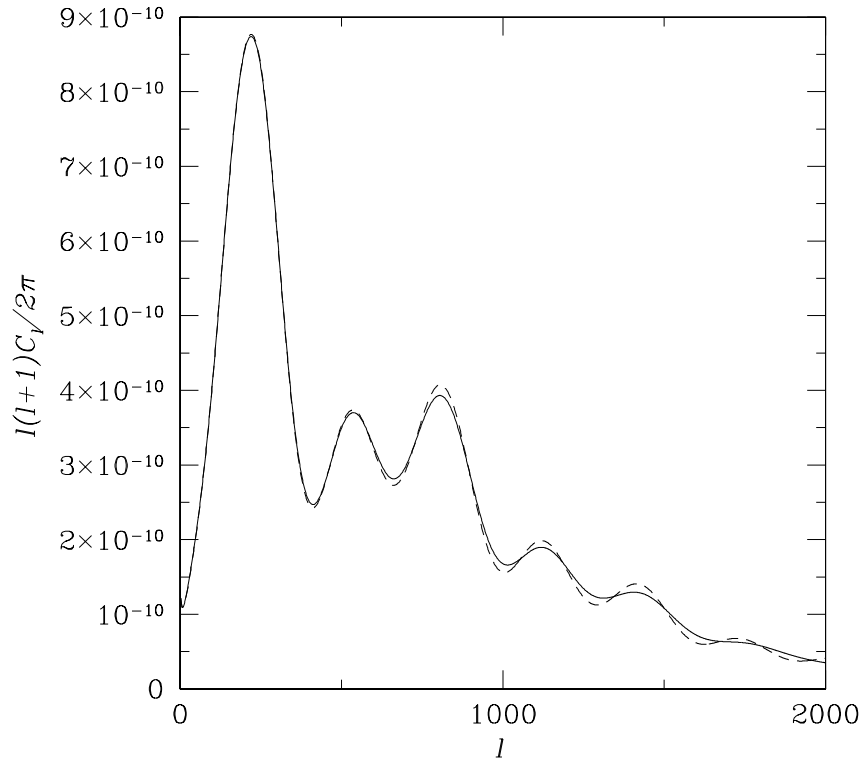

Figure 2.6: The solid curve shows the $\Theta\Theta$ power spectrum for the Gaussian correlation function for φ with angular correlation scale $\theta_c = 1^\circ$ and variance $\sigma^{(\varphi\varphi)} = 9 \times 10^{-4}$ (fluctuations in φ at the 3% level). The dashed curve shows the power spectrum without fluctuations in φ.

Collecting all terms we find that to leading order the average power spectrum including fluctuations in φ is

$$C_l^{\Theta\Theta} = C_l^{\Theta_0\Theta_0} + \sigma^{(\varphi\varphi)} C_l^{\Theta_0\partial^2\Theta_0} + 2\sigma^{(\varphi\partial^2\Theta_0)} C_l^{\Theta_0\varphi} + \left[C^{\varphi\varphi} \star C^{\partial\Theta_0\partial\Theta_0}\right]_l + \left[C^{\partial\Theta_0\varphi} \star C^{\partial\Theta_0\varphi}\right]_l . \quad (2.60)$$

Note that the corrections to $C_l^{\Theta_0\Theta_0}$ involve couplings between the derivative power spectra, which are calculated as described in Section 2.3.2, and $C_l^{\varphi\varphi}$ and the various cross power spectra, which are specified by the model that generates spatial variations in φ.

In general Θ_0 and φ may be correlated if, for instance, they are generated by a common mechanism or if they are strongly coupled through evolution equations. We discuss in Section 2.7 why we do not expect the latter source of correlations to be important as long as the energy in the φ field and its fluctuations are small compared to the dominant radiation and matter perturbations. In the

case where Θ_0 and φ have no cross-correlation, the expression simplifies to

$$C_l^{\Theta\Theta} = C_l^{\Theta_0\Theta_0} + \sigma^{(\varphi\varphi)} C_l^{\Theta_0\partial^2\Theta_0} + \left[C^{\varphi\varphi} \star C^{\partial\Theta_0\partial\Theta_0}\right]_l . \tag{2.61}$$

The result for $C_l^{\Theta\Theta}$ will of course depend on $C_l^{\varphi\varphi}$. To illustrate we consider a simple model in which φ is highly correlated on angular scales smaller than the correlation angle θ_c, and uncorrelated on larger scales. (Below we discuss a physical model that may produce such a correlation function). We thus have,

$$\langle\varphi(0)\varphi(\theta)\rangle = \sigma^{(\varphi\varphi)} e^{-(\theta/\theta_c)^2} . \tag{2.62}$$

In the flat-sky approximation we have

$$C_l^{\varphi\varphi} = \int d^2\theta \, \langle\varphi(0)\varphi(\theta)\rangle e^{-i\mathbf{l}\cdot\boldsymbol{\theta}} , \tag{2.63}$$

which implies that

$$C_l^{\varphi\varphi} = \pi\theta_c^2 \sigma^{(\varphi\varphi)} e^{-\frac{1}{4}l^2\theta_c^2} . \tag{2.64}$$

In this case the average $\Theta\Theta$ power spectrum is

$$C_l^{\Theta\Theta} = C_l^{\Theta_0\Theta_0} + \sigma^{(\varphi\varphi)} \left[C_l^{\Theta_0\partial^2\Theta_0} + \frac{\theta_c^2}{2} \int dl' l' e^{-\frac{1}{4}(l^2+l'^2)\theta_c^2} I_0\left(\frac{\theta_c^2}{2} l l'\right) C_{l'}^{\partial\Theta_0\partial\Theta_0}\right] , \tag{2.65}$$

where I_n is the n^{th}-order modified Bessel function of the first kind. We show this average power spectrum in Fig. 2.6. The main effect of φ fluctuations is to reduce the amplitude of oscillatory features in the damping tail. This effect can be understood by noting that patches of the sky with different values of φ will have different power spectra, and that, as can be seen in Fig. 4.13, the location of of the peaks in the damping tail of these power spectra shift as φ changes. These patch power spectra add incoherently, so that the amplitude of the oscillatory component of the average power spectrum is reduced. This is the same type of effect as in weak gravitational lensing [2·47]. As we will see, this is the predominant effect in the other power spectra as well.

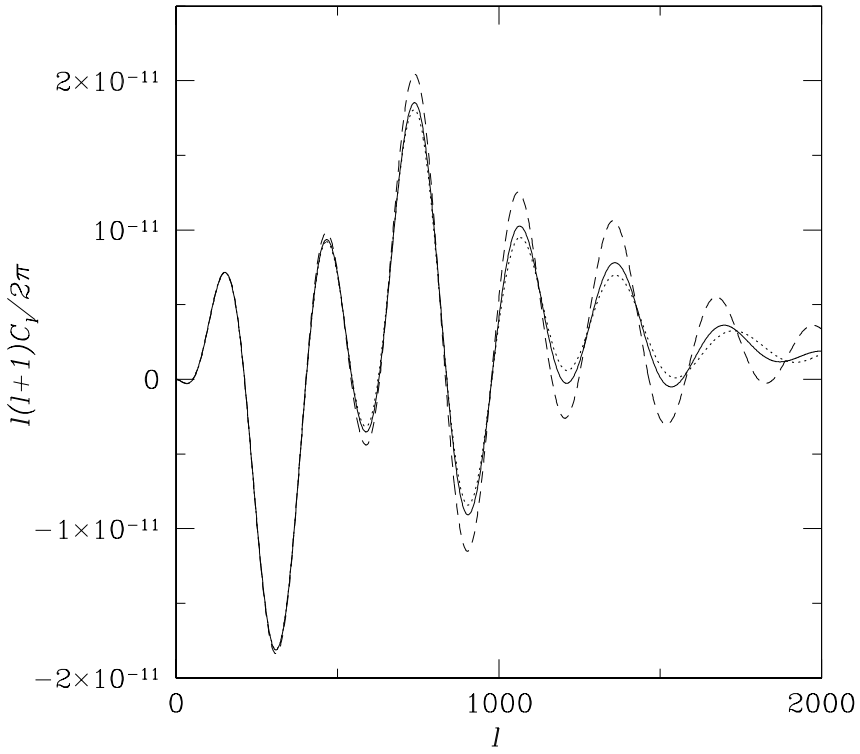

Figure 2.7: The ΘE power spectra for the Gaussian correlation function for φ with $\sigma^{(\varphi\varphi)} = 9 \times 10^{-4}$ and $\theta_c = 1^\circ$ (solid) and $\theta_c = 2^\circ$ (dotted). The dashed line shows the power spectrum without fluctuations in φ.

2.4.3 The ΘE Power Spectrum

Using Eqs. (2.49) and (2.52), and immediately dropping the vanishing correlators involving an odd number of fields, we find that the expansion for the power spectrum is

$$\begin{aligned}\langle\Theta(\mathbf{l})E(\mathbf{l}')\rangle = \langle\Theta_0(\mathbf{l})E_0(\mathbf{l}')\rangle &+ \frac{1}{2}\langle\Theta_0(\mathbf{l})\left[\partial_\varphi^2 E_0 \star_{\mathbf{c}} \varphi\star\varphi\right](\mathbf{l}')\rangle - \frac{1}{2}\langle\Theta_0(\mathbf{l})\left[\partial_\varphi^2 B_0 \star_{\mathbf{s}} \varphi\star\varphi\right](\mathbf{l}')\rangle \\ &+ \langle[\partial_\varphi\Theta_0\star\varphi](\mathbf{l})\left[\partial_\varphi E_0\star_{\mathbf{c}}\varphi\right](\mathbf{l}')\rangle - \langle[\partial_\varphi\Theta_0\star\varphi](\mathbf{l})\left[\partial_\varphi B_0\star_{\mathbf{s}}\varphi\right](\mathbf{l}')\rangle \\ &+ \frac{1}{2}\langle\left[\partial_\varphi^2\Theta_0\star\varphi\star\varphi\right](\mathbf{l})E_0(\mathbf{l}')\rangle \end{aligned} \tag{2.66}$$

to $O(\varphi^2)$.

By writing out the convolutions, Wick expanding the correlators, and noting that terms involving B_0 vanish due to parity, we find that the non-vanishing terms are

$$\langle\Theta_0(\mathbf{l})\left[\partial_\varphi^2 E_0\star\varphi\star\varphi\right](\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\left[\sigma^{(\varphi\varphi)}C_l^{\Theta_0\partial^2 E_0} + 2\sigma^{(\varphi\partial^2 E_0)}C_l^{\Theta_0\varphi}\right], \tag{2.67}$$

$$\langle\left[\partial_\varphi^2\Theta_0\star\varphi\star\varphi\right](\mathbf{l})E_0(\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\left[\sigma^{(\varphi\varphi)}C_l^{E_0\partial^2\Theta_0} + 2\sigma^{(\varphi\partial^2\Theta_0)}C_l^{E_0\varphi}\right], \tag{2.68}$$

$$\langle[\partial_\varphi\Theta_0\star\varphi](\mathbf{l})\,[\partial_\varphi E_0\star_{\mathbf{c}}\varphi](\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\{\left[C^{\partial\Theta_0\partial E_0}\star_{\mathbf{c}}C^{\varphi\varphi}\right]_l + \left[C^{\partial E_0\varphi}\star_{\mathbf{c}}C^{\partial\Theta_0\varphi}\right]_l\}. \tag{2.69}$$

Collecting these terms we find that, to leading order, the average cross power-spectrum including fluctuations in φ is

$$\begin{aligned}C_l^{\Theta E} = C_l^{\Theta_0 E_0} + \sigma^{(\varphi\varphi)}\left[\frac{1}{2}C_l^{\Theta_0\partial^2 E_0} + \frac{1}{2}C_l^{E_0\partial^2\Theta_0}\right] + \sigma^{(\varphi\partial^2 E_0)}C_l^{\Theta_0\varphi} + \sigma^{(\varphi\partial^2\Theta_0)}C_l^{E_0\varphi}\\ + \left[C^{\partial\Theta_0\partial E_0}\star_{\mathbf{c}}C^{\varphi\varphi}\right]_l + \left[C^{\partial E_0\varphi}\star_{\mathbf{c}}C^{\partial\Theta_0\varphi}\right]_l.\end{aligned} \tag{2.70}$$

If φ has no correlation with the density field, the expression simplifies to

$$C_l^{\Theta E} = C_l^{\Theta_0 E_0} + \sigma^{(\varphi\varphi)}\left[\frac{1}{2}C_l^{\Theta_0\partial^2 E_0} + \frac{1}{2}C_l^{E_0\partial^2\Theta_0}\right] + \left[C^{\partial\Theta_0\partial E_0}\star_{\mathbf{c}}C^{\varphi\varphi}\right]_l. \tag{2.71}$$

Inserting the power spectrum from Eq. (2.64), we obtain the expression

$$\begin{aligned}C_l^{\Theta E} = C_l^{\Theta_0 E_0} + \sigma^{(\varphi\varphi)}\left[\frac{1}{2}C_l^{\Theta_0\partial^2 E_0} + \frac{1}{2}C_l^{E_0\partial^2\Theta_0}\right.\\ \left. + \frac{\theta_c^2}{2}\int dl'l'e^{-\frac{1}{4}(l^2+l'^2)\theta_c^2}I_2\left(\frac{\theta_c^2}{2}ll'\right)C_{l'}^{\partial\Theta_0\partial\Theta_0}\right].\end{aligned} \tag{2.72}$$

We show this average power spectrum in Fig 2.7. The major effect of φ fluctuations on the ΘE power spectrum is to reduce the peak amplitudes on small scales.

2.4.4 The EE Power Spectrum

Using Eq. (2.52), and dropping the correlators involving an odd number of fields and those that vanish due to parity, we find that the expansion for the power spectrum is

$$\begin{aligned}\langle E(\mathbf{l})E(\mathbf{l}')\rangle = \langle E_0(\mathbf{l})E_0(\mathbf{l}')\rangle + \frac{1}{2}\langle E_0(\mathbf{l})\left[\partial_\varphi^2 E_0\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l}')\rangle + \frac{1}{2}\langle\left[\partial_\varphi^2 E_0\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l})E_0(\mathbf{l}')\rangle\\ + \langle[\partial_\varphi E_0\star_{\mathbf{c}}\varphi](\mathbf{l})\,[\partial_\varphi E_0\star_{\mathbf{c}}\varphi](\mathbf{l}')\rangle + \langle[\partial_\varphi B_0\star_{\mathbf{s}}\varphi](\mathbf{l})\,[\partial_\varphi B_0\star_{\mathbf{s}}\varphi](\mathbf{l}')\rangle\end{aligned} \tag{2.73}$$

to $O(\varphi^2)$.

After evaluating the convolutions, and Wick expanding these correlators, we find

$$\begin{aligned}\langle E_0(\mathbf{l})\left[\partial_\varphi^2 E_0\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l}')\rangle &= \langle\left[\partial_\varphi^2 E_0\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l}')E_0(\mathbf{l})\rangle\\ &= (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\left[\sigma^{(\varphi\varphi)}C_l^{E_0\partial^2 E_0} + 2\sigma^{(\varphi\partial^2 E_0)}C_l^{E_0\varphi}\right],\end{aligned} \tag{2.74}$$

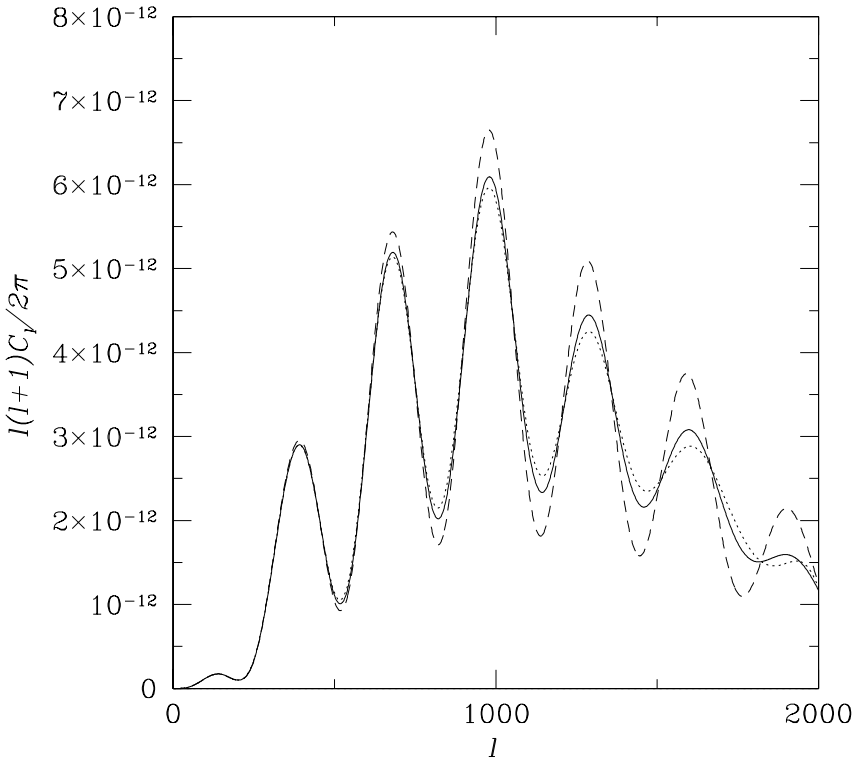

Figure 2.8: The EE power spectra for the Gaussian correlation function for φ with $\sigma^{(\varphi\varphi)} = 9 \times 10^{-4}$ and $\theta_c = 1^\circ$ (solid) and $\theta_c = 2^\circ$ (dotted). The dashed line shows the power spectrum without fluctuations in φ.

$$\langle [\partial_\varphi E_0 \star_{\mathbf{c}} \varphi]\,(\mathbf{l})\,[\partial_\varphi E_0 \star_{\mathbf{c}} \varphi]\,(\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\int \frac{d^2\mathbf{l}_1}{(2\pi)^2}\left\{\cos^2(2\phi_{l_1})C_{l_1}^{\partial E_0 \partial E_0}C_{|\mathbf{l}-\mathbf{l}_1|}^{\varphi\varphi} + \cos(2\phi_{l_1})\left[\frac{2{l_1}^2\sin^2(\phi_{l_1})}{|\mathbf{l}-\mathbf{l}_1|^2}-1\right]C_{l_1}^{\partial E_0\varphi}C_{|\mathbf{l}-\mathbf{l}_1|}^{\partial E_0\varphi}\right\}, \tag{2.75}$$

$$\langle\,[\partial_\varphi B_0 \star_{\mathbf{s}} \varphi]\,(\mathbf{l})\,[\partial_\varphi B_0 \star_{\mathbf{s}} \varphi]\,(\mathbf{l}')\rangle = (2\pi)^2\delta^2(\mathbf{l}+\mathbf{l}')\int \frac{d^2\mathbf{l}_1}{(2\pi)^2}\sin^2(2\phi_{l_1})C_{l_1}^{\partial B_0\partial B_0}C_{|\mathbf{l}-\mathbf{l}_1|}^{\varphi\varphi}. \tag{2.76}$$

Collecting these terms we find that

$$C_l^{EE} = C_l^{E_0E_0} + \sigma^{(\varphi\varphi)}C_l^{E_0\partial^2E_0} + 2\sigma^{(\varphi\partial^2E_0)}C_l^{E_0\varphi} + \int\frac{d^2\mathbf{l}'}{(2\pi)^2}\left\{[\cos^2(2\phi_{l'})C_{l'}^{\partial E_0\partial E_0} + \sin^2(2\phi_{l'})C_{l'}^{\partial B_0\partial B_0}]C_{|\mathbf{l}-\mathbf{l}'|}^{\varphi\varphi} + \cos(2\phi_{l'})\left[\frac{2{l'}^2\sin^2(\phi_{l'})}{|\mathbf{l}-\mathbf{l}'|^2}-1\right]C_{l'}^{\partial E_0\varphi}C_{|\mathbf{l}-\mathbf{l}'|}^{\partial E_0\varphi}\right\}, \tag{2.77}$$

is the expression for the average E-mode power spectrum including fluctuations in φ.

If E_0 and φ are uncorrelated, the expression simplifies to

$$C_l^{EE} = C_l^{E_0E_0} + \sigma^{(\varphi\varphi)}C_l^{E_0\partial^2E_0} + \int\frac{d^2\mathbf{l}'}{(2\pi)^2}[\cos^2(2\phi_{l'})C_{l'}^{\partial E_0\partial E_0} + \sin^2(2\phi_{l'})C_{l'}^{\partial B_0\partial B_0}]C_{|\mathbf{l}-\mathbf{l}'|}^{\varphi\varphi}. \tag{2.78}$$

For the φ power spectrum in Eq (2.64), dropping the negligible primordial B-mode term, we obtain the expression

$$C_l^{EE} = C_l^{E_0E_0} + \sigma^{(\varphi\varphi)}\left\{C_l^{E_0\partial^2E_0} + \frac{\theta_c^2}{4}\int dl'l'e^{-\frac{1}{4}(l^2+{l'}^2)\theta_c^2}\left[I_0\left(\frac{\theta_c^2}{2}ll'\right)+I_4\left(\frac{\theta_c^2}{2}ll'\right)\right]C_{l'}^{\partial E_0\partial E_0}\right\}. \tag{2.79}$$

We plot this average power spectrum in Fig. 2.8. Again, the effect of φ fluctuations is to reduce the amplitude of the oscillatory component at small angular scales.

2.4.5 The BB Power Spectrum

Using Eq. (2.53), and once again dropping the correlators involving an odd number of fields and those that vanish due to parity, we find that the expansion for the two point correlation function is

$$\langle B(\mathbf{l})B(\mathbf{l}')\rangle = \langle B_0(\mathbf{l})B_0(\mathbf{l}')\rangle + \frac{1}{2}\langle B_0(\mathbf{l})\left[\partial_\varphi^2 B_0\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l}')\rangle + \frac{1}{2}\langle\left[\partial_\varphi^2 B_0\star_{\mathbf{c}}\varphi\star\varphi\right](\mathbf{l})B_0(\mathbf{l}')\rangle + \langle[\partial_\varphi B_0\star_{\mathbf{c}}\varphi]\,(\mathbf{l})\,[\partial_\varphi B_0\star_{\mathbf{c}}\varphi]\,(\mathbf{l}')\rangle + \langle[\partial_\varphi E_0\star_{\mathbf{s}}\varphi]\,(\mathbf{l})\,[\partial_\varphi E_0\star_{\mathbf{s}}\varphi]\,(\mathbf{l}')\rangle \tag{2.80}$$

to $O(\varphi^2)$.

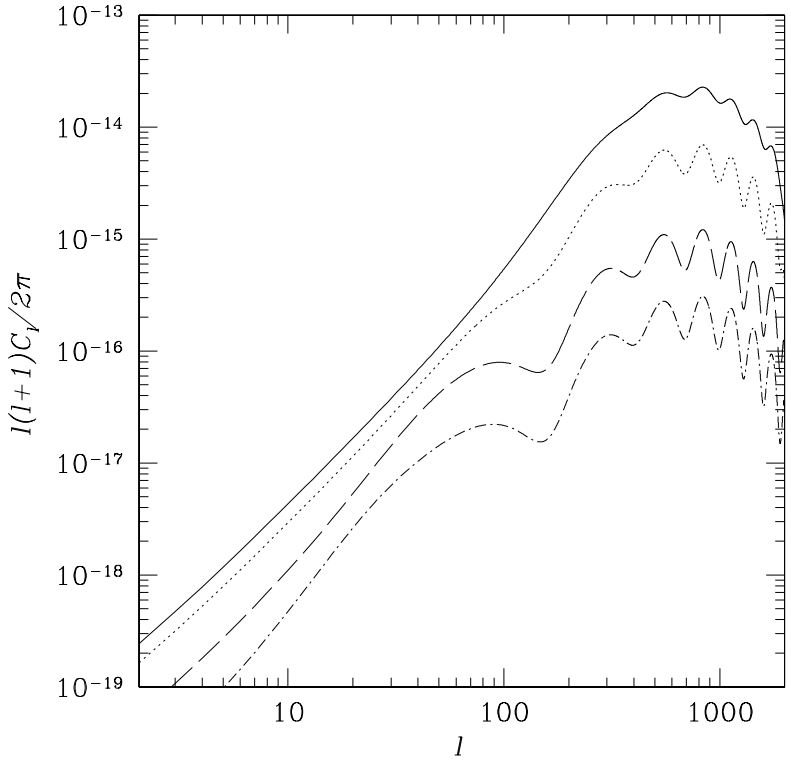

Figure 2.9: The BB power spectra for the Gaussian correlation function for φ with $\sigma^{(\varphi\varphi)} = 9 \times 10^{-4}$ and $\theta_c = 1^\circ$ (solid), $\theta_c = 2^\circ$ (dotted), $\theta_c = 5^\circ$ (dashed), and $\theta_c = 10^\circ$ (dot-dashed).

Evaluating the convolutions, and Wick expanding these correlators, leads to the expressions

$$\langle B_0(\mathbf{l}) \left[\partial_\varphi^2 B_0 \star_{\mathbf{c}} \varphi \star \varphi\right](\mathbf{l}')\rangle = \langle \left[\partial_\varphi^2 B_0 \star_{\mathbf{c}} \varphi \star \varphi\right](\mathbf{l}') B_0(\mathbf{l})\rangle = (2\pi)^2 \delta^2(\mathbf{l}+\mathbf{l}') \left[\sigma^{(\varphi\varphi)} C_l^{B_0 \partial^2 B_0}\right], \quad (2.81)$$

$$\langle \left[\partial_\varphi B_0 \star_{\mathbf{c}} \varphi\right](\mathbf{l}) \left[\partial_\varphi B_0 \star_{\mathbf{c}} \varphi\right](\mathbf{l}')\rangle = (2\pi)^2 \delta^2(\mathbf{l}+\mathbf{l}') \int \frac{d^2\mathbf{l}_1}{(2\pi)^2} \cos^2(2\phi_{l_1}) C_{l_1}^{\partial B_0 \partial B_0} C_{|\mathbf{l}-\mathbf{l}_1|}^{\varphi\varphi}, \quad (2.82)$$

$$\begin{aligned}\langle \left[\partial_\varphi E_0 \star_{\mathbf{s}} \varphi\right](\mathbf{l}) \left[\partial_\varphi E_0 \star_{\mathbf{s}} \varphi\right](\mathbf{l}')\rangle = (2\pi)^2 \delta^2(\mathbf{l}+\mathbf{l}') \int \frac{d^2\mathbf{l}_1}{(2\pi)^2} \Big\{ &\sin^2(2\phi_{l_1}) C_{l_1}^{\partial E_0 \partial E_0} C_{|\mathbf{l}-\mathbf{l}_1|}^{\varphi\varphi} \\ &+ \sin(2\phi_{l_1}) \left[\frac{l_1^2 \sin(2\phi_{l_1}) - 2 l l_1 \sin(\phi_{l_1})}{|\mathbf{l}-\mathbf{l}_1|^2}\right] \times C_{l_1}^{\partial E_0 \varphi} C_{|\mathbf{l}-\mathbf{l}_1|}^{\partial E_0 \varphi} \Big\}. \end{aligned} \quad (2.83)$$

Any φ-B_0 cross correlations must vanish due to parity.

Collecting all terms we find that the expression for the average B-mode power spectrum is

$$C_l^{BB} = C_l^{B_0B_0} + \sigma^{(\varphi\varphi)} C_l^{B_0\partial^2 B_0} + \int \frac{d^2\mathbf{l}'}{(2\pi)^2} \left\{ \left[\cos^2(2\phi_{l'}) C_{l'}^{\partial B_0 \partial B_0} + \sin^2(2\phi_{l'}) C_{l'}^{\partial E_0 \partial E_0} \right] C^{\varphi\varphi}_{|\mathbf{l}-\mathbf{l}'|} + \sin(2\phi_{l'}) \left[\frac{l'^2 \sin(2\phi_{l'}) - 2ll' \sin(\phi_{l'})}{|\mathbf{l}-\mathbf{l}'|^2} \right] C_{l'}^{\partial B_0 \varphi} C^{\partial B_0 \varphi}_{|\mathbf{l}-\mathbf{l}'|} \right\}, \quad (2.84)$$

and if E_0 and φ are uncorrelated the expression simplifies to

$$C_l^{BB} = C_l^{B_0B_0} + \sigma^{(\varphi\varphi)} C_l^{B_0\partial^2 B_0} + \int \frac{d^2\mathbf{l}'}{(2\pi)^2} [\cos^2(2\phi_{l'}) C_{l'}^{\partial B_0 \partial B_0} + \sin^2(2\phi_{l'}) C_{l'}^{\partial E_0 \partial E_0}] C^{\varphi\varphi}_{|\mathbf{l}-\mathbf{l}'|}. \quad (2.85)$$

If the intrinsic B-modes due to gravitational waves are negligible then for our φ power spectrum, Eq. (2.64), the average B-mode power spectrum is

$$C_l^{BB} = \sigma^{(\varphi\varphi)} \left\{ \frac{\theta_c^2}{4} \int dl' l' e^{-\frac{1}{4}(l^2+l'^2)\theta_c^2} \left[I_0\left(\frac{\theta_c^2}{2} ll'\right) - I_4\left(\frac{\theta_c^2}{2} ll'\right) \right] C_{l'}^{\partial E_0 \partial E_0} \right\}. \quad (2.86)$$

This expression shows that E modes modulated by φ fluctuations can induce B modes. We note here that this effect is more general than the specific application of α variation we focus on here, and is a generic feature of any fluctuations that modulate the power spectrum but have negligible effects on the evolution of the dominant density perturbations. In Fig. 2.9 we plot the induced B-mode power spectra due to φ fluctuations. The induced B-mode power spectrum inherits the oscillatory features of the unperturbed E_0-mode power spectrum as long as correlation angle is larger than the horizon size at recombination, while the amplitude of the power spectrum decreases as the correlation angle increases.

2.5 Bispectra

We have shown in the previous section that modulation of the temperature and polarization power spectra by spatial variations of α alters the mean power spectra of CMB anisotropies. In this section and the one that follows we show how this modulation introduces higher order correlations into the CMB temperature field.

The temperature bispectrum is defined in terms of the connected piece (the terms remaining after the Gaussian piece is subtracted out) of the three-point correlation function in Fourier space as

$$\langle \Theta(\mathbf{l}_1)\Theta(\mathbf{l}_2)\Theta(\mathbf{l}_3)\rangle_c = (2\pi)^2 \delta^2(\mathbf{l}_1 + \mathbf{l}_2 + \mathbf{l}_3) B^{\Theta\Theta\Theta}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3). \quad (2.87)$$

This expression must be invariant under the exchange of any two fields, or equivalently, $\mathbf{l}$ vectors. We insert Eq. (2.49) into the three-point correlation function, and after some straightforward algebra

we find that

$$B^{\Theta\Theta\Theta}(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3) = \sum_{\substack{i,j=1\\ i<j}}^{3} \mathrm{B}(\mathbf{l}_i,\mathbf{l}_j,\mathbf{l}_{6-i-j})\,, \tag{2.88}$$

where the sum runs over the distinct permutations of $\{1,2,3\}$ and, to leading order,

$$\mathrm{B}(\mathbf{l}_i,\mathbf{l}_j,\mathbf{l}_k) = C_{l_i}^{\Theta_0\partial\Theta_0} C_{l_j}^{\Theta_0\varphi} + C_{l_i}^{\Theta_0\varphi} C_{l_j}^{\Theta_0\partial\Theta_0}\,. \tag{2.89}$$

Thus, the temperature bispectrum vanishes unless φ and Θ_0 are correlated. In the models we consider in this work, φ and Θ_0 are not expected to be highly correlated. However, if a model of spatial variations of α predicts that variations of φ and Θ_0 are strongly correlated the bispectrum may be a strong signature of such a model as, unlike the power spectrum or the trispectrum (discussed below), it is first order in φ. We note that expressions analogous to Eqs. (2.88) and (2.89) will hold for the polarization and cross-bispectra as well.

2.6 Trispectra

In analogy with the bispectrum, the temperature trispectrum is defined in terms of the connected piece of the four-point correlation function in Fourier space as

$$\langle\Theta(\mathbf{l}_1)\Theta(\mathbf{l}_2)\Theta(\mathbf{l}_3)\Theta(\mathbf{l}_4)\rangle_c = (2\pi)^2\delta^2(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3+\mathbf{l}_4)T^{\Theta\Theta\Theta\Theta}(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3,\mathbf{l}_4)\,. \tag{2.90}$$

This expression must also be invariant under the exchange of any two fields, or equivalently, $\mathbf{l}$ vectors. We insert Eq. (2.49) into the four-point correlation function, and after some straightforward algebra we find that

$$T^{\Theta\Theta\Theta\Theta}(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3,\mathbf{l}_4) = \sum_{\substack{i,j=1\\ i<j}}^{4} \sum_{\substack{k,l=1(\neq i,j)\\ k<l}}^{4} \mathrm{T_A}(\mathbf{l}_i,\mathbf{l}_j,\mathbf{l}_k,\mathbf{l}_l) + \sum_{\substack{i,j,k=1\\ i<j<k}}^{4} \sum_{\substack{l=1\\ l\neq i,j,k}}^{4} \mathrm{T_B}(\mathbf{l}_i,\mathbf{l}_j,\mathbf{l}_k,\mathbf{l}_l)\,, \tag{2.91}$$

where the sums run over the distinct permutations of $\{1,2,3,4\}$,

$$\begin{aligned}\mathrm{T_A}(\mathbf{l}_i,\mathbf{l}_j,\mathbf{l}_k,\mathbf{l}_l) = {} & C_{l_i}^{\Theta_0\partial\Theta_0} C_{l_j}^{\Theta_0\partial\Theta_0} \left(C^{\varphi\varphi}_{|\mathbf{l}_j+\mathbf{l}_k|} + C^{\varphi\varphi}_{|\mathbf{l}_j+\mathbf{l}_l|}\right) \\ & + \left(C_{l_i}^{\Theta_0\partial\Theta_0} C_{l_j}^{\Theta_0\varphi} + C_{l_i}^{\Theta_0\varphi} C_{l_j}^{\Theta_0\partial\Theta_0}\right)\left(C^{\partial\Theta_0\varphi}_{|\mathbf{l}_j+\mathbf{l}_k|} + C^{\partial\Theta_0\varphi}_{|\mathbf{l}_j+\mathbf{l}_l|}\right) \\ & + C_{l_i}^{\Theta_0\varphi} C_{l_j}^{\Theta_0\varphi}\left(C^{\partial\Theta_0\partial\Theta_0}_{|\mathbf{l}_j+\mathbf{l}_k|} + C^{\partial\Theta_0\partial\Theta_0}_{|\mathbf{l}_j+\mathbf{l}_l|}\right)\,,\end{aligned} \tag{2.92}$$

and

$$\mathrm{T_B}(\mathbf{l}_i,\mathbf{l}_j,\mathbf{l}_k,\mathbf{l}_l) = 2C_{l_i}^{\Theta_0\partial^2\Theta_0}C_{l_j}^{\Theta_0\varphi}C_{l_k}^{\Theta_0\varphi} + 2C_{l_i}^{\Theta_0\varphi}(C_{l_j}^{\Theta_0\varphi}C_{l_k}^{\Theta_0\partial^2\Theta_0} + C_{l_j}^{\Theta_0\partial^2\Theta_0}C_{l_k}^{\Theta_0\varphi})\,. \tag{2.93}$$

If Θ_0 and φ have no cross correlation, the second term vanishes and the first term simplifies to

$$\mathrm{T_A}(\mathbf{l}_i,\mathbf{l}_j,\mathbf{l}_k,\mathbf{l}_l) = C_{l_i}^{\Theta_0\partial\Theta_0}C_{l_j}^{\Theta_0\partial\Theta_0}\left(C_{|\mathbf{l}_j+\mathbf{l}_k|}^{\varphi\varphi} + C_{|\mathbf{l}_j+\mathbf{l}_l|}^{\varphi\varphi}\right)\,. \tag{2.94}$$

Expressions similar to Eqs. (2.91)-(2.93) will hold for the polarization and cross-trispectra as well.

2.6.1 The Kurtosis

The trispectrum is a nontrivial function of six variables and extracting the full trispectrum from CMB data will be a challenging experimental endeavor. It is therefore worthwhile to examine simpler statistical signatures of non-Gaussianity in the CMB. If the trispectrum is nonvanishing, the probability distribution function of $\Theta(\hat{\mathbf{n}})$ on the sky will no longer be precisely Gaussian. As in the case of weak lensing, the deviation from Gaussianity can be parametrized by the kurtosis [2·48], which is a measure of how flattened out or peaked the distribution is relative to a Gaussian distribution.

The kurtosis of a non-Gaussian random field may be written in terms of the trispectrum as

$$\begin{aligned} K(\theta) = \frac{1}{\sigma^4(\theta)}\int \frac{d^2\mathbf{l}_1}{(2\pi)^2}\frac{d^2\mathbf{l}_2}{(2\pi)^2}\frac{d^2\mathbf{l}_3}{(2\pi)^2}\frac{d^2\mathbf{l}_4}{(2\pi)^2}\left\{(2\pi)^2\delta^2(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3+\mathbf{l}_4)T^{\Theta\Theta\Theta\Theta}(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3,\mathbf{l}_4)\right. \\ \left.\times W(l_1\theta)W(l_2\theta)W(l_3\theta)W(l_4\theta)\right\}, \end{aligned} \tag{2.95}$$

where $W(l\theta)$ is a smoothing function with smoothing scale θ, and

$$\sigma^2(\theta) = \int \frac{d^2\mathbf{l}}{(2\pi)^2}C_l^{\Theta\Theta}W^2(l\theta) \tag{2.96}$$

is the smoothed variance.

If we adopt a Gaussian smoothing function

$$W(l\theta) = e^{-\frac{1}{2}\sigma_b^2 l^2}\,, \tag{2.97}$$

with $\sigma_b = \theta/(\sqrt{8\ln 2})$ and insert the expression for the trispectrum from Eq. (2.94) we find after some algebra that

$$K(\theta) = \frac{3}{2\pi^3\sigma^4(\theta)}\int l_1 dl_1 l_2 dl_2 l_3 dl_3\left\{C_{l_1}^{\Theta\partial\Theta}C_{l_2}^{\varphi\varphi}C_{l_3}^{\Theta\partial\Theta}\left[I_0(\sigma_b^2 l_1 l_2)I_0(\sigma_b^2 l_2 l_3)e^{-\sigma_b^2(l_1^2+l_2^2+l_3^2)}\right]\right\}. \tag{2.98}$$

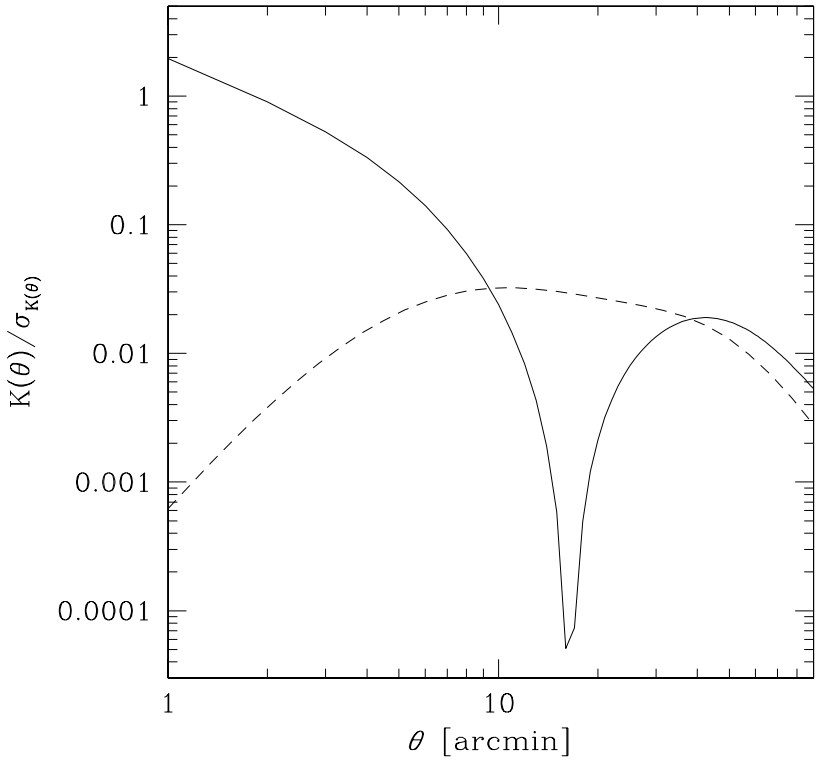

Figure 2.10: The ratio of the kurtosis to the variance in the estimator of the kurtosis due to Gaussian fluctuations for spatial α fluctuations with $\sigma^{(\varphi\varphi)} = 9 \times 10^{-4}$ and $\theta_c = 1^\circ$ (solid) and the same ratio for weak lensing (dashed). The kurtosis is undetectable for weak lensing at any angular resolution, while for spatial α fluctuations of this amplitude the kurtosis may in principle be detected by a high resolution no-noise experiment.

It was shown in Ref. [2·48] that the sample variance of the kurtosis of a Gaussian random field smoothed over an angular scale θ for a full-sky experiment is

$$\sigma^2_{K(\theta)} = \frac{3}{2}\theta^2 \,. \tag{2.99}$$

In Fig. 2.10 we plot the ratio of the kurtosis to the standard deviation of the kurtosis as a function of smoothing scale. Despite the improved variance at high resolution the kurtosis due to weak lensing cannot be detected because in the limit of infinite resolution the kurtosis of weak lensing vanishes more quickly than the variance. This can be understood by noting that weak lensing is power conserving because it maps the temperature at one point on the sky to another, and so with infinite resolution the probability distribution function is Gaussian. In contrast, the kurtosis due to spatial α fluctuations approaches a constant value in the limit of infinite resolution, and so for low θ the signal-

to-noise increases as θ^{-1}. This occurs because α fluctuations do not lead to a pure remapping of of the temperature pattern (they do not conserve power), and thus in the limit of infinite resolution modulate the variance of the (Gaussian) temperature probability distribution function from one patch of the sky to another. The resulting mean probability distribution is no longer Gaussian. Note that the dip in the kurtosis at $\theta \approx 15$ arcminutes occurs because when smoothed over that scale modulation of the CMB due to spatial α fluctuations becomes approximately power conserving. An ideal noise-free experiment at an angular resolution of 1 arcminute could detect the kurtosis due to α fluctuations at the level of $\sigma^{(\varphi\varphi)} = 9 \times 10^{-4}$, while higher resolution experiments could observe lower amplitude α fluctuations.

2.6.2 A Discriminating Filter

As discussed above, weak lensing of the CMB by matter along the line of sight must induce a contribution to the trispectrum

$$L^{\Theta\Theta\Theta\Theta}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4) = \sum_{\substack{i,j=1 \\ i<j}}^{4} \sum_{\substack{k,l=1(\neq i,j) \\ k<l}}^{4} \mathrm{L}(\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k, \mathbf{l}_l)\,, \tag{2.100}$$

where

$$\mathrm{L}(\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k, \mathbf{l}_l) = -\,C_{l_k}^{\Theta_0\Theta_0} C_{l_l}^{\Theta_0\Theta_0} \left[C^{\phi\phi}_{|\mathbf{l}_i+\mathbf{l}_k|} [(\mathbf{l}_i + \mathbf{l}_k)\cdot\mathbf{l}_k][(\mathbf{l}_i + \mathbf{l}_k)\cdot\mathbf{l}_l] + C^{\phi\phi}_{|\mathbf{l}_j+\mathbf{l}_k|} [(\mathbf{l}_j + \mathbf{l}_k)\cdot\mathbf{l}_k][(\mathbf{l}_j + \mathbf{l}_k)\cdot\mathbf{l}_l \right] . \tag{2.101}$$

We emphasize that in this expression $C_l^{\phi\phi}$ is the power spectrum of the projected lensing potential and not the power spectrum of α fluctuations $C_l^{\varphi\varphi}$.

Since this contribution to the trispectrum must be in the CMB, we now derive a filter to distinguish between the weak-lensing and α-fluctuation trispectra.

For compactness, we introduce the notation

$$X_{ijk} \equiv X^{\Theta\Theta\Theta\Theta}(\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k, -\mathbf{l}_i - \mathbf{l}_j - \mathbf{l}_k)\,, \tag{2.102}$$

and

$$\sum_{ijk} \equiv \int \frac{d^2\mathbf{l}_i}{(2\pi)^2} \frac{d^2\mathbf{l}_j}{(2\pi)^2} \frac{d^2\mathbf{l}_k}{(2\pi)^2}\,. \tag{2.103}$$

The signal for detection of the α-fluctuation trispectrum can be in general written as a windowed

integral over trispectrum configurations (quadrilaterals in **l**-space),

$$S = \sum_{ijk} W_{ijk} T_{ijk} \,, \tag{2.104}$$

where W_{ijk} is the to-be-determined window function.

To determine the least-square quartic discriminator we treat the trispectrum due to weak gravitational lensing as a source of noise that should be minimized, and write

$$N_L = \sum_{ijk} W_{ijk} L_{ijk} \,. \tag{2.105}$$

In addition, we at minimum also have the Gaussian noise due to cosmic variance

$$N_G^2 = \sum_{ijk} W_{ijk}^2 G_{ijk} \,, \tag{2.106}$$

where

$$G_{ijk} = \frac{4!}{4\pi f_{sky}} \widehat{C_{l_i}^{\Theta\Theta}} \widehat{C_{l_j}^{\Theta\Theta}} \widehat{C_{l_k}^{\Theta\Theta}} \widehat{C_{|\mathbf{l}_i+\mathbf{l}_j+\mathbf{l}_k|}^{\Theta\Theta}} \,, \tag{2.107}$$

and

$$\widehat{C_l^{\Theta\Theta}} = C_l^{\Theta\Theta} + \frac{4\pi f_{sky} s^2}{t_{exp} T_{CMB}^2} e^{\sigma_b^2 l^2} \tag{2.108}$$

is the sum of the actual CMB power spectrum and the noise power spectrum introduced by an experiment that observes a fraction f_{sky} of the sky with beam width σ_b for a time t_{exp} with detectors of noise-equivalent-temperature s.

The total noise is then just

$$N^2 = N_G^2 + N_L^2 \,. \tag{2.109}$$

We want to solve for the W_{ijk} that maximizes the signal to noise ratio S/N. To do this we set

$$\frac{\delta}{\delta W_{ijk}} \left(\frac{S^2}{N^2} \right) = 0 \,. \tag{2.110}$$

After some straightforward algebra, and making use of the fact that S^2/N^2 is invariant under renormalization of W_{ijk}, we find that the signal-to-noise ratio is extremized only if W_{ijk} takes the

form

$$W_{ijk} = \frac{T_{ijk} - \lambda L_{ijk}}{G_{ijk}} \,. \tag{2.111}$$

Thus, we have

$$\frac{S^2}{N^2} = \frac{\mathcal{T}^2 - 2\lambda\mathcal{T}\mathcal{X} + \lambda^2\mathcal{X}^2}{\mathcal{T} + \mathcal{X}^2 - (2\lambda\mathcal{X} - \lambda^2\mathcal{L})(1+\mathcal{L})} \,, \tag{2.112}$$

where

$$\mathcal{T} = \sum_{ijk} \frac{T_{ijk}^2}{G_{ijk}} \,, \tag{2.113}$$

$$\mathcal{X} = \sum_{ijk} \frac{L_{ijk}T_{ijk}}{G_{ijk}} \,, \tag{2.114}$$

$$\mathcal{L} = \sum_{ijk} \frac{L_{ijk}^2}{G_{ijk}} \,, \tag{2.115}$$

and the optimum weighting is

$$\lambda = \frac{\mathcal{X}}{1+\mathcal{L}} \,. \tag{2.116}$$

In the absence of lensing this expression reduces to the conventional signal-to-noise measure

$$\frac{S^2}{N^2} = \mathcal{T} = \sum_{ijk} \frac{T_{ijk}^2}{G_{ijk}} \,. \tag{2.117}$$

In Fig. 2.11 we show the maximum signal-to-noise ratio for detection of the α-fluctuation trispectrum that would be detected were weak lensing not present. For a given weak lensing power spectrum $C_l^{\phi\phi}$ the actual signal-to-noise will be less than the bound shown.

2.7 Theoretical Models of Variable α

Recent theoretical work [2·11–2·16,2·18–2·21] has considered the ingredients required of field-theory models for variable fine-structure parameter in order to explain a small time variation of α. Here we briefly amend those discussions to consider the field-theory requirements for spatial variations of the sort we consider in this chapter.

The simplest way to introduce spatial variation of α is to couple the photon to a scalar field $\phi(x)$

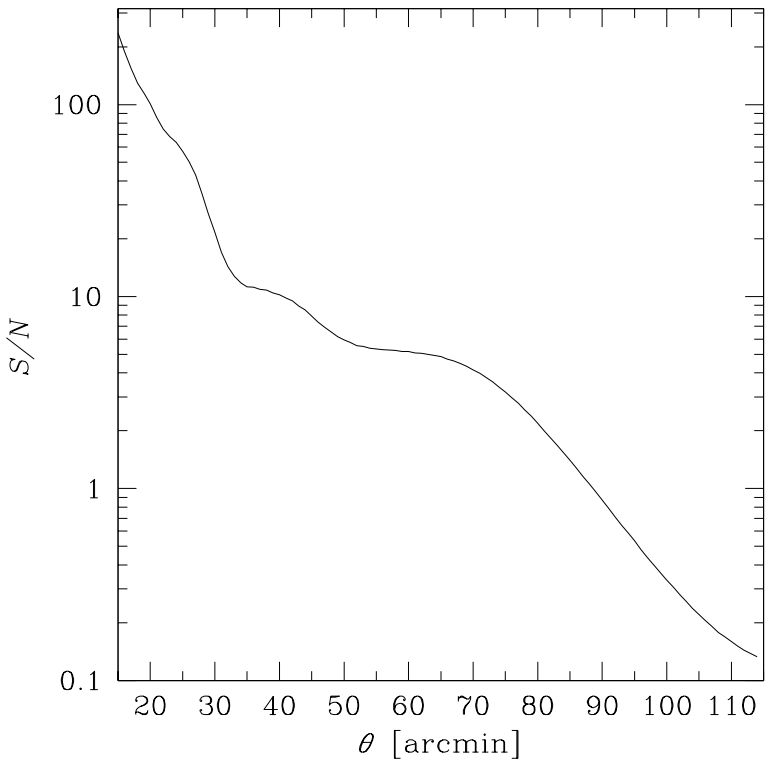

Figure 2.11: The signal-to-noise ratio of the α-fluctuation trispectrum that would be observed by an experiment with $s = 12.42\ \mu\mathrm{K}\sqrt{\mathrm{sec}}$ observing for 1 year (the effective parameters of the Planck satellite) with resolution θ if weak lensing were not present for $\sigma_{(\varphi\varphi)} = 9 \times 10^{-4}$. Since in practice we know weak lensing must be present in the CMB this curve serves as an upper bound on the observable signal-to-noise ratio.

through a term (see, e.g., Ref. [2·20] and references therein)

$$\mathcal{L}_{\phi\gamma\gamma} = -\frac{\lambda}{4} g\left(\frac{\phi}{f_\phi}\right) F^{\mu\nu} F_{\mu\nu}, \tag{2.118}$$

in the Lagrangian that appears in addition to the usual electromagnetic Lagrangian, $F^{\mu\nu}F_{\mu\nu}/4$. Here, f_ϕ is a constant with dimensions of mass, λ is an overall coupling constant, and $g(x) = \sum_{n=1}^{\infty} c_n x^n$ is a dimensionless function of the ratio ϕ/f_ϕ normalized so that $g(1) = 1$. This interaction then leads to a fractional change in α,

$$\frac{\delta\alpha}{\alpha} = -\lambda g\left(\frac{\phi}{f_\phi}\right) = \varphi\,. \tag{2.119}$$

The best present constraint to g comes from the evolution of globular-cluster stars; this requires $[\lambda(dg/d\phi)]^{-1} \gtrsim 1.6 \times 10^{10}$ GeV [2·49]. If we consider fractional fluctuations in α of $\lesssim 10^{-2}$, then $\lambda \sim 10^{-2}$, $\delta\phi \lesssim f_\phi$, and $\sigma^{(\varphi\varphi)} \sim \lambda^2$.

Cosmological mechanisms that might induce spatial variations in ϕ are analogous to those that have been considered, e.g., for spatial variations in the axion field [2·50]. These variations in ϕ can arise either during or after inflation. If ϕ is a spectator field during inflation, then fluctuations in ϕ can be induced quantum mechanically during inflation resulting in a nearly scale-invariant spectrum for ϕ fluctuations. Alternatively, if ϕ is a pseudo-Nambu-Goldstone field in a model with an approximate global symmetry, it could fall during the symmetry-breaking phase transition to different random points on the vacuum manifold (which is periodic in ϕ with period $2f_\phi$) in different causally-disconnected regions. In this case, the scalar-photon interaction will in general be an arbitrary Fourier series in ϕ. A simple interaction of this form, containing only the first harmonic and (as motivated below) constraining $g(\phi/f_\phi)$ to be an even function, is

$$\mathcal{L}_{\phi\gamma\gamma} = -\frac{\lambda}{8}\left[1 - \cos\left(\frac{\pi\phi}{f_\phi}\right)\right] F^{\mu\nu} F_{\mu\nu}\,. \tag{2.120}$$

For such an interaction ϕ will be uncorrelated on scales larger than the horizon, and $\delta\alpha/\alpha$ is fixed by the magnitude of the explicit symmetry breaking term λ, rather than by f_ϕ. In either case (inflation or spontaneous symmetry breaking), the gradient term in the scalar-field Lagrangian will tend to align the scalar field within causally-connected regions of the Universe. Thus, the value of α at the surface of last scatter should be constant within square-degree patches, but will vary from one square-degree patch to another, as we have assumed throughout this chapter.

Now consider constraints to the potential-energy density $V(\phi)$ for the scalar field. If it is quadratic, $V(\phi) = m_\phi^2\phi^2/2$, then (neglecting for the moment the interaction with photons) the

equation of motion for ϕ is

$$\ddot{\phi} + 3H\dot{\phi} + m_\phi^2 \phi = 0\,, \tag{2.121}$$

and has the solution

$$\phi(t) = \phi_0 \frac{\sin(m_\phi t)}{m_\phi t} \tag{2.122}$$

in the matter dominated era. At early times when $H \gg m_\phi$ the value of ϕ is frozen at $\phi \approx \phi_0$, while for $H \lesssim m_\phi$ the field in each horizon will oscillate about its minimum with an amplitude decaying like $1/t$. We thus require $m_\phi \lesssim H_{\rm rec} \simeq 10^{-28}$ eV so that the α fluctuations are frozen in at recombination. On the other hand, observations of the Lyman-α forest at redshifts $z \sim 4$ show no evidence of spatial fluctuations in α larger than one part in 10^4. Thus, if we consider $(\delta\alpha/\alpha) \gtrsim 10^{-4}$ at the surface of last scatter, then we must require that $\lambda g|_{z\approx 4} \lesssim 10^{-4}$. If the leading order term in g is quadratic in ϕ (as we argue below) this leads to the constraint $m_\phi \gtrsim \sqrt{\lambda/0.01} \times 10^{-31}$ eV. Finally, laboratory experiments constrain the value $\dot{\alpha}/\alpha$ today to be $\lesssim 5 \times 10^{-15}$ yr^{-1} [2·51]. To satisfy this requirement we require that $g \propto \phi^2$ to leading order at small ϕ so that $\dot{\alpha}/\alpha \propto t^{-2}$. This leads to the constraint $m_\phi \gtrsim 2 \times (\lambda/0.01) \times 10^{-30}$ eV if the oscillatory factor is unity. We note that if $g \propto \phi^n$ with $n > 2$ for small ϕ this constraint can be relaxed, and that late-time Λ domination does not significantly alter the limit. These constraints will insure that the model conforms to upper limits to α variations in the low-redshift Universe, but we caution that they are constraints based on the root-mean-squared value of the scalar field. If we happen to live in a region where the amplitude of ϕ is randomly small (large) the constraints based on local observations will be weaker (stronger) than discussed above. Also, the nonlinear evolution of the mass distribution may modify constraints based on recent terrestrial phenomena [2·52].

We must also be sure that fluctuations in the scalar field do not lead to density perturbations that exceed those of amplitude 10^{-5}. This constraint requires that the gradient-energy density, $k^2(\Delta\phi)^2 \lesssim 10^{-5}\,\rho_m(z_{\rm rec})$, where $k \simeq H_{\rm rec}$ is the largest wavenumber for which perturbations are significant at the redshift $z_{\rm rec} \simeq 1100$ of decoupling, and $H_{\rm rec} = \Omega_m^{1/2} H_0 (1+z_{\rm rec})^{3/2}$ is the Hubble parameter at decoupling. However, from the Friedmann equation, $\rho_m(z_{\rm rec}) \sim H_{\rm rec}^2 m_{\rm Pl}^2$, where $m_{\rm Pl} \simeq 10^{19}$ GeV is the Planck mass, so we find a constraint $\Delta\phi \lesssim f_\phi \lesssim 3 \times 10^{16}$ GeV.

Now lets examine the effect of the scalar-photon Lagrangian on the scalar field dynamics. If we include the interaction in the Lagrangian it is easy to verify the equation of motion for the scalar field is modified to include a forcing term

$$\ddot{\phi} + 3H\dot{\phi} + m_\phi^2 \phi + \frac{\lambda}{4}\frac{dg}{d\phi}\,\langle F_{\mu\nu}F^{\mu\nu}\rangle = 0\,, \tag{2.123}$$

where

$$\langle F_{\mu\nu}F^{\mu\nu}\rangle = \frac{1}{8\pi}\left(\left\langle E^2\right\rangle - \left\langle B^2\right\rangle\right) \tag{2.124}$$

and $\left\langle E^2\right\rangle$ and $\left\langle B^2\right\rangle$ are the spatially-averaged, squared, electric and magnetic fields respectively. A bath of thermal photons has $\left\langle E^2\right\rangle = \left\langle B^2\right\rangle$ and so will not contribute to the forcing term. However, nonrelativistic matter is a source for electromagnetic fields with $\left\langle E^2\right\rangle \neq \left\langle B^2\right\rangle$, and so the forcing term should be proportional to the density of nonrelativistic matter ρ_m. If, as we discussed above, $g(\phi/f_\phi)$ is dominated by its quadratic term then in the matter dominated era the equation of motion reads

$$\ddot{\phi} + \frac{2}{t}\dot{\phi} + (m_\phi^2 + \frac{\eta}{t^2})\phi = 0\,, \tag{2.125}$$

where $\eta = (2\lambda\xi\rho_{m_0}t_0^2)/f_\phi^2$, and ξ is the fraction of matter density ρ_m due to electromagnetic energy.[1] Thus, if η is small the effect of the forcing term is to introduce a small time-dependent mass term into the equation of motion. The solution will qualitatively behave like that of Eq. (2.122), except that the time that field begins to oscillate may shift by a small η-dependent factor, and the power law of decay will shift from -1 to $-1+\eta$. Generically η need not be small, but for $f_\phi \gtrsim 5\times 10^{15}$ GeV and $\lambda \approx 0.01$ we find $|\eta| \lesssim 0.1$.[2] This brings us uncomfortably close to the limits on f_ϕ from gradient energy density. In general, η should be replaced by the quantity $\tilde{\eta} = \eta + \omega$, where $\tilde{\eta}$ may in principle be small even if η is not. The term ω arises from the terms of the effective non-renormalizable Lagrangian of the form $\lambda_n m_\psi \phi^{2n}\bar{\psi}\psi$ or $\sigma_n m_\chi^2\phi^{2n}\chi^2$ that couple the scalar field ϕ to the energy density in other matter fields. Theories that do not lead to 'fifth forces' which violate the weak-equivalence-principle (WEP) must have $\tilde{\eta} = 0$ precisely.

We thus have a constraint $[\lambda(dg/d\phi)]^{-1} \gtrsim 10^{10}$ GeV to the scalar-photon coupling, and a constraint $(2\lambda/0.01)\times 10^{-30}$ eV $\lesssim m_\phi \lesssim 10^{-28}$ eV to the scalar-particle mass. We note that these relations may seem hard to reconcile, as contributions to the scalar-particle propagator from divergent loop diagrams containing photons should generically be large. Likewise, if ϕ is a Goldstone mode for some spontaneously broken global symmetry, it may be unusual from the point of view of Planck-scale physics [2·53]. It was also pointed out in Ref. [2·54] that variations of a light scalar should give rise to large variations in vacuum energy density. On the other hand, similar problems arise in models for inflation and for scalar-field models for dark energy and dark matter. Here, we take the view that our ϕ-photon coupling is simply a low-energy effective Lagrangian, and anticipate that the light mass can be protected, e.g., by a mechanism such as that proposed in Ref. [2·55] where

[1] Naively, one might guess that ∼nG intergalactic magnetic fields would suggest $\xi \sim 10^{-11}$. However, the quantity of interest here is $\langle B^2\rangle$, rather than $\langle B\rangle^2$, as well as $\langle E^2\rangle$. Thus, strong microscopic magnetic fields in the vicinity of electrons and protons may contribute $\xi_b \sim 10^{-5}$ [2·20].

[2] This assumes a baryon fraction $\Omega_b/\Omega_m = 1/6$, that magnetic energy contributes $|\xi_b| \approx 2\times 10^{-5}$ [2·20], and that the dark matter has negligible electromagnetic energy so that $\xi = (\Omega_b/\Omega_m)\xi_b$.

the scale of an inflaton potential is fixed by the size of an extra dimension. If so then one might assume that whatever mechanism alleviates the cosmological-constant problem may also alleviate problems associated with vacuum-energy gradients.

2.8 Discussion

In this chapter we have studied the effects on the CMB of spatial fluctuations in the fine-structure parameter between causally disconnected regions of the Universe at the time of recombination. Although we have focused on the particular case of fluctuations in the fine-structure parameter, the formalism we have presented may be applied *mutatis mutandis* to other modifications of recombination physics that do not alter the evolution of the dominant density perturbations. As discussed above, such fluctuations will alter the predicted CMB power spectra, introduce a B-mode polarization signal, and introduce temperature and polarization trispectra and perhaps bispectra. We stress here that these results are not dependent on the particular model for α variation discussed in Section 2.7.

From the point of view of effective field theory, variations of the fine-structure parameter can be phrased in terms of a scalar-photon interaction Lagrangian $\mathcal{L}_{\phi\gamma\gamma}$, the parameters of which can be chosen to be consistent with current experimental limits. Ultimately, if such light, cosmologically interesting, scalars do exist in nature their mass must be protected by some yet unknown mechanism.

Acknowledgments

We thank A. Cooray and M. Wise for useful discussions, and A. Cooray and M. Kesden for providing the results of their weak-lensing kurtosis calculations for comparison. KS acknowledges the support of a Canadian NSERC Postgraduate Scholarship. This work was supported in part by NASA NAG5-9821, DoE DE-FG03-92-ER40701 and DE-FG03-02ER41215, and NSF PHY00-71856.

Bibliography

[2·1] P. Jordan, Naturwiss. **25**, 513 (1937); P. Jordan, Z. f. Physik **113**, 660 (1939).

[2·2] E. Teller, Phys. Rev. **73**, 801 (1948).

[2·3] R. H. Dicke, Science **129**, 621 (1959).

[2·4] K. P. Stanyukovich, Doklady Akad. Nauk SSSR **147**, 1348 (1962) [Sov. Phys. Doklady **7**, 1150 (1963)].

[2·5] R. H. Dicke, *The Theoretical Significance of Experimental Relativity* (Gordon and Breach, New York, 1965).

[2·6] G. Gamow, Phys. Rev. Lett. **19**, 759 (1967); G. Gamow, Phys. Rev. Lett. **19**, 913 (1967).

[2·7] P. Forgacs and Z. Horvath, Gen. Rel. Grav. **10**, 931 (1979); P. Forgacs and Z. Horvath, Gen. Rel. Grav. **11**, 205 (1979).

[2·8] J. D. Bekenstein, Phys. Rev. D **25**, 1527 (1982).

[2·9] J. K. Webb *et al.*, Phys. Rev. Lett. **87**, 091301 (2001); J. K. Webb, M. T. Murphy, V. V. Flambaum and S. J. Curran, Astrophys. Space Sci. **283**, 565 (2003).

[2·10] J. N. Bahcall, C. L. Steinhardt and D. Schlegel, Astrophys. J. **600**, 520 (2004).

[2·11] J. P. Uzan, Rev. Mod. Phys. **75**, 403 (2003).

[2·12] R. A. Battye, R. Crittenden and J. Weller, Phys. Rev. D **63**, 043505 (2001).

[2·13] C. J. A. Martins, A. Melchiorri, G. Rocha, R. Trotta, P. P. Avelino and P. Viana, Phys. Lett. B **585**, 29 (2004).

[2·14] H. B. Sandvik, J. D. Barrow and J. Magueijo, Phys. Rev. Lett. **88**, 031302 (2002); J. D. Barrow, H. B. Sandvik and J. Magueijo, Phys. Rev. D **65**, 063504 (2002); J. Magueijo, J. D. Barrow and H. B. Sandvik, Phys. Lett. B **549**, 284 (2002).

[2·15] D. Youm, Mod. Phys. Lett. A **17**, 175 (2002).

[2·16] J. D. Barrow and D. F. Mota, Class. Quant. Grav. **20**, 2045 (2003).

[2·17] V. A. Kostelecky, R. Lehnert and M. J. Perry, Phys. Rev. D **68**, 123511 (2003).

[2·18] Q. G. Huang and M. Li, JHEP **0305**, 026 (2003).

[2·19] G. R. Dvali and M. Zaldarriaga, Phys. Rev. Lett. **88**, 091303 (2002).

[2·20] J. D. Bekenstein, Phys. Rev. D **66**, 123514 (2002).

[2·21] T. Chiba and K. Kohri, Prog. Theor. Phys. **107**, 631 (2002).

[2·22] K. M. Nollett and R. E. Lopez, Phys. Rev. D **66**, 063507 (2002).

[2·23] G. Jungman, M. Kamionkowski, A. Kosowsky and D. N. Spergel, Phys. Rev. Lett. **76**, 1007 (1996); G. Jungman, M. Kamionkowski, A. Kosowsky and D. N. Spergel, Phys. Rev. D **54**, 1332 (1996).

[2·24] P. de Bernardis *et al.* [Boomerang Collaboration], Nature **404**, 955 (2000); S. Hanany *et al.*, Astrophys. J. **545**, L5 (2000); N. W. Halverson *et al.*, Astrophys. J. **568**, 38 (2002); B. S. Mason *et al.*, Astrophys. J. **591**, 540 (2003); A. Benoit *et al.* [the Archeops Collaboration], Astron. Astrophys. **399**, L25 (2003).

[2·25] D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148**, 175 (2003).

[2·26] M. Kamionkowski and A. Kosowsky, Ann. Rev. Nucl. Part. Sci. **49**, 77 (1999).

[2·27] S. Hannestad, Phys. Rev. D **60**, 023515 (1999).

[2·28] M. Kaplinghat, R. J. Scherrer and M. S. Turner, Phys. Rev. D **60**, 023516 (1999).

[2·29] P. J. E. Peebles, Astrophys. J. **153**, 1 (1968);

[2·30] P. Boschan and P. Biltzinger, Astron. Astrophys. **336**, 1 (1998).

[2·31] W. J. Karzas, and R. Latter, Astrophys. J. Suppl. **6**, 167 (1961); G. B. Rybicki, and A. P. Lightman, *Radiative Processes in Astrophysics* (John Wiley & Sons, New York, 1979).

[2·32] C. P. Ma and E. Bertschinger, Astrophys. J. **455**, 7 (1995).

[2·33] S. P. Goldman, Phys. Rev. A **40**, 1185 (1989).

[2·34] G. Breit and E. Teller, Astrophys. J. **91**, 215 (1940). J. Shapiro and G. Breit, Phys. Rev. **113**, 179 (1959) [Erratum-ibid. **115**, 1779 (1959)].

[2·35] S. Seager, D. D. Sasselov and D. Scott, Astrophys. J. Suppl. **128**, 407 (2000).

[2·36] P. J. Mohr and B. N. Taylor, Rev. Mod. Phys. **72**, 351 (2000).

[2·37] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer and E. L. Wright, Astrophys. J. **473**, 576 (1996).

[2·38] G. F. Smoot *et al.*, Astrophys. J. **396**, L1 (1992).

[2·39] J. Kovac, E. M. Leitch, C. Pryke, J. E. Carlstrom, N. W. Halverson and W. L. Holzapfel, Nature **420**, 772 (2002); A. Kogut *et al.*, Astrophys. J. Suppl. **148**, 161 (2003).

[2·40] W. Hu, N. Sugiyama and J. Silk, Nature **386**, 37 (1997).

[2·41] J. R. Bond and G. Efstathiou, Mon. Not. Roy. Astron. Soc. **226**, 655 (1987); A. Kosowsky, Annals Phys. **246**, 49 (1996).

[2·42] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997); M. Kamionkowski, A. Kosowsky and A. Stebbins, Phys. Rev. D **55**, 7368 (1997); W. Hu and M. J. White, Phys. Rev. D **56**, 596 (1997).

[2·43] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962). J. N. Goldberg *et al.*, J. Math. Phys. **8**, 2155 (1967); K. S. Thorne, Rev. Mod. Phys. **52**, 299 (1980).

[2·44] W. Hu, Phys. Rev. D **62**, 043007 (2000).

[2·45] W. Hu, U. Seljak, M. J. White and M. Zaldarriaga, Phys. Rev. D **57**, 3290 (1998).

[2·46] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996).

[2·47] U. Seljak, Astrophys. J. **463**, 1 (1996); M. Zaldarriaga and U. Seljak, Phys. Rev. D **58**, 023003 (1998).

[2·48] M. H. Kesden, A. Cooray and M. Kamionkowski, Phys. Rev. D **66**, 083007 (2002).

[2·49] K. Hagiwara *et al.* [Particle Data Group], Phys. Rev. D **66**, 010001 (2002).

[2·50] M. S. Turner, Phys. Rept. **197**, 67 (1990); G. G. Raffelt, Phys. Rept. **198**, 1 (1990); L. J. Rosenberg and K. A. van Bibber, Phys. Rept. **325**, 1 (2000).

[2·51] Y. Sortais *et al.*, Phys. Scripta **T95**, 50 (2001).

[2·52] D. F. Mota and J. D. Barrow, Phys. Lett. B **581**, 141 (2004).

[2·53] M. Kamionkowski and J. March-Russell, Phys. Lett. B **282**, 137 (1992); M. Kamionkowski and J. March-Russell, Phys. Rev. Lett. **69**, 1485 (1992); R. Holman, S. D. H. Hsu, T. W. Kephart, E. W. Kolb, R. Watkins and L. M. Widrow, Phys. Lett. B **282**, 132 (1992); R. Holman, S. D. H. Hsu, E. W. Kolb, R. Watkins and L. M. Widrow, Phys. Rev. Lett. **69**, 1489 (1992).

[2·54] T. Banks, M. Dine and M. R. Douglas, Phys. Rev. Lett. **88**, 131301 (2002).

[2·55] N. Arkani-Hamed, H. C. Cheng, P. Creminelli and L. Randall, Phys. Rev. Lett. **90**, 221302 (2003).

Chapter 3

Charged-Particle Decay and Suppression of Primordial Power on Small Scales

We study the suppression of the small-scale power spectrum due to the decay of charged matter to dark matter prior to recombination. Prior to decay, the charged particles couple to the photon-baryon fluid and participate in its acoustic oscillations. After these charged particles decay to neutral dark matter the photon-baryon fluid is coupled only gravitationally to the newly-created dark matter. This generically leads to suppression of power on length scales that enter the horizon prior to decay. For decay times of $\sim$3.5 years this leads to suppression of power on subgalactic scales, bringing the observed number of Galactic substructures in line with observation. Decay times of a few years are possible if the dark matter is purely gravitationally interacting, such as the gravitino in supersymmetric models or a massive Kaluza-Klein graviton in models with universal extra dimensions.

Originally published as K. Sigurdson and M. Kamionkowski, *Phys. Rev. Lett.* **92**, 171302 (2004).

3.1 Introduction

The standard inflation-inspired cosmological model, with its nearly scale-invariant power spectrum of primordial perturbations, is in remarkable agreement with observation. It predicts correctly the detailed pattern of temperature anisotropies in the cosmic microwave background (CMB) [3·1], and accurately describes the large scale clustering of matter in the Universe [3·2]. However, on subgalactic scales there are possible problems with the standard cosmology that warrant further investigation. Namely, the model overpredicts the number of subgalactic halos by an order of magnitude compared to the eleven observed dwarf satellite galaxies of the Milky Way [3·3]. Several possible resolutions have been proposed to this apparent discrepancy, ranging from astrophysical mechanisms that suppress dwarf-galaxy formation in subgalactic halos (see, for example, Ref. [3·4])

to features in the inflaton potential that suppress small-scale power and thus reduce the predicted number of subgalactic halos [3·5].

In this chapter, we show that if dark matter is produced by the out-of-equilibrium decay of a long-lived charged particle, then power will be suppressed on scales smaller than the horizon at the decay epoch. Unlike some other recent proposals, which suppress small-scale power by modifying the dark-matter particle properties [3·6], ours modifies the dark-matter production mechanism. In the model we discuss here, prior to decay, the charged particles coupled electromagnetically to the primordial plasma and participate in its acoustic oscillations. After decay, the photon-baryon fluid is coupled only gravitationally to the neutral dark matter. This generically leads to suppression of power for scales that enter the horizon prior to decay. This suppression, reduces the amount of halo substructure on galactic scales while preserving the successes of the standard hierarchical-clustering paradigm on larger scales. Apart from the changes to the model due to the decay process, we adopt the standard flat-geometry ΛCDM cosmological model with present-day dark-matter density (in units of the critical density) $\Omega_d = 0.25$, baryon density $\Omega_b = 0.05$, cosmological constant $\Omega_\Lambda = 0.70$, Hubble parameter $H_0 = 72\ \mathrm{km\,s^{-1}Mpc^{-1}}$, and spectral index $n = 1$.

3.2 The Standard Case

In the standard ΛCDM model the initial curvature perturbations of the Universe, presumably produced by inflation or some inflation-like mechanism, are adiabatic (perturbations in the total density but not the relative density between species) and Gaussian with a nearly scale-invariant spectrum of amplitudes. These initial perturbations grow and react under the influence of gravity and other forces, with the exact nature of their behavior dependent upon the species in question. Because dark-matter particles are, by assumption, cold and collisionless the fractional dark-matter-density perturbation $\delta_d \equiv \delta\rho_d/\rho_d$ can only grow under the influence of gravity. The baryonic species being charged, are tightly coupled by Coulomb scattering to the electrons, which are themselves tightly coupled to the photons via Thomson scattering. The baryons and photons can thus be described at early times as a single baryon-photon fluid, with the photons providing most of the pressure and inertia and the baryons providing only inertia. Gravity will tend to compress this baryon-photon fluid, while the radiation pressure will support it against this compression. The result is acoustic oscillation, and the baryon density perturbation $\delta_b \equiv \delta\rho_b/\rho_b$ and photon density perturbation $\delta_\gamma \equiv \delta\rho_\gamma/\rho_\gamma$ will oscillate in time for length scales inside the horizon (on length scales larger than the horizon, the pressure can have no effect). At early times these perturbations are very small and linear perturbation theory can be applied. This allows an arbitrary density field to be decomposed into a set of independently evolving Fourier modes, labeled by a wavenumber k. Fig. 3.1 shows the growth of a dark-matter density perturbation under the influence of gravity, and the oscillatory

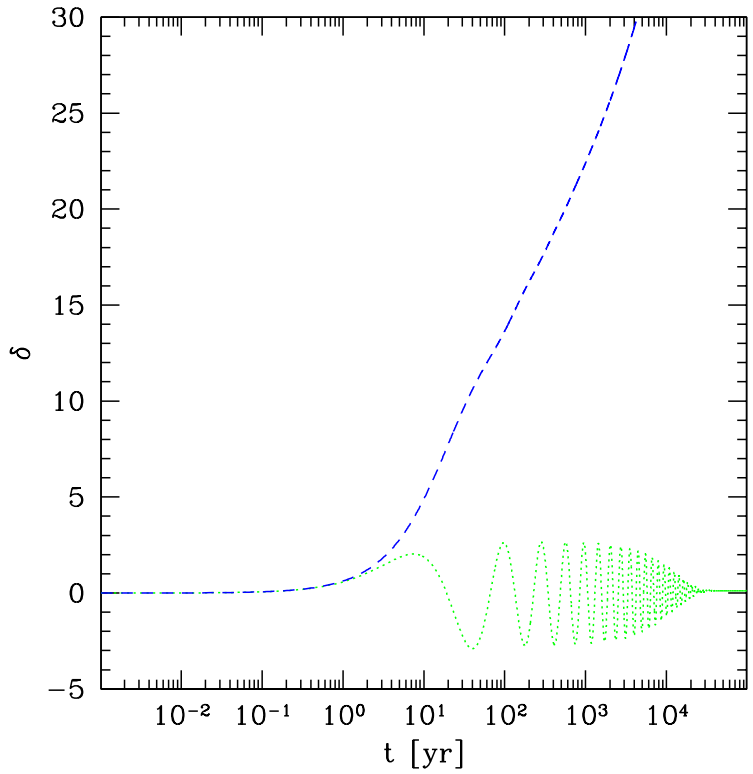

Figure 3.1: The evolution of the comoving wavenumber $k = 3.0$ Mpc^{-1} density perturbations in the early Universe for dark matter (dashed line) and baryons (dotted line). The dark-matter perturbation always grows under the influence of gravity while the baryonic perturbation oscillates due to a competition between gravity and the photon pressure.

behavior of the baryon perturbation for the same wavenumber.

We choose to work in the synchronous gauge where the time slicing is fixed to surfaces of constant proper time so that particle decays proceed everywhere at the same rate. In the synchronous gauge the standard linearized evolution equations for perturbations in Fourier space are (e.g., [3·7])

$$\dot{\delta}_d = -\theta_d - \frac{1}{2}\dot{h}\,, \tag{3.1}$$

$$\dot{\theta}_d = -\frac{\dot{a}}{a}\theta_d\,, \tag{3.2}$$

$$\dot{\delta}_b = -\theta_b - \frac{1}{2}\dot{h}\,, \tag{3.3}$$

$$\dot{\theta}_b = -\frac{\dot{a}}{a}\theta_b + c_s^2 k^2 \delta_b + \frac{4\rho_\gamma}{3\rho_b} a n_e \sigma_T (\theta_\gamma - \theta_b)\,, \tag{3.4}$$

$$\dot{\delta}_\gamma = -\frac{4}{3}\theta_\gamma - \frac{2}{3}\dot{h}\,, \tag{3.5}$$

and

$$\dot{\theta}_\gamma = k^2 \left(\frac{1}{4}\delta_\gamma - \Theta_\gamma\right) + a n_e \sigma_T (\theta_b - \theta_\gamma)\,, \tag{3.6}$$

where θ_b, θ_d, and θ_γ are the divergence of the baryon, dark-matter, and photon fluid velocities respectively and an overdot represents a derivative with respect to the conformal time η. Here h is the trace of the spatial metric perturbations h_{ij}. Its evolution is described by the linearized Einstein equations, which close this system of linearized equations. The last terms on the right-hand-sides of Eqs. (3.4) and (3.6) account for Thomson scattering between baryons and photons, and are responsible for keeping them tightly coupled in the early Universe. In these equations σ_T is the Thomson cross section, n_e is the electron number density, and c_s is the intrinsic sound speed of the baryons. During tight coupling the second moment Θ_γ of the photon distribution and other higher moments can be neglected, and the radiation can reliably be given the fluid description described above.

3.3 Charged-Particle Decay

We now show how Eqs. (3.1)–(3.6) are modified by the decay of a long-lived metastable charged particle to dark matter in the early Universe. We assume that the decay is of the form $q^\pm \to \ell^\pm d$, so the decay of each charged particle $q^\pm$ produces a dark-matter particle d and a charged lepton $\ell^\pm$. Denoting the decaying charged component by the subscript 'q', the background density ρ_q evolves according to the equation,

$$\dot{\rho}_q = -3\frac{\dot{a}}{a}\rho_q - \frac{a}{\tau}\rho_q\,, \tag{3.7}$$

where τ is the lifetime of $q^\pm$. The first term just accounts for the normal a^{-3} scaling of non-relativistic matter in an expanding universe, while the second leads to the expected exponential decay of the comoving density. For the dark matter we have

$$\dot{\rho}_d = -3\frac{\dot{a}}{a}\rho_d + \lambda\frac{a}{\tau}\rho_q\,, \tag{3.8}$$

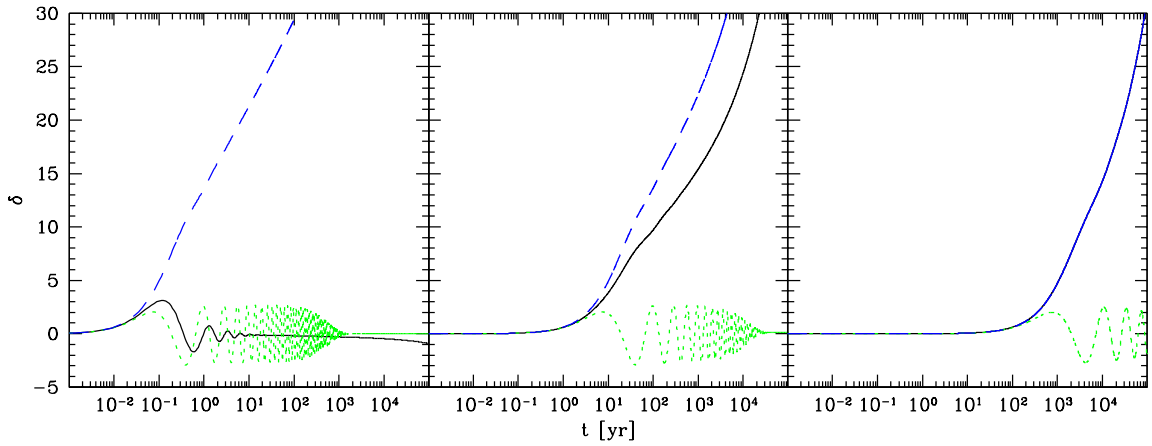

Figure 3.2: The evolution of the comoving wavenumber $k = 30.0$ Mpc^{-1} (left panel), $k = 3.0$ Mpc^{-1} (center panel), and $k = 0.3$ Mpc^{-1} (right panel) density perturbations in the early Universe for dark matter in the ΛCDM model (dashed line) and in the model with $\tau = 3.5$ yr (solid line). The 'β' perturbation is represented by the dotted line. Due to being sourced by the low amplitude 'β' perturbations at early times the dark matter perturbation in the model with a decaying charged component is suppressed relative to the standard ΛCDM case for $k = 3.0$ Mpc^{-1}. For $k \gg 3.0$ Mpc^{-1} (very small scales) δ_d tracks the oscillations in δ_β before decay, while for $k \ll 3.0$ Mpc^{-1} (large scales) δ_d follows the standard growing evolution.

where $\lambda = m_d/m_q$ is the ratio of the masses of the dark matter particle to the charged particle. The energy density in photons evolves according to

$$\dot{\rho}_\gamma = -4\frac{\dot{a}}{a}\rho_\gamma + (1-\lambda)\frac{a}{\tau}\rho_q\,. \tag{3.9}$$

This last equation follows from the assumption that the produced lepton initiates an electromagnetic cascade which rapidly (compared to the expansion timescale) thermalizes with the photon distribution. In practice, the last term on the right-hand-side of Eq. (3.9) is negligibly small because the decay takes place during the radiation dominated era when $\rho_\gamma \gg \rho_q$. Furthermore, limits on the magnitude of μ-distortions to the blackbody spectrum of the CMB constrain $|1-\lambda|$ to be a small number, as we discuss below.

Using covariant generalizations of Eqs. (3.7)–(3.9) we can derive how Eqs. (3.1)–(3.6) are modified by the transfer of energy and momentum from the 'q' component to the dark matter during the decay process. Since the charged 'q' component and the baryons are tightly coupled via Coulomb scattering they share a common velocity $\theta_\beta = \theta_b = \theta_q$. This makes it useful to describe them in terms of a total charged-species component with energy density $\rho_\beta = \rho_b + \rho_q$, which we denote here by the subscript 'β'. Because in the synchronous gauge the decay proceeds everywhere at the same rate this description is even more useful as $\delta_\beta = \delta_b = \delta_q$ is maintained at all times for adiabatic initial conditions. In terms of these 'β' variables, then, we have

$$\dot{\delta}_d = -\theta_d - \frac{1}{2}\dot{h} + \lambda\frac{\rho_q}{\rho_d}\frac{a}{\tau}(\delta_\beta - \delta_d)\,, \tag{3.10}$$

$$\dot{\theta}_d = -\frac{\dot{a}}{a}\theta_d + \lambda\frac{\rho_q}{\rho_d}\frac{a}{\tau}(\theta_\beta - \theta_d)\,, \tag{3.11}$$

$$\dot{\delta}_\beta = -\theta_\beta - \frac{1}{2}\dot{h}\,, \tag{3.12}$$

$$\dot{\theta}_\beta = -\frac{\dot{a}}{a}\theta_\beta + c_s^2 k^2\delta_\beta + \frac{4\rho_\gamma}{3\rho_\beta}an_e\sigma_T(\theta_\gamma - \theta_\beta)\,, \tag{3.13}$$

$$\dot{\delta}_\gamma = -\frac{4}{3}\theta_\gamma - \frac{2}{3}\dot{h} + (1-\lambda)\frac{\rho_q}{\rho_\gamma}\frac{a}{\tau}(\delta_\beta - \delta_\gamma)\,, \tag{3.14}$$

and

$$\dot{\theta}_\gamma = k^2\left(\frac{1}{4}\delta_\gamma - \Theta_\gamma\right) + an_e\sigma_T(\theta_\beta - \theta_\gamma) + (1-\lambda)\frac{\rho_q}{\rho_\gamma}\frac{a}{\tau}\left(\frac{3}{4}\theta_\beta - \theta_\gamma\right)\,. \tag{3.15}$$

We now describe how small-scale modes that enter the horizon prior to decay are suppressed relative to those modes that enter the horizon after decay. Due to the Thomson collision terms, the 'β' component and the photons will be tightly coupled as a 'β'-photon fluid at early times and this fluid will support acoustic oscillations. Furthermore, Eqs. (5.5) and (5.6) show that the dark-matter perturbations are strongly sourced by the perturbations of the 'β' component prior to decay, when the ratio ρ_q/ρ_d is large. Dark-matter modes that enter the horizon prior to decay will thus track the oscillations of the 'β'-photon fluid rather than simply growing under the influence of gravity. After decay, when the ratio ρ_q/ρ_d is small, the source term shuts off and dark-matter modes that enter the horizon undergo the standard growing evolution. In Fig. 3.2 we follow the evolution of the dark-matter perturbations through the epoch of decay. We modified `CMBFAST` [3·8] to carry out these calculations.

3.4 Discussion

In order to suppress power on subgalactic scales the decay lifetime must be roughly the age of the Universe when the mass enclosed in the Hubble volume is equal to a galaxy mass; this occurs when $\tau \sim$ years. In Fig. 3.3 we plot the linear power spectrum of matter density fluctuations at the

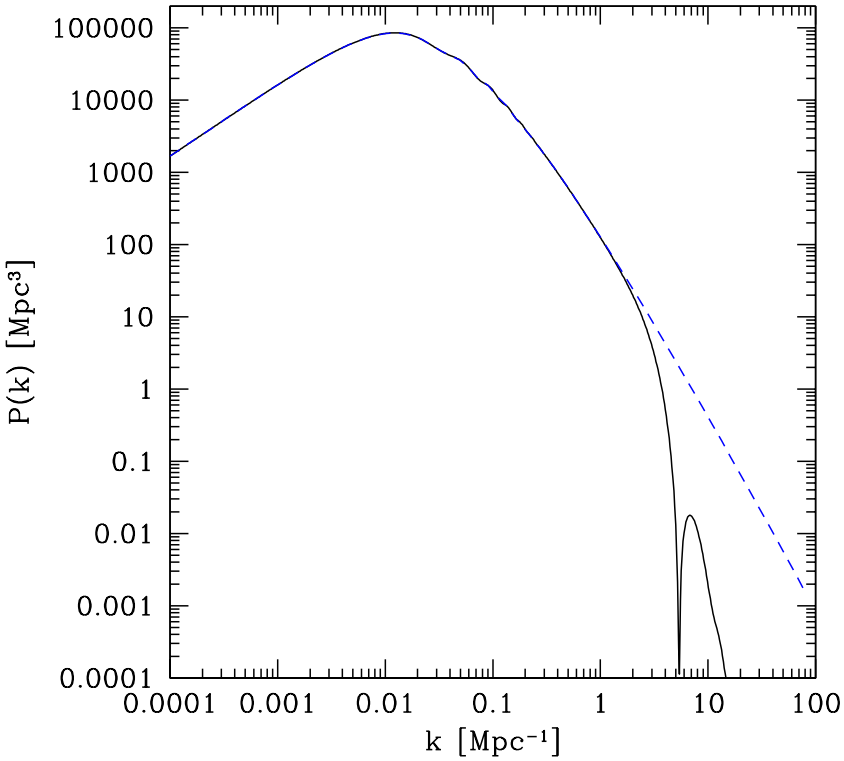

Figure 3.3: The linear-theory power spectrum of matter density fluctuations in the standard ΛCDM model (dashed line), and in the charged decay to dark matter model (solid line) with $\tau = 3.5$ yr. The charged decay model matches the standard ΛCDM model on length scales larger than 0.3 Mpc, but power drops of sharply below 0.3 Mpc.

present day for a charged-particle lifetime $\tau = 3.5$ yr assuming a scale-invariant primordial power spectrum. We see that power is suppressed on scales smaller than $k^{-1} \sim 0.3$ Mpc relative to the standard ΛCDM power spectrum. Suppression of power on these length scales reduces the expected number of subgalactic halos, bringing the predictions in line with observation [3·5] without violating constraints from the Lyman-alpha forest [3·9]. It leaves the shapes of galactic halo density profiles essentially unchanged [3·10]. Of course, the model reproduces the successes of the standard ΛCDM model on larger scales and in the CMB.

The requirements of the charged-particle species are that it has a comoving mass density equal to the dark-matter density today and has a lifetime of $\tau \sim 3.5$ yr. In order to satisfy the constraint to the CMB chemical potential [3·11], the fractional mass difference between the charged and neutral particles must be $\Delta m/m < 3.6 \times 10^{-3}$, and in order for the decay to be allowed kinematically the mass difference must be greater than the electron mass. One possibility is the superweakly interacting massive particle scenario of Ref. [3·12] in which a charged particle may decay to an

exclusively gravitationally interacting particle. For example, in supersymmetric models, the decay of a selectron to an electron and gravitino $\widetilde{e} \to e\,\widetilde{G}$ with $m_{\widetilde{e}} \approx m_{\widetilde{G}} > 122$ TeV would satisfy these constraints, as would the decay of a KK-electron to an electron and KK-graviton $e^1 \to e\,G^1$ with $m_{e^1} \approx m_{G^1} > 72$ TeV in the case of the single universal extra dimension Kaluza-Klein (KK) model discussed in Refs. [3·12, 3·13]. Such masses are close to the unitary bound for thermal production [3·14], but might be accommodated through nonthermal mechanisms or if the next-to-lightest partner is a squark which might then interact more strongly and thus evade this bound. There may also be viable scenarios involving nearly-degenerate charged and neutral higgsinos.

It should be noted that the recent Wilkinson Microwave Anisotropy Probe evidence for early star formation [3·16] argues against the suppression of small-scale power, but these results are not yet conclusive. If it does turn out that traditional astrophysical mechanisms can explain the dearth of dwarf galaxies, then our arguments can be turned around to provide constraints to an otherwise inaccessible region of the parameter space for decaying dark matter [3·15]. Finally, if the mechanism we propose here is realized in nature, then the dearth of small-scale power, along with the detection of a non-zero CMB chemical potential, would be a powerful probe of the particle spectrum of the new physics responsible for dark matter.

Acknowledgments

KS acknowledges the support of a Canadian NSERC Postgraduate Scholarship. This work was supported in part by NASA NAG5-9821 and DoE DE-FG03-92-ER40701.

Bibliography

[3·1] P. de Bernardis *et al.* [Boomerang Collaboration], Nature **404**, 955 (2000); S. Hanany *et al.*, Astrophys. J. **545**, L5 (2000); N. W. Halverson *et al.*, Astrophys. J. **568**, 38 (2002); B. S. Mason *et al.*, Astrophys. J. **591**, 540 (2003); A. Benoit *et al.* [the Archeops Collaboration], Astron. Astrophys. **399**, L25 (2003); D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148**, 175 (2003);

[3·2] J. A. Peacock *et al.*, Nature **410**, 169 (2001); W. J. Percival, Mon. Not. Roy. Astron. Soc. **327**, 1313 (2001); M. Tegmark *et al.* [SDSS Collaboration], Astrophys. J. **606**, 702 (2004).

[3·3] G. Kauffmann, S. D. M. White and B. Guiderdoni, Mon. Not. Roy. Astron. Soc. **264**, 201 (1993); A. A. Klypin, A. V. Kravtsov, O. Valenzuela and F. Prada, Astrophys. J. **522**, 82 (1999); B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel and P. Tozzi, Astrophys. J. **524**, L19 (1999).

[3·4] A. J. Benson, C. S. Frenk, C. G. Lacey, C. M. Baugh and S. Cole, Mon. Not. Roy. Astron. Soc. **333**, 177 (2002); R. S. Somerville, Astrophys. J. **572**, L23 (2002); L. Verde, S. P. Oh and R. Jimenez, Mon. Not. Roy. Astron. Soc. **336**, 541 (2002).

[3·5] M. Kamionkowski and A. R. Liddle, Phys. Rev. Lett. **84**, 4525 (2000).

[3·6] D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84**, 3760 (2000); C. Boehm, P. Fayet and R. Schaeffer, Phys. Lett. B **518**, 8 (2001); X. l. Chen, M. Kamionkowski and X. m. Zhang, Phys. Rev. D **64**, 021302 (2001); X. l. Chen, S. Hannestad and R. J. Scherrer, Phys. Rev. D **65**, 123515 (2002); C. Boehm, H. Mathis, J. Devriendt and J. Silk, arXiv:astro-ph/0309652.

[3·7] C. P. Ma and E. Bertschinger, Astrophys. J. **455**, 7 (1995).

[3·8] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996).

[3·9] M. J. White and R. A. C. Croft, Astrophys. J. **539**, 497 (2000).

[3·10] B. Moore, T. Quinn, F. Governato, J. Stadel and G. Lake, Mon. Not. Roy. Astron. Soc. **310**, 1147 (1999).

[3·11] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer and E. L. Wright, Astrophys. J. **473**, 576 (1996).

[3·12] J. L. Feng, A. Rajaraman and F. Takayama, Phys. Rev. Lett. **91**, 011302 (2003).

[3·13] T. Appelquist, H. C. Cheng and B. A. Dobrescu, Phys. Rev. D **64**, 035002 (2001).

[3·14] K. Griest and M. Kamionkowski, Phys. Rev. Lett. **64**, 615 (1990).

[3·15] S. Profumo, K. Sigurdson, P. Ullio and M. Kamionkowski, Phys. Rev. D **71**, 023518 (2005).

[3·16] A. Kogut *et al.*, Astrophys. J. Suppl. **148**, 161 (2003).

Chapter 4

Dark-Matter Electric and Magnetic Dipole Moments

We consider the consequences of a neutral dark-matter particle with a nonzero electric and/or magnetic dipole moment. Theoretical constraints, as well as constraints from direct searches, precision tests of the standard model, the cosmic microwave background and matter power spectra, and cosmic gamma rays, are included. We find that a relatively light particle with mass between an MeV and a few GeV and an electric or magnetic dipole as large as $\sim 3\times10^{-16}e$ cm (roughly $1.6\times10^{-5}\,\mu_B$) satisfies all experimental and observational bounds. Some of the remaining parameter space may be probed with forthcoming more sensitive direct searches and with the Gamma-Ray Large Area Space Telescope.

Originally published as K. Sigurdson, M. Doran, A. Kurylov, R. R. Caldwell and M. Kamionkowski, *Phys. Rev. D* **70**, 083501 (2004).

4.1 Introduction

A wealth of observational evidence indicates the existence of considerably more mass in galaxies and clusters of galaxies than seen in stars and gas. The source of the missing mass has been a problem since Zwicky's 1933 measurement of the masses of extragalactic systems [4·1]. Given the evidence from galaxy clusters, galaxy dynamics and structure formation, big-bang nucleosynthesis, and the cosmic microwave background that baryons can only account for $\sim 1/6$ of this matter, most of it must be nonbaryonic. Although neutrinos provide the cosmological density of dark matter if their masses sum to $\sim$ 12 eV, such particles cannot (essentially from the Pauli principle) have a sufficiently high phase-space density to account for galactic dark-matter halos [4·2]; moreover, such masses are now inconsistent with neutrino-mass measurements [4·3]. Theorists have thus taken to considering for dark-matter candidates new physics beyond the standard model. To date, the most promising candidates—those that appear in fairly minimal extensions of the standard model

(SM) and which coincidentally have a cosmological density near the critical density—are a weakly-interacting massive particle (WIMP), such as the neutralino, the supersymmetric partner of the photon, Z^0 boson, and/or Higgs boson [4·4], or the axion [4·5]. A considerable theoretical literature on the properties and phenomenology of these particles has arisen, and there are considerable ongoing experimental efforts to detect these particles.

In the absence of discovery of such particles, it may be well worth exploring other possibilities. Thus, an alternative line of investigation takes a more model-independent approach and seeks to explore phenomenologically the possible properties of a dark-matter particle. Along these lines, for example, constraints to strongly-interacting dark matter were considered in Ref. [4·6]; self-interacting dark matter has been considered [4·7, 4·8], and some have studied whether dark matter might be charged [4·9] or have a millicharge [4·10, 4·11].

Our investigation follows in spirit the latter possibility. In particular, dark matter is so called because the coupling to photons is assumed to be nonexistent or very weak, or else we would have presumably seen such particles either through absorption or emission of radiation or in laboratory experiments. In this chapter, we ask the question, "How dark is 'dark'?" In other words, how weak must the coupling of the dark-matter particle to the photon be in order to be consistent with laboratory and astrophysical constraints? In the work on millicharged particles, a dark-matter coupling to photons was assumed to arise from a tiny charge.

In this chapter we consider the possibility that the dark matter possesses an electric or magnetic dipole moment. The punch line, illustrated in Fig. 4.1, is that a Dirac particle with an electric or magnetic dipole moment of order $\sim 10^{-17} e$ cm with a mass between an MeV and a few GeV can provide the dark matter while satisfying all experimental and observational constraints.[1]

In the following section, we introduce the effective Lagrangian for the dipolar dark matter (DDM) interaction with photons. We discuss the relic abundance in Section 4.3. Section 4.4 presents constraints on dark-matter dipole moments and masses that arise from direct searches at low-background detectors as well as constraints from high-altitude experiments. Section 4.5 discusses constraints due to precision tests of the standard model, while Section 4.6 discusses constraints due to the cosmic microwave background and the growth of large-scale structure. We provide some concluding remarks in Section 4.8. An Appendix provides details of the calculation of the drag force between the baryon and DDM fluids used in Section 4.6.

[1] Throughout, we will quote numbers for both the electric and magnetic dipole moments in units of e cm, where e is the electron charge. For reference, the Bohr magneton $\mu_B = e\hbar/2m_e = 1.93 \times 10^{-11}\, e$ cm in these units. Also note that we work in rationalized Gaussian units so that the fine-structure constant $\alpha \equiv e^2/4\pi\hbar c \approx 1/137$, and in particle-physics units ($\hbar = c = 1$) $e^2 \approx 4\pi/137$ and $e \approx 0.303$.

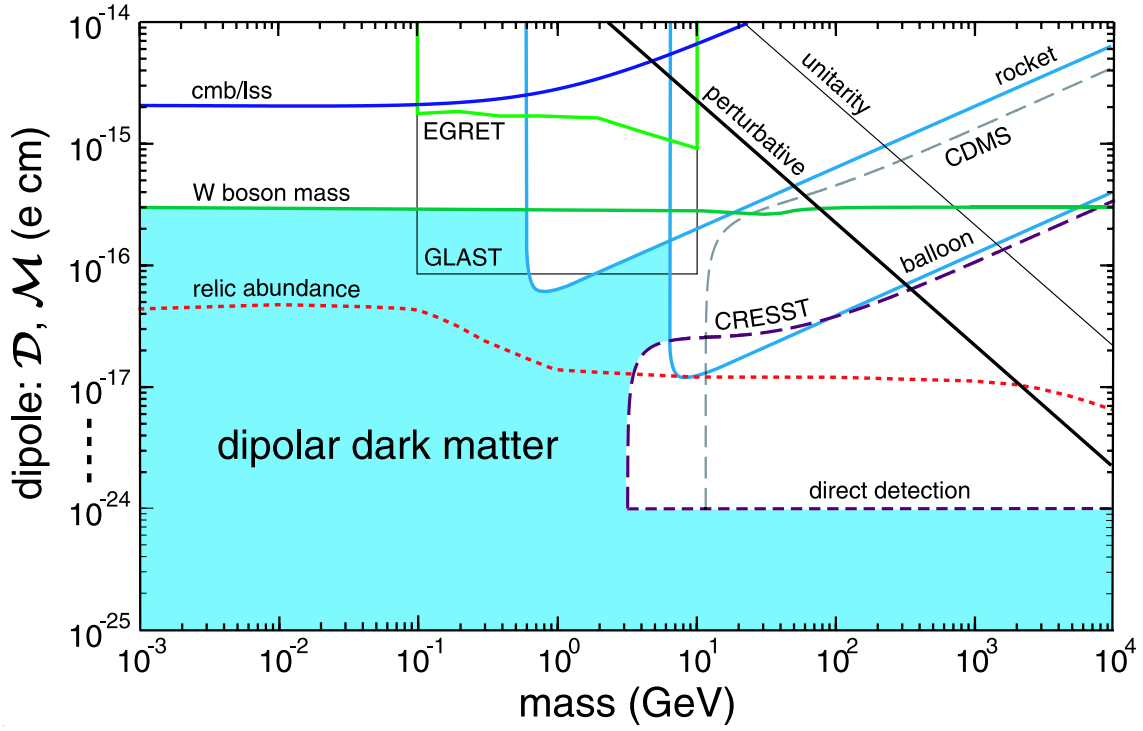

Figure 4.1: The constraints on the dipolar-dark-matter parameter space $[m_\chi, (\mathcal{D}, \mathcal{M})]$ that come from present-day searches and experiments. Viable candidates must lie in the shaded region, below the solid lines and outside the long-dashed lines. The short-dashed "relic abundance" curve shows where the dark matter would have a cosmological density $\Omega_\chi h^2 = 0.135$, assuming standard freeze-out, no particle-antiparticle asymmetry, and no interactions with standard-model particles apart from the dipole coupling to photons. Note that the EGRET and GLAST curves constrain the combination $(\mathcal{D}^4 + \mathcal{M}^4)^{1/4}$, the perturbative and unitarity curves apply to the stronger of $(\mathcal{D}, \mathcal{M})$, while all other curves restrict $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2}$.

4.2 Theory of Dipole Moments

A particle with a permanent electric and/or magnetic dipole moment must have a nonzero spin; we thus consider spin-1/2 particles. Moreover, Majorana particles cannot have permanent dipole moments, so we consider Dirac fermions. Since the spin and the magnetic dipole are both axial vectors, a magnetic dipole moment can arise without violating any discrete symmetries. However, the electric dipole moment is a vector and thus requires time-reversal and parity violation.

The effective Lagrangian for coupling of a Dirac fermion χ with a magnetic dipole moment $\mathcal{M}$ and an electric dipole moment $\mathcal{D}$ to the electromagnetic field $F^{\mu\nu}$ is

$$\mathcal{L}_{\gamma\chi} = -\frac{i}{2}\bar{\chi}\sigma_{\mu\nu}(\mathcal{M} + \gamma_5\mathcal{D})\chi F^{\mu\nu}. \tag{4.1}$$

At energies that are low compared to the dark-matter mass, the photon is blind to distinctions between $\mathcal{M}$ and $\mathcal{D}$ (unless time-reversal-violating observables are considered). Hence, we can discuss limits to $\mathcal{D}$ which equally apply to $\mathcal{M}$, except where noted.

On dimensional grounds, we expect the electric dipole moment to satisfy the limit $\mathcal{D} \lesssim e\, m_\chi^{-1} \simeq 2\times10^{-14}\,(m_p/m_\chi)\,e$ cm, where m_p is the proton mass. Similar arguments also apply to the magnetic dipole moment.[2] This limit is shown as the perturbative bound in Fig. 4.1, as violation of this bound would signal some non-trivial or non-perturbative field configuration. As we will see below, more rigorous but slightly weaker upper limits can be set with unitarity arguments.

These upper limits already simplify our analysis. The phenomenology of charged dark-matter particles is determined largely by the ability of these particles to form atom-like bound states with electrons, protons, or each other. However, dipolar dark matter cannot form such bound states. A neutral particle with a magnetic moment will not form bound states with charged particles. Curiously enough, a neutral particle with an electric dipole moment (EDM) can indeed form a bound state with an electron, as first noted by no less than Fermi and Teller [4·12], but only if the dipole moment is greater than $0.639\, e\, a_0 = 3.4\times10^{-9}\, e$ cm (assuming $m_\chi \gg m_e$) where a_0 is the Bohr radius. For smaller values of the dipole, the electron "sees" both poles of the dipole and finds no stable orbit. This critical electric dipole moment scales inversely with the dipole-electron reduced mass, so a bound state with a proton can occur if the dipole mass is $\gg m_p$ and $\mathcal{D} \gtrsim 1.8\times10^{-12} e$ cm. As we will see below, such values for the EDM cannot occur for a pointlike DDM. Likewise, the weakness of the dipole-dipole interaction prevents the formation of any stable dark-matter atoms.

The first cosmological constraint is that from big-bang nucleosynthesis (BBN). BBN requires that the effective number of relativistic degrees of freedom at $T \sim$ MeV does not exceed the equivalent of roughly 0.2 neutrino species [4·13]. Since the particles we are considering are Dirac particles, they

[2] The limit is satisfied if m_χ is the lightest scale relevant for the DDM sector (see Section 4.5 for discussion). Note, however, that the actual magnitude of the dipole moments in a particular theory can be significantly below this limit.

act like two effective neutrino species and thus cannot be relativistic and in equilibrium at BBN. Generally, such considerations rule out $m_\chi \lesssim$ MeV, and so we restrict our attention in Fig. 4.1 to masses above an MeV. Strictly speaking, if the dipole moment is $(\mathcal{D}, \mathcal{M}) \lesssim 10^{-22}\, e$ cm, and if the particle has no other interactions with standard-model particles, then a particle of mass $\lesssim$ MeV can decouple at a temperature $\gtrsim 10$ GeV, and if so, it will evade the BBN limit.

4.3 Dark Matter Annihilation and Relic Abundance

DDM particles can exist in thermal equilibrium in the early Universe when the temperature $T \gg m_\chi$, and their interactions will freeze out when T drops below m_χ resulting in some cosmological relic abundance. The mass density of relic DDM particles is fixed by the cross section $\sigma_{\rm ann}$ for annihilation to all lighter particles times the relative velocity v through (see, e.g., Eq. (5.47) in Ref. [4·14]),

$$\begin{aligned} \Omega_\chi h^2 &\simeq 3.8\times 10^7 \left(\frac{m_\chi}{m_p}\right) \ln\left(A/\sqrt{\ln A}\right)/A \\ &= 0.135(g_*/10)^{-1/2} \frac{\ln\left[A/\sqrt{\ln A}\right]}{21} \left(\frac{\sigma_{\rm ann} v}{5.3\times 10^{-26}\,{\rm cm}^3\ {\rm sec}^{-1}}\right)^{-1}, \end{aligned} \tag{4.2}$$

where

$$A = 0.038\sqrt{g_*} m_{pl} m_\chi (\sigma_{\rm ann} v) = \frac{6.3\times 10^9\,(g_*/10)^{1/2}}{(m_\chi/{\rm GeV})} \left(\frac{\sigma_{\rm ann} v}{5.3\times 10^{-26}\,{\rm cm}^3\ {\rm sec}^{-1}}\right), \tag{4.3}$$

assuming that annihilation takes place (as it does in our case) through the s wave. Here, g_* is the effective number of relativistic degrees of freedom at the temperature $T_f \sim m_\chi/A$ of freezeout.

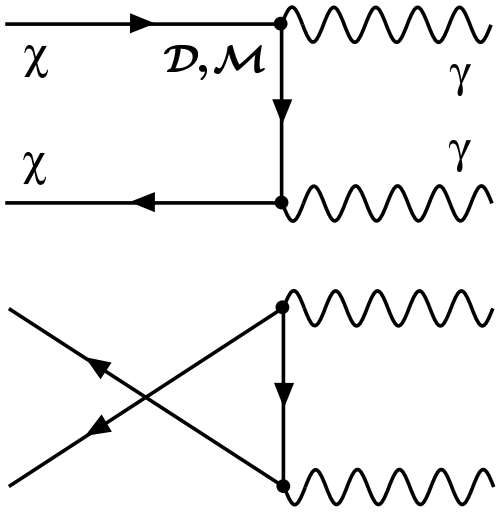

Figure 4.2: Feynman diagrams for annihilation of a DDM–anti-DDM pair to two photons.

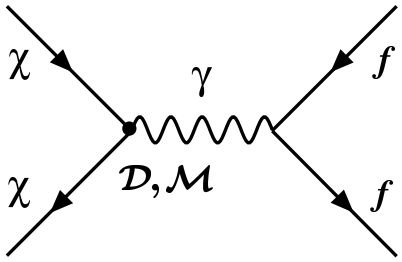

Figure 4.3: Feynman diagram for DDM–anti-DDM annihilation to fermion-antifermion pairs.

For the interaction of Eq. (4.1), DDM–anti-DDM pairs can annihilate to either two photons or to charged particle-antiparticle pairs through the diagrams shown in Figs. 4.2 and 4.3. The cross sections for these two processes (to lowest order in v) are

$$\sigma_{\chi\bar{\chi}\to 2\gamma} v = (\mathcal{D}^4 + \mathcal{M}^4) m_\chi^2 / 2\pi = 1.0 \times 10^{-33} m_{\rm GeV}^2 (\mathcal{D}_{17}^4 + \mathcal{M}_{17}^4)\,{\rm cm}^3\ {\rm sec}^{-1},$$

$$\sigma_{\chi\bar{\chi}\to f\bar{f}} v = N_{\rm eff}\,\alpha(\mathcal{D}^2 + \mathcal{M}^2) = 2 \times 10^{-27}\, N_{\rm eff}\,(\mathcal{D}_{17}^2 + \mathcal{M}_{17}^2)\,{\rm cm}^3\ {\rm sec}^{-1}, \tag{4.4}$$

where $(\mathcal{D}_{17},\, \mathcal{M}_{17}) = (\mathcal{D},\, \mathcal{M})/(10^{-17}\ e\ {\rm cm})$, and $m_{\rm GeV} \equiv m_\chi/{\rm GeV}$. Here, $N_{\rm eff} = \sum_f Q_f^2 N_{cf}$ is the effective number of fermion-antifermion pairs with mass $m_f < m_\chi$, Q_f is the charge of fermion f, and N_{cf} is the number of color degrees of freedom for fermion f. ($N_{cf} = 1$ for electrons.) In the standard model, annihilation can also occur to W^+W^- pairs above threshold. For $(\mathcal{D}_{17}, \mathcal{M}_{17})(m_\chi/m_p) \lesssim 5000$, fermions are the dominant final state. The present-day mass density of DDM particles thus depends primarily on the dipole moment. If such particles are to account for the dark matter, then $\Omega_\chi h^2 = 0.135$ [4·15], and $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} \simeq 1.0 \times 10^{-17}\, e$ cm for $m_\chi \sim 1$ GeV. The full mass dependence of this result is shown in Fig. 4.1.

The cross sections in Eq. (4.4) are s-wave cross sections. According to partial-wave unitarity, the total s-wave annihilation cross section must be $\sigma \lesssim 4\pi/m_\chi^2$ [4·16], so that $(\mathcal{D},\, \mathcal{M}) m_\chi \lesssim 3$, fixed by the cross section for annihilation to two photons. This limit is shown in Fig. 4.1, as is the more stringent, but less rigorous, limit $(\mathcal{D},\, \mathcal{M}) \lesssim e/m_\chi$.

Of course, the present-day mass density of DDM particles could differ from the estimates obtained above. To obtain these results, we assumed (1) that the dipole interaction with photons is the only interaction of these particles; and (2) that there is no particle-antiparticle asymmetry. It is reasonable to assume that in any realistic model, a dark-matter dipole interaction will arise from loop diagrams involving other standard-model and new particles. If so, then there may be other contributions to the annihilation cross sections. In this case, the relic abundance will be smaller than we have estimated above. We thus conclude that if there is no particle-antiparticle asymmetry, our estimates should be treated as an upper bound to the relic abundance, and the $\Omega_\chi h^2$ curve in

Fig. 4.1 should be considered an upper limit to the desired values of $\mathcal{D}$ and $\mathcal{M}$. On the other hand, the relic abundance could also be increased if exotic processes increase the expansion rate during freezeout [4·17].

If there is an asymmetry between χ and $\bar{\chi}$, then the relic abundance is increased relative to our estimate. In this case, however, the present-day Universe should contain predominantly either particles or antiparticles. Although there is no *a priori* reason to expect there to be a particle-antiparticle asymmetry, the observed baryon-antibaryon asymmetry might lead us to expect an analogous dark-matter asymmetry, should the dark matter be composed of Dirac particles. It is possible such asymmetries have a common origin.

Finally, we have assumed above that the particles freeze out when they are nonrelativistic. However, as the dipole moment is lowered for a given mass, freezeout occurs earlier. If the dipole moment is reduced beyond a certain value, and if there are no other couplings to standard-model particles, then freezeout will be relativistic. These particles will then be roughly as abundant as photons, and they will overclose the Universe by huge margins unless their masses are $\lesssim$ few eV; even in this case they will violate constraints to hot dark matter from the CMB and large-scale structure, and they will also be unable, from the Tremaine-Gunn argument, to make up the dark matter in Galactic halos. The transition from nonrelativistic to relativistic freezeout occurs (again, assuming no non-dipole interactions with standard-model particles) for $m_\chi \mathcal{D}_{17}^2 \lesssim 10^{-10}$ GeV for $m_\chi \gtrsim$ MeV, and for $m_\chi \lesssim$ MeV, at $m_\chi \mathcal{D}_{17}^{4/3} \lesssim 200$ MeV.

4.4 Direct Detection

The diagram for scattering of a DDM particle with a particle of charge Ze occurs through the exchange of a photon, as shown in Fig. 4.4 (not unlike the electron-neutron interaction [4·18]). In the nonrelativistic limit, the differential cross section for this process is given by,

$$\frac{d\sigma}{d\Omega} = \frac{Z^2 e^2 \left(\mathcal{D}^2 + \mathcal{M}^2\right)}{8\pi^2 v^2 (1 - \cos\theta)}, \tag{4.5}$$

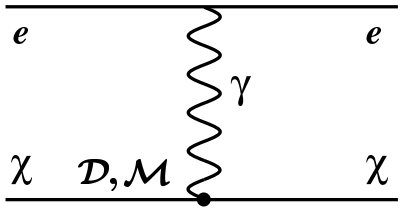

Figure 4.4: Feynman diagrams for elastic scattering of an electron from a DDM particle.

where v is the relative velocity. Through this interaction the DDM might produce a signal in a direct-detection experiment. Although the total cross section is formally infinite, the divergence comes from the small-momentum-transfer scatterings that will be screened by atomic electrons. Roughly speaking, then, the DDM-nucleus elastic-scattering cross section will be $\sigma \sim (Ze)^2(\mathcal{D}^2+\mathcal{M}^2)/2\pi v^2 \simeq 6.4\times10^{-32}\,Z^2(\mathcal{D}_{17}^2+\mathcal{M}_{17}^2)$ cm^2, using $v\sim 10^{-3}\,c$. Current null searches in germanium detectors $[(A,Z)=(76,32)]$ correspond for masses $m_\chi\sim 10$ GeV to a rough upper limit to the cross section $\sim 10^{-42}$ cm^2 [4·19], thus ruling out any dipole moment $(\mathcal{D}_{17}^2+\mathcal{M}_{17}^2)^{1/2}\gtrsim 10^{-7}$. This is shown in Fig. 4.1 as the horizontal dashed line at $\mathcal{D}=10^{-24}\,e$ cm. Note that the cross-section limit depends (and increases) with mass at higher masses; the curve appears as a horizontal line simply because of the break in scale on the y axis.

This seems like a very stringent limit, especially considering the value, $\mathcal{D}_{17}\sim 1$, favored for the correct cosmological density. However, if the dipole moment becomes large enough, the particles will be slowed in the rock above the detector and thus evade detection in these underground experiments. In order to determine the magnitude of the dipole moment for this to occur, we need to calculate the stopping power dE/dx for the particle as it passes through the atmosphere and then the rock. Since elastic scattering takes place through exchange of a photon, it leads to a long-range interaction and, as we have seen above, a formally divergent elastic-scattering cross section. The calculation of the stopping power thus parallels that for ionization loss due to Coulomb collisions, with two important differences. First of all, since the long-range force is $\propto r^{-3}$, as opposed to r^{-2} for Rutherford scattering, stopping occurs via scattering from nuclei, rather than electrons. Second, since this interaction falls more rapidly with radius than the Coulomb interaction, the stopping power is due primarily to hard scatters at small impact parameter, rather than soft scatters at a wide range of impact parameters.

Our result for the stopping power due to scattering from nuclei of charge Z is

$$\frac{dE}{dx} = -n_N \int T\,d\sigma = -n_N \frac{(Ze)^2(\mathcal{D}^2+\mathcal{M}^2)\mu^2c^2}{2\pi m_N}, \tag{4.6}$$

where the kinetic-energy transfer in a single collision is $T=p^2(1-\cos\theta)/m_N$, x is the depth, and $\mu=\mu[m_\chi,m_N]=m_\chi m_N(m_\chi+m_N)^{-1}$ is the reduced mass. For very weak WIMP interactions with nuclei, the most restrictive limits on the WIMP-proton cross section (the smallest upper bounds) are obtained from null searches from experiments that are deepest underground (so as to reduce the background). However, the most restrictive constraints on the cross section at the upper end of the excluded range of cross sections will come from the shallowest underground experiment with a null result. From Eq. (4.6), we find that the dark-matter particles will only penetrate to a depth $x=(E_i-E_f)/|dE/dx|$ where $E_i=\frac{1}{2}m_\chi v^2$ is the initial dark-matter kinetic energy and E_f is the final energy. For a stopped particle, $E_f=0$. However, the particle only needs to lose enough energy

for E_f to drop below the detection threshold for a particular experiment. Equating the maximum kinetic energy transferable in a collision ($\theta = \pi$) to the threshold detectable nuclear-recoil energy ($E_{\rm th}$), we find the velocity must be slowed to $v_f^2 = m_d E_{\rm th}/2\mu[m_\chi, m_d]^2$, where m_d is the mass of the nuclei in the detector, and $\mu[m_\chi, m_d]$ is the DDM-nucleus reduced mass. Hence, the final dark-matter energy must be $E_f = m_\chi m_d E_{\rm th}/4\mu[m_\chi, m_d]^2$.

So far we have assumed that the particle loses energy but is not significantly deflected in each scatter; this will be a good approximation if $m_\chi \gg m_N$. However, when $m_\chi \lesssim m_{\rm Si} \simeq 26$ GeV, the dark-matter particle may be backscattered upon encountering a terrestrial nucleus, rather than simply being slowed without deflection. In this case, the particle will diffuse, undergoing $N \sim m_N/2m_\chi$ scatters before coming to rest. If so, the penetration depth will be reduced by an additional factor of $\sim N^{-1/2}$. We thus replace $dE/dx \to dE/dx[1 + (m_N/2m_\chi)^{1/2}]$ in our expression for the penetration depth.

Then, for a given shielding thickness L, in meters water equivalent (mwe), we invert the expression for the stopping distance to obtain the following bound on the dipole strength:

$$\mathcal{D}^2 + \mathcal{M}^2 > \frac{\frac{1}{2}m_\chi v^2 - \frac{1}{4}\frac{m_\chi m_d}{\mu[m_\chi,m_d]^2}E_{\rm th}}{\frac{e^2}{2\pi}L\sum_i f_i Z_i^2 \frac{\mu[m_\chi,m_i]^2}{m_i^2}[1 + (m_i/2m_\chi)^{1/2}]}\,, \tag{4.7}$$

where the index i sums over the composition of the shielding material, and f_i is the fractional composition by weight. We use a realistic model of the composition of the Earth (chemical composition by weight: O [46.6%], Si [27.7%], Al [8.1%], Fe [5%], Ca [3.6%], Na [2.8%], K [2.6%], Mg [2.1%]) and atmosphere (10 mwe consisting of a 4:1 ratio of nitrogen to oxygen), although the resulting bounds do not change substantially if we ignore the atmosphere and approximate the Earth's crust as entirely composed of Si. We take the initial DDM velocity to be 300 km sec^{-1}.

The shallowest underground experiment with a strong null result is the Stanford Underground Facility (SUF) run of the Cryogenic Dark Matter Search (CDMS) [4·20], which was situated at a depth of 16 mwe. With a detector energy threshold of $E_{\rm th} = 5$ keV, it is sensitive to DDM masses down to $m_\chi \sim 10$ GeV. Near this threshold we find that DDM particles are stopped by the shielding for $\mathcal{D}_{17} \gtrsim 20$. This bound grows more prohibitive with increasing mass, as illustrated in Fig. 4.1. The Cryogenic Rare Event Search with Superconducting Thermometers (CRESST) [4·21], though at a depth of 3800 mwe, extends to slightly lower masses, having a detector threshold energy $E_{\rm th} = 0.6$ keV. Near $m_\chi \sim 1$ GeV the minimum dipole strength is $\mathcal{D}_{17} \gtrsim 2$. However, there are no limits from underground experiments for DDM masses below 1 GeV.

Two airborne experiments—unobscured by the atmosphere or rock—which have closed the windows on some forms of strongly-interacting dark matter [4·6, 4·22], also restrict dark-matter dipole moments. To determine the predicted signal at a detector, we recast Eq. (4.5) as the cross section per energy transfer, whereby $d\sigma/dT = Z^2e^2(\mathcal{D}^2 + \mathcal{M}^2)/4\pi v^2 T$. The event rate (per time, energy,

and unit mass of detector) is

$$R = N_N(0.3\ \mathrm{GeV\,cm^{-3}})\frac{v}{m_\chi}\frac{d\sigma}{dT} = 2.3\,(\mathcal{D}_{17}^2+\mathcal{M}_{17}^2)\frac{m_p}{m_\chi}\left(\frac{\mathrm{keV}}{T}\right)\mathrm{sec^{-1}\,keV^{-1}\,g^{-1}}, \tag{4.8}$$

where N_N is the number of nuclei per gram of material.

The silicon semiconductor detector flown on a balloon in the upper atmosphere by Rich *et al.* [4·23] observed an event rate of $\sim 0.5\,\mathrm{counts\,sec^{-1}\,keV^{-1}\,g^{-1}}$ nuclear recoils in the lowest energy bin at 2 keV. For dark-matter masses above the threshold $\sim$ 7 GeV, we thus require $(\mathcal{D}_{17}^2+\mathcal{M}_{17}^2)(m_p/m_\chi) < 0.2$.

The X-ray Quantum Calorimeter (XQC) detector flown on a rocket by McCammon *et al.* [4·24], was designed to probe the soft x-ray background. However, it serendipitously provides a tight constraint on dipolar dark matter. To predict the expected number of events, we start by computing the number of DDM particles that could impinge on the detector: $N_\chi = n_\chi v A t = 3\times10^7 m_p/m_\chi$, where n_χ is the galactic number density of dark-matter particles, v is their velocity, the cross-sectional area of the XQC detector is $A = 0.33$ cm^2, and the rocket flight was about $t \sim 120$ seconds. The chief property of the XQC detector is the 14 μm thick Si substrate above the thermistor, providing a target of $N_N \sim 6.5\times10^{19}$ nuclei/cm^2. Thus, the event/energy count $N_N N_\chi d\sigma/dT$ integrated over the 25–100 eV energy bin gives a predicted $\sim 0.38\,(\mathcal{D}_{17}^2+\mathcal{M}_{17}^2)(m_p/m_\chi)$ events compared to the ~ 10 observed events. Since the detector has a 25 eV threshold, energy transfer by dark-matter particles as light as $\sim$ 1 GeV can be detected. Altogether, the balloon and rocket experiments exclude a wide range of dipole strengths and masses, as illustrated in Fig. 4.1.

4.5 Constraints from Precision Measurements

We now consider the limits placed on DDM due to precision tests of the standard model. Our use of perturbation theory is valid provided the energy scale of the interaction $\mathcal{E}$ satisfies $(\mathcal{D},\,\mathcal{M})\mathcal{E} \ll 1$. In addition, we require that the DDM mass satisfies $(\mathcal{D},\,\mathcal{M})m_\chi \lesssim 1$, equivalent to the unitary bound [4·16], which ensures the self-consistency of the local operator in Eq. (4.1). Indeed, if Λ is the energy scale at which a dipole is generated then one generally expects $(\mathcal{D},\,\mathcal{M})\,\Lambda \sim 1$. In

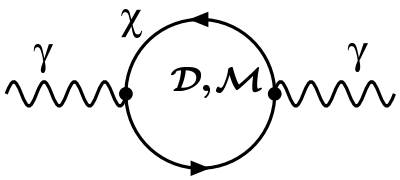

Figure 4.5: One-loop correction to the photon self-energy induced by dipole moments $\mathcal{M}, \mathcal{D}$ of the dark-matter particle.

$\mathcal{L}_{\gamma\chi}$ we assume that all interacting particles with masses greater than Λ have been integrated out. Consequently, one must have at least $m_\chi < \Lambda$ for the dark matter to be dynamical, which also yields $(\mathcal{D},\, \mathcal{M})m_\chi \lesssim 1$.

4.5.1 Muon Anomalous Magnetic Moment

The interaction in Eq. (4.1) contributes to the photon propagator via the diagram shown in Fig. 4.5. The photon-DDM interaction vertices are either both electric or magnetic dipolar; the mixed diagram where one vertex is magnetic and the other is electric is proportional to $\epsilon_{\mu\nu\rho\lambda}F^{\mu\nu}F^{\rho\lambda} = 0$ for photons with equal momenta. The sum of the diagrams produces the following contribution to the photon vacuum-polarization tensor:

$$\begin{aligned} \Pi^{\mu\nu}(q^2) &= \Pi(q^2)\left(q^2 g^{\mu\nu} - q^\mu q^\nu\right) \approx \beta q^2 \left(q^2 g^{\mu\nu} - q^\mu q^\nu\right), \\ \beta &= \frac{\mathcal{D}^2 + \mathcal{M}^2}{8\pi^2}\left(1 - \frac{1}{3}\ln\frac{m_\chi^2}{\mu^2}\right), \end{aligned} \tag{4.9}$$

where the photon momentum is taken to be small, $q^2 \ll m_\chi^2$ (resulting in $\beta q^2 \ll 1$), and μ is the renormalization scale, which should be smaller than Λ. We take $\mu \lesssim 1$ TeV for our estimates. With this self-energy correction, the photon propagator for $\beta q^2 \ll 1$ can be written as

$$-ig^{\mu\nu}\left(\frac{1}{q^2} - \frac{1}{q^2 - 1/\beta}\right). \tag{4.10}$$

The second term above generates a correction to the muon gyromagnetic ratio $\delta a_\mu = -\alpha m_\mu^2 \beta/3\pi$. Interestingly, this contribution is not explicitly suppressed by the DDM mass. In view of recent measurements [4·25] and comparison with the SM predictions, we require that δa_μ does not exceed 10^{-9}, whereby $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} < 6 \times 10^{-15}$ e cm. The order of magnitude of this result can be obtained on dimensional grounds, if we consider that the DDM dipole moment contributes to a_μ via at least a two-loop graph (see Fig. 4.6), with two electromagnetic couplings and two dipole couplings.

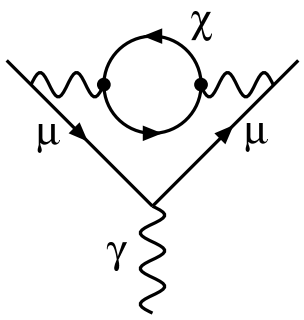

Figure 4.6: Lowest-order correction to the muon anomalous magnetic moment induced by dipole moments of the dark-matter particle.

Including a factor $1/16\pi^2$ per loop, one obtains the estimate,

$$\delta a_\mu \sim \frac{e^2}{(16\pi^2)^2} \left(\mathcal{D}^2 + \mathcal{M}^2\right) \mathcal{E}^2, \tag{4.11}$$

where $\mathcal{E}$ is the characteristic energy scale for the process. In the case of the muon, $\mathcal{E} \sim m_\mu$, which reproduces the rigorous result to within an order of magnitude.

4.5.2 Electric Dipole Moments

Furry's theorem tells us that in evaluating radiative corrections to a process one should only keep the diagrams with an even number of photons attached to a closed loop.[3] Contributions with an odd number of photons attached sum to zero. On the other hand, one must have an odd number of time-reversal-odd (T-odd) EDM vertices in the DDM loop to generate a T-odd operator. These considerations show that the lowest-order non-vanishing diagram must have four photons attached to the DDM loop; diagrams with two photons attached, similar to the one in Fig. 4.6, vanish (see above). Out of the four photons attached to the DDM loop, either one or three can have EDM coupling to DDM. Note that in this scenario both electric and magnetic DDM moments are necessary to generate a dipole moment for a SM fermion. With these considerations in mind, the lowest-order three-loop diagram that induces an EDM for a charged fermion is shown in Fig. 4.7. One obtains the following estimate for the induced electric dipole moment:

$$\mathcal{D}_f \sim \left[\mathcal{D}\mathcal{M}(\mathcal{D}^2 + \mathcal{M}^2)\right] \frac{e^3 m_f^3}{(16\pi^2)^3} \ln^2 \frac{m_\chi}{m_f}, \tag{4.12}$$

[3]Since the theorem is valid for interactions that preserve charge conjugation invariance we can apply it to electric and magnetic dipole moment interactions.

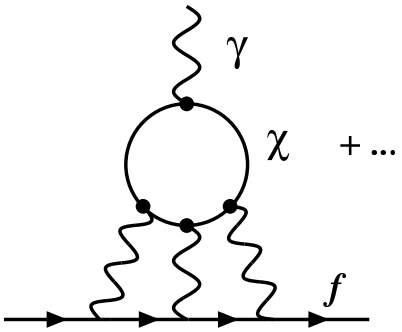

Figure 4.7: Three-loop contributions to the EDM of a charged fermion f. Either one or three of the DDM-photon interaction vertices must be of EDM type. The dots indicate all other diagrams which can be obtained from the one shown by permutation of the interaction vertices.

where a possible double-logarithmic enhancement is included. For the electron, the present limit is $\mathcal{D}_e < 4 \times 10^{-27}\ e$ cm, which implies $(\mathcal{D},\, \mathcal{M}) < 3 \times 10^{-13}\ e$ cm for $m_\chi \sim 100$ GeV and $\mathcal{D} \sim \mathcal{M}$. For smaller m_χ the limit becomes weaker.

There are constraints on the EDMs of other systems, such as the neutron and the mercury atom. It is non-trivial to translate such constraints into limits on the underlying interaction. In case of the neutron, one may attempt to treat the neutron as a point particle for virtual-photon energies below 1 GeV. For higher loop momenta, photons begin seeing the quarks and the contribution to the EDM becomes suppressed by the quark masses. In this case one may use the above equation with $m_f \to m_n$ and no logarithmic enhancement, in order to estimate the neutron EDM:

$$\mathcal{D}_n = \left[\mathcal{D}\mathcal{M}(\mathcal{D}^2 + \mathcal{M}^2)\right] \frac{e^3 |\kappa_n|^3 m_n^3}{(16\pi^2)^3} < 6 \times 10^{-26}\ e \text{ cm}, \tag{4.13}$$

which results in the limit $(\mathcal{D},\, \mathcal{M}) \lesssim 4 \times 10^{-15}\ e$ cm (assuming $\mathcal{D} \simeq \mathcal{M}$. In the above equation, $\kappa_n = -1.91$ is the magnetic moment of the neutron. It appears because the neutron is neutral, and couples to the photon in Fig. 4.7 via a magnetic-moment interaction. The limit for the EDM of the mercury atom is much stronger than the neutron, $\mathcal{D}_{Hg} \lesssim 10^{-28}\ e$ cm. Unfortunately, the mercury atom is a complicated system for which the EDM is influenced by many sources. Therefore, we leave the Hg limit for future study.

4.5.3 W Boson Mass

The DDM can contribute to the running of the fine-structure constant for momenta ranging up to the Z^0 mass. Such running will affect the relationship between the Fermi constant G_F, the mass of the $W^\pm$ boson, and the fine-structure constant at zero momentum:

$$m_W^2 = \frac{\pi\alpha}{\sqrt{2} G_F} \frac{1}{(1 - m_W^2/m_Z^2)(1 - \Delta r)}, \tag{4.14}$$

where Δr is a correction calculable in a given theory. The interaction in Eq. (4.1) modifies the standard expression for Δr, whereby $\Delta r^{New} = \Pi(m_Z^2) - \Pi(0)$. In the standard model $\Delta r^{SM} = 0.0355 \pm 0.0019 \pm 0.0002$. On the other hand, one can use experimentally measured values for α, $m_{W,Z}$, and G_F in Eq. (4.14) to infer $\Delta r^{Exp} = 0.0326 \pm 0.0023$, which gives $\Delta r^{New} < 0.003$ at the 95% confidence level. Therefore, we obtain the limit $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} \lesssim 3 \times 10^{-16}\ e$ cm. A full calculation of the vacuum polarization yields the constraint shown in Fig. 4.1. This turns out to be the strongest constraint due to precision tests of the standard model.

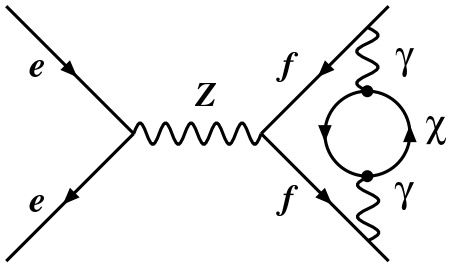

Figure 4.8: Lowest-order correction to Z^0-pole observables induced by dipole moments of the dark-matter particle.

4.5.4 Z-Pole Observables

The DDM will contribute to various Z^0-pole observables through two-loop diagrams similar to the one shown in Fig. 4.8, at the level $\alpha(\mathcal{D}^2+\mathcal{M}^2)m_Z^2/64\pi^3$. Requiring that these contributions do not exceed the $\sim 0.1\%$ precision to which Z^0-pole observables are typically known [4·26] results in the constraint $(\mathcal{D}^2+\mathcal{M}^2)^{1/2} < 10^{-14}$ e cm. Note that in order for perturbation theory to apply to energies $\mathcal{E} \sim m_Z$, one must have $(\mathcal{D}, \mathcal{M})m_Z < 1$, which means $(\mathcal{D}, \mathcal{M}) \lesssim 7\times 10^{-16}$ e cm. Interestingly, consistency with a perturbative treatment at the Z^0 pole imposes much stronger constraints on the DDM than the Z^0-pole observables themselves.

4.5.5 Direct Production

If kinematically allowed, DDM can be directly produced in various scattering and decay experiments. In this case one may use the "missing energy" signature to constrain the DDM couplings. Here, we consider missing-energy constraints from both low-energy (B^+ and K^+ meson decays) as well as collider (LEP, CDF) experiments.

4.5.5.1 B^+ and K^+ decays

Searching for light ($m_\chi \lesssim 1$ GeV) dark matter using missing-energy signatures in rare B^+ meson decays was originally suggested in Ref. [4·27].

There, data from BABAR [4·28] and CLEO [4·29] were used to set a limit $Br(B^+ \to K^+ + \text{invisible}) \lesssim 10^{-4}$ [derived from $Br(B^+ \to K^+\bar{\nu}\nu)$]. This limit can be used to constrain the dipole moments of dark matter. The diagram for $B^+ \to K^+\bar{\chi}\chi$ decay is shown in Fig. 4.9(a). The rate for this decay can be related to the photon-exchange contribution to $B^+ \to K^+l^+l^-|_\gamma$ shown in Fig. 4.9(b). Since the graphs have identical topologies, the difference in rates will come from the difference in effective couplings and the final-state phase-space integrals. One can estimate,

$$\frac{Br\left(B^+ \to K^+\bar{\chi}\chi\right)}{Br\left(B^+ \to K^+l^+l^-\right)|_\gamma} \sim \frac{(\mathcal{D}^2+\mathcal{M}^2)m_{B^+}^2}{e^2}\frac{PS\left(K^+\bar{\chi}\chi\right)}{PS\left(K^+l^+l^-\right)}, \tag{4.15}$$

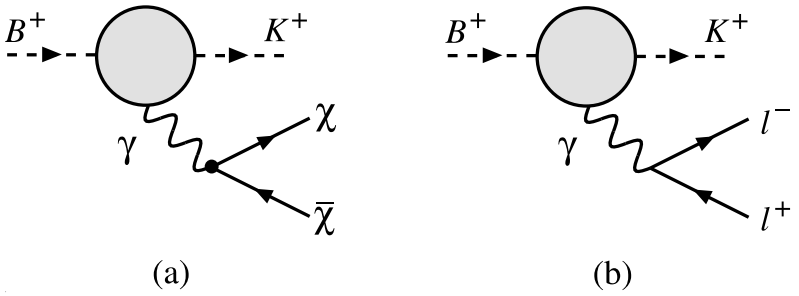

Figure 4.9: Photon-exchange contributions to (a) $Br\,(B^+ \to K^+\bar{\chi}\chi)$ and (b) $Br\,(B^+ \to K^+l^+l^-)$. The blobs collectively represent quark flavor-changing interactions.

where $PS(\cdots)$ stand for the corresponding final-state phase-space integrals, and $m_{B^+} = 5.279$ GeV is the B^+ mass. Belle [4·30] and BABAR [4·31] find $Br(B^+ \to K^+l^+l^-) \lesssim 10^{-6}$. Since the ratio of the phase-space integrals is of order unity, and since in the absence of accidental cancellations $Br(B^+ \to K^+l^+l^-)|_\gamma \lesssim Br(B^+ \to K^+l^+l^-)$, one obtains the constraint

$$2 \times 10^{-6} \left(\frac{(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} m_{B^+}}{e} \right) \lesssim 10^{-4}, \tag{4.16}$$

which leads to $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} \lesssim 3.8 \times 10^{-14}$ e cm. This constraint is relevant for $m_\chi < (m_{B^+} - m_{K^+})/2 = 2.38$ GeV.

Rare K^+ decays can be treated in a similar manner. The relevant branching ratios are $Br(K^+ \to \pi^+e^+e^-) = 2.88^{+0.13}_{-0.13} \times 10^{-7}$ and $Br(K^+ \to \pi^+\bar{\nu}\nu) = 0.157^{+0.175}_{-0.082} \times 10^{-9}$ [4·26]. The resulting constraint on the dipole moment is $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} \lesssim 1.5 \times 10^{-15}$ e cm. This constraint applies for $m_\chi < (m_{K^+} - m_{\pi^+})/2 = 0.18$ GeV. We see that constraints from B^+ and K^+ decays are not competitive with other constraints shown in Fig. 4.1.

4.5.5.2 Collider Experiments

A typical example of a process where DDM can be directly produced in a collider experiment is shown in Fig. 4.10. Here, two fermions f scatter to produce a final state containing some set of visible particles X (photon, multiple jets, etc.), along with particles that are not detected. In the SM, the latter are neutrinos.

Limits on the rate for such processes have been set by, e.g., the L3 and CDF collaborations [4·32]. At LEP, X consisted of a single photon whereas at CDF it consisted of one or more hadronic jets.

In order to translate constraints from collider experiments into limits on DDM couplings one needs an analytical expression for the rate for $f\bar{f} \to X\bar{\chi}\chi$. Naive application of the effective Lagrangian in Eq. (4.1) would result in upper limits from these missing-energy searches of roughly $10^{-17}\,e$ cm. However, this constraint does not actually exclude larger values of the dipole moments.

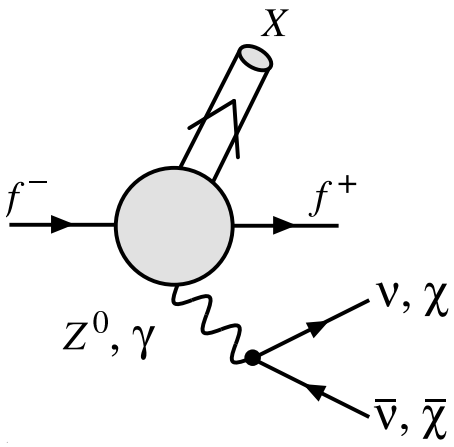

Figure 4.10: A typical process that would produce a missing-energy signature in a collider experiment. Here, X stands for the visible portion of the final state. Neutrinos or DDM may carry a large fraction of the energy but are not detected.

Indeed, as discussed above, perturbation theory will break down when the energy scale for the process $\mathcal{E}$ satisfies $(\mathcal{D}, \mathcal{M})\mathcal{E} \gtrsim 1$. This means that missing-energy searches from L3 ($\mathcal{E} \approx 200$ GeV) and CDF ($\mathcal{E} = 1.8$ TeV) cannot be used to probe effective dipole moments $(\mathcal{D}, \mathcal{M}) > 10^{-16}$ e cm and $(\mathcal{D}, \mathcal{M}) > 10^{-17}$ e cm, respectively, unless the underlying physics that gives rise to the dipole moment is specified.

4.5.6 Other Laboratory Constraints

Important constraints can be obtained for millicharged particles from the Lamb shift [4·10,4·33] and from a targeted experiment at SLAC [4·34]. We have checked, however, that due to the different energy dependence of the photon-dipole vertex, as opposed to the photon-millicharge vertex, the DDM-induced correction to the Lamb shift is small for dipole moments not eliminated by other precision measurements, such as the running of the fine-structure constant. Likewise, although the SLAC experiment is in principle sensitive to neutral particles with a dipole, the production and energy deposition of dipole particles is sufficiently small, for dipole moments consistent with accelerator experiments, to evade detection in the SLAC experiment.

4.6 Constraints from Large-Scale Structure and the CMB

We now consider the effects of the interaction $\mathcal{L}_{\gamma\chi}$ on the evolution of cosmological perturbations and their resulting imprints on the matter power spectrum and the CMB. A dipole moment can induce a coupling of the dark matter to the baryon-photon fluid by scattering from photons through the diagrams shown in Fig. 4.11, or by scattering from protons, helium nuclei, and/or electrons through

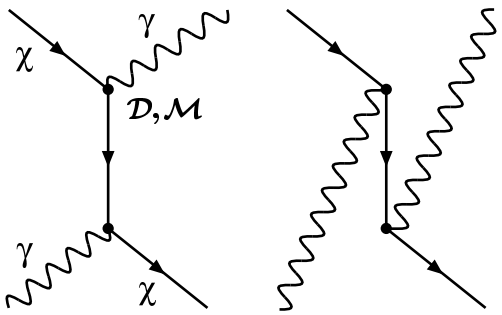

Figure 4.11: The photon-DDM scattering diagram, the analogue of the Compton-scattering diagram for electric or magnetic moments.

the diagram shown in Fig. 4.4. What we will show below is that the dark matter is coupled to the baryon-photon fluid at early times, and decouples at later times. When the dark matter is coupled to the photon-baryon fluid, the pressure of the plasma resists the growth of gravitational potential wells. Thus, the short-wavelength modes of the density field that enter the horizon at early times will have their growth suppressed relative to the standard calculation resulting in a suppression of small-scale power. The evolution of the longer-wavelength modes that enter the horizon after the dark matter has decoupled remain unaffected. Before presenting the results of our detailed analysis, we begin with some simple estimates.

We first show that DDM-photon scattering is negligible compared with DDM-baryon scattering in providing the drag force between the DDM fluid and the baryon-photon fluid. To do so, we first estimate the drag force per unit mass (i.e., the deceleration) on a DDM particle that moves with a velocity V with respect to the rest frame of a blackbody at temperature T. The diagrams in Fig. 4.11 will lead to a photon-DDM scattering cross section $\sigma_{\chi\gamma} \sim (\mathcal{D}^4 + \mathcal{M}^4)E_\gamma^2$. Considering that the momentum transfer to the DDM particle in each scatter is $\sim E_\gamma$ and that the difference of the fluxes of photons moving in the same versus opposite direction to the DDM particle is $\sim T^3(V/c)$, we conclude that the deceleration due to photon scattering is $a_\gamma \sim (\mathcal{D}^4 + \mathcal{M}^4)T^6(V/c)/m_\chi$.

We next estimate a_b, the drag force per unit mass due to DDM-proton scattering. We first note that the peculiar velocity of the baryon-photon fluid (obtained from the continuity equation) in the early Universe will be $V \sim (H/k)c\delta$, where H is the Hubble parameter, k is the physical wavenumber of the mode in question, and $\delta \sim 10^{-5}$ is the amplitude of the fractional density perturbation. Since $(H/k) \lesssim 1$ for modes inside the horizon, we must have $V \lesssim 10^{-5}\,c$. On the other hand, the proton thermal velocity dispersion is $\bar{v}_p \sim (T/m_p)^{1/2}c \gtrsim 10^{-4.5}\,c$ before recombination. Thus, for the early times of interest to us here, the relative velocity between the DDM and the baryon-photon fluid is small compared with the thermal proton velocities. Thus, the appropriate relative velocity to use in Eq. (4.5) in estimating the proton-DDM cross section is $\bar{v}_p$, resulting in a DDM-proton cross section $\sigma_{\chi b} \sim e^2(\mathcal{D}^2 + \mathcal{M}^2)/\bar{v}_p^2$. The momentum transfer per scatter is $\sim \mu\bar{v}_p$, where μ is the proton-DDM

reduced mass, and the difference of the fluxes of protons moving with as opposed to against the DDM fluid is $n_p V$, where n_p is the proton density. The drag force per unit mass on the DDM fluid due to scattering with protons is thus $a_b \sim e^2(\mathcal{D}^2 + \mathcal{M}^2)(\mu/m_\chi)(V/\bar{v}_p)n_p$. We also conclude from the appearance of μ in this result that drag due to scattering from electrons is negligible compared with baryon drag.

Since $n_p \propto T^3$ and $\bar{v}_p \propto T^{1/2}$, we find $a_b \propto T^{2.5}$ as opposed to $a_\gamma \propto T^6$. Thus, at early times, photon drag dominates while baryon drag dominates at later times. The transition occurs at a temperature $T \sim$ GeV for values of m_χ and ($\mathcal{D}$, $\mathcal{M}$) of interest to us, and such high temperatures correspond to (comoving) horizon scales considerably smaller than the distance scales ($\gtrsim$ Mpc) probed by large-scale structure. We can thus neglect photon drag. From $a_b \propto T^{2.5}$ we infer a deceleration time for the DDM fluid $t_{\rm dec} = V/a_b \propto T^{-2.5}$. Since this decreases more rapidly than the Hubble time $t_U \sim m_{\rm Pl} T^{-2}$ (where $m_{\rm Pl} \sim 10^{19}$ GeV is the Planck mass), we conclude that DDM particles are tightly coupled to the plasma at early times and then are decoupled at later times. With these rough estimates, the transition temperature is $T \sim 10\,{\rm keV}\,(\mathcal{D}_{15}^2 + \mathcal{M}_{15}^2)^{-2}(1 + m_\chi/m_p)^2$ suggesting that power on scales smaller than $\lambda \sim 10^{-2}\,(\mathcal{D}_{15}^2 + \mathcal{M}_{15}^2)^2(1 + m_\chi/m_p)^{-2}$ Mpc will be suppressed. The $T^{0.5}$ dependence of the ratio of the deceleration and expansion times suggests furthermore that the small-scale suppression will change gradually, rather than exponentially, with wavenumber k. Knowing that the linear-theory power spectrum is measured and roughly consistent with scale invariance down to distances ∼Mpc leads us to conclude that dipole moments $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} \gtrsim 5 \times 10^{-15}\,(1 + m_\chi/m_p)^{1/2}\,e$ cm will be ruled out. Strictly speaking, when $m_\chi < m_p$, the detailed calculation must take into account the velocity dispersion of the DDM particles; our detailed calculation below includes these effects. As seen below, the detailed analysis leads to a slightly stronger constraint.

4.6.1 Exact Equations

The standard calculation of perturbations in an expanding universe requires solution of the combined Einstein and Boltzmann equations for the distribution functions of the dark matter, baryons, photons, and neutrinos including all relevant standard-model interactions (see, e.g., Refs. [4·35,4·36] and references therein). Since the perturbations are initially very small, linear perturbation theory is an excellent approximation; this allows us to solve the perturbation equations in Fourier space at each wavenumber k independently of all other wavenumbers (modes are uncoupled). The scattering of photons and baryons by DDM through the interaction $\mathcal{L}_{\gamma\chi}$ influences the growth of cosmological perturbations by introducing additional collision terms to the Boltzmann equations, which ultimately result in a drag force between the DDM and the colliding species in the equations describing the cosmological fluid (see, e.g., Refs. [4·37,4·38], which consider similar effects). Below we present the exact perturbation equations including the effects of dark matter with electric or magnetic dipole

moments. Since solutions to these equations are numerically intensive when photons and baryons are tightly coupled through Compton scattering, we also discuss the equations appropriate for solving for the DDM, photon, and baryon perturbations during the epoch of tight coupling.

In the synchronous gauge the equations describing the evolution of baryons, photons, and dark matter with an electric or magnetic dipole moment are,

$$\begin{aligned}
\dot{\delta}_\gamma &= -\frac{4}{3}\theta_\gamma - \frac{2}{3}\dot{h}\,, \qquad \dot{\delta}_b = -\theta_b - \frac{1}{2}\dot{h}\,, \qquad \dot{\delta}_\chi = -\theta_\chi - \frac{1}{2}\dot{h}\,, \\
\dot{\theta}_\gamma &= k^2\left(\frac{1}{4}\delta_\gamma - \Sigma_\gamma\right) + an_e\sigma_T(\theta_b - \theta_\gamma) + an_\chi\langle\sigma\rangle_{\chi\gamma}(\theta_\chi - \theta_\gamma), \\
\dot{\theta}_b &= -\frac{\dot{a}}{a}\theta_b + c_{sb}^2 k^2\delta_b + \frac{4\rho_\gamma}{3\rho_b}an_e\sigma_T(\theta_\gamma - \theta_b) + an_\chi\langle\sigma v\rangle_{\chi b}(\theta_\chi - \theta_b), \\
\dot{\theta}_\chi &= -\frac{\dot{a}}{a}\theta_\chi + c_{s\chi}^2 k^2\delta_\chi + \frac{\rho_b}{\rho_\chi}an_\chi\langle\sigma v\rangle_{\chi b}(\theta_b - \theta_\chi) + \frac{4\rho_\gamma}{3\rho_\chi}an_\chi\langle\sigma\rangle_{\chi\gamma}(\theta_\gamma - \theta_\chi). \qquad (4.17)
\end{aligned}$$

While the evolution equations for the density contrast $\delta_j = \delta\rho_j/\rho_j$ for each species $j \in \{\gamma, b, \chi\}$ are as in the standard case [4·35], as discussed above, the evolution equations for the fluid-velocity perturbations have additional drag-force terms due to the photon-DDM interaction. Note that in these equations and what follows the variable $\theta_j = ikV_j$ is the divergence of the fluid velocity in Fourier space, Σ_j is the shear, c_{sj} is the intrinsic sound speed, and n_j and ρ_j are the background number and energy densities of a particular species j, respectively. The variable h is the trace of the scalar metric perturbation in Fourier space (not to be confused with the Hubble parameter), a is the cosmological scale factor, and an overdot represents a derivative with respect to the conformal time τ. Furthermore, σ_T is the Thomson cross section, while

$$\langle\sigma\rangle_{\chi\gamma} = \frac{80}{21}\pi(\mathcal{D}^4 + \mathcal{M}^4)T_\gamma^2, \qquad (4.18)$$

is the appropriately thermally-averaged DDM-photon cross section, which can be obtained from the differential cross section [4·39, 4·40],

$$\frac{d\sigma_{\chi\gamma}}{d\Omega} = \frac{(\mathcal{D}^4 + \mathcal{M}^4)E_\gamma^2}{8\pi^2}(3 - \cos^2\theta), \qquad (4.19)$$

for photon-DDM scattering. As argued above, the photon-DDM drag term is small, and we consider it no further in Eq. (4.17).

The quantity

$$\langle\sigma v\rangle_{\chi b} = \frac{4(1 + \xi Y)}{3\pi^2\sqrt{\langle v_p\rangle^2 + \langle v_\chi\rangle^2}}\frac{m_\chi}{m_\chi + m_{\rm p}}e^2(\mathcal{D}^2 + \mathcal{M}^2), \qquad (4.20)$$

is the appropriate thermally-averaged cross section times relative velocity for the baryon-DDM

coupling, and

$$\xi = 8\frac{m_\chi + m_{\rm p}}{m_\chi + 4m_{\rm p}}\sqrt{\frac{\langle v_p\rangle^2 + \langle v_\chi\rangle^2}{\langle v_p\rangle^2 + 4\langle v_\chi\rangle^2}} - 1 \tag{4.21}$$

is the relative efficiency for coupling to helium nuclei compared to protons. Appendix A provides a derivation of this collision coefficient. In these expressions, $Y = \rho_{He}/\rho_b \simeq 0.24$ is the cosmological helium mass fraction (approximating $m_{He} \simeq 4m_p$), $\langle v_p\rangle = \sqrt{8T_b/\pi m_p}$ is the average thermal speed of the protons, $\langle v_\chi\rangle = \sqrt{8T_\chi/\pi m_\chi}$ is the average thermal speed of the DDM, and T_γ, T_b, and T_χ are the photon, baryon, and DDM temperatures respectively. The dark-matter temperature evolves according to

$$\dot{T}_\chi = -2\frac{\dot{a}}{a}T_\chi + \frac{2a\rho_b\langle\sigma v\rangle^{\rm T}_{\chi b}}{m_\chi + m_p}(T_b - T_\chi) + \frac{8a\rho_\gamma\langle\sigma\rangle_{\chi\gamma}}{3m_\chi}(T_\gamma - T_\chi), \tag{4.22}$$

where $\langle\sigma v\rangle^{\rm T}_{\chi b}$ is the same as the expression given in Eq. (4.20) with the replacement of ξ by $\xi^{\rm T}$ which is given by the expression in Eq. (4.21) with the factor $(m_\chi + m_p)/(m_\chi + 4m_p)$ replaced by $[(m_\chi + m_p)/(m_\chi + 4m_p)]^2$. The final term, describing the dark matter heating by photons, is important at very early times. For the dipole strength and mass range considered, the influence is manifest only on very small length scales, below the range of interest.

At early times, the DDM temperature $T_\chi \simeq T_b$, but at later times, when the DDM decouples, T_χ drops relative to T_b. The DDM-proton cross section is $\propto v^{-2}$, which leads to $\langle\sigma v\rangle_{\chi b} \propto \langle v_p\rangle^{-1}$. As a result, we cannot directly apply the results of Ref. [4·37], wherein a velocity-independent dark matter-baryon interaction was assumed. However, we have verified that we recover their results if we take a velocity-independent cross section as the source of dark-matter–baryon drag.

4.6.2 Tightly Coupled Equations

At early times when $\tau_c^{-1} \equiv an_e\sigma_T \gg \dot{a}/a$ the rapid scattering of baryons and photons forces these species to have nearly equal fluid velocities, and consequently the solution of the equations shown in Eq. (4.17) is numerically intensive. Following standard procedures [4·35, 4·41] we derive a set of equations to leading order in the (conformal) Compton scattering time τ_c that are appropriate for evolving the fluid variables through this epoch of tight coupling. We first write down an equation for the time derivative of $\theta_b - \theta_\gamma$ which is usually termed the baryon-photon 'slip' to leading order in τ_c,

$$\begin{aligned}\dot{\theta}_b - \dot{\theta}_\gamma = \frac{2R_{\gamma b}}{1+R_{\gamma b}}\frac{\dot{a}}{a}(\theta_b - \theta_\gamma) + \frac{\tau_c}{1+R_{\gamma b}}\Bigg[-\frac{\ddot{a}}{a}\theta_b + k^2\left(c_{sb}^2\dot{\delta}_b - \frac{1}{4}\dot{\delta}_\gamma - \frac{1}{2}\frac{\dot{a}}{a}\delta_\gamma\right)\\ + \frac{1}{\tau_\chi}(\dot{\theta}_\chi - \dot{\theta}_b) - \frac{1}{2}\frac{\dot{a}}{a}\frac{1}{\tau_\chi}(\theta_\chi - \theta_b)\Bigg],\end{aligned} \tag{4.23}$$

where we have introduced the (conformal) DDM-baryon scattering time $\tau_\chi^{-1} \equiv a n_\chi \langle \sigma v \rangle_{\chi b}$, and $R_{\gamma b} \equiv 4\rho_\gamma/(3\rho_b)$. It is useful to separate this equation as a sum of the terms not containing τ_χ (this is just the time derivative of the standard slip, which we denote $S_{b\gamma}$), and the new terms introduced by the DDM coupling,

$$\dot{\theta}_b - \dot{\theta}_\gamma = \dot{S}_{b\gamma} + \beta \left[(\dot{\theta}_\chi - \dot{\theta}_b) - \frac{1}{2}\frac{\dot{a}}{a}(\theta_\chi - \theta_b) \right], \tag{4.24}$$

where

$$\beta = \frac{1}{1 + R_{\gamma b}} \frac{\tau_c}{\tau_\chi} \tag{4.25}$$

is the parameter that controls how strongly the new interaction affects the evolution of the slip. In terms of these definitions the baryon-velocity evolution equation is

$$\dot{\theta}_b = \frac{1}{1 + R_{\gamma b} + \beta R_{\gamma b}} \left\{ -\frac{\dot{a}}{a}\theta_b + k^2 \left[c_{sb}^2 \delta_b + R_{\gamma b} \left(\frac{1}{4}\delta_\gamma - \Sigma_\gamma \right) \right] \right. \\ \left. + R_{\gamma b} \left[\dot{S}_{b\gamma} + \beta \left(\dot{\theta}_\chi - \frac{1}{2}\frac{\dot{a}}{a}(\theta_\chi - \theta_b) \right) \right] \right\}. \tag{4.26}$$

The photon-evolution equation is then given by the exact expression

$$\dot{\theta}_\gamma = -\frac{1}{R_{\gamma b}} \left[\dot{\theta}_b + \frac{\dot{a}}{a}\theta_b - c_{sb}^2 k^2 \delta_b - \frac{1}{\tau_\chi}(\theta_\chi - \theta_b) \right] + k^2 \left(\frac{1}{4}\delta_\gamma - \Sigma_\gamma \right). \tag{4.27}$$

We use these equations to follow the initial evolution of the baryon and photon fluid variables and switch to the exact equations of Eq. (4.17) at later times. For the evolution of the DDM fluid variables we always use the exact form of Eq. (4.17).

4.6.3 Effects on the Matter and CMB Power Spectra

In Fig. 4.12 we show the linear matter power spectrum and in Fig. 4.13 we show the angular power spectrum of the CMB for several values of the dipole moment and for DDM mass $m_\chi = 1$ GeV. Physically, the effects of DDM on the matter power spectrum and CMB can be simply understood. Prior to matter-radiation equality the photons have a much larger density than the baryons or the DDM and so to a first approximation completely drive the behavior of the baryon perturbations through Compton scattering. In turn, the baryon perturbations drive the behavior of the DDM perturbations, very efficiently before DDM decoupling so that the DDM density contrast δ_χ on scales that enter the horizon during this epoch track the oscillations of the baryon-photon fluid before growing, and less efficiently after DDM decoupling so that the baryons simply cause a drag on the growth of δ_χ. In either case, the matter power spectrum is suppressed relative to the standard

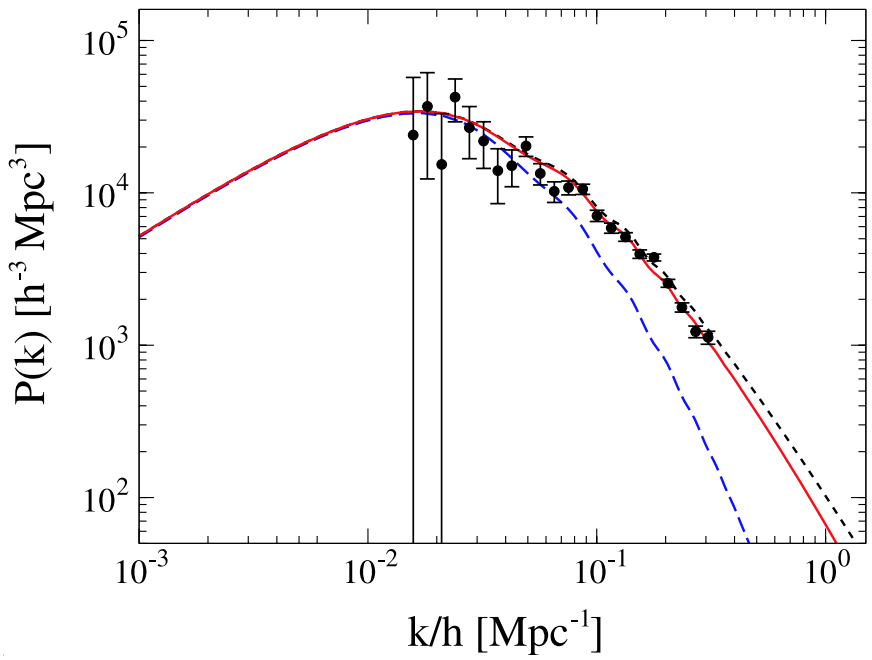

Figure 4.12: Matter power spectra including baryon-DDM drag. The solid (red) curve is for is for $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} = 1.4 \times 10^{-15}\, e$ cm The short-dash (black) is for $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} = 1.0 \times 10^{-16}\, e$ cm The long-dash (blue) curve is for is for $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} = 5 \times 10^{-15}\, e$ cm. These are all for a mass of 1 GeV. The curves are all for the standard concordance cosmological parameters, and the data points are from SDSS [4·43]

case. The behavior of the CMB angular power spectrum can be similarly understood. Roughly speaking, the coupling of the DDM and baryons increases the effective baryon loading of the plasma at early times so that the CMB power spectra look similar to those from high-baryon models. This is of course an imperfect correspondence as modes of larger wavelength enter the horizon when the coupling is weaker, and so at later and later times the evolution of the photon perturbations becomes more and more like the standard-CDM case. But due to geometrical projection effects modes of wavenumber k contribute to all $l \lesssim k d_A$ where d_A is the angular-diameter distance to the last-scattering surface, and so the effects of DDM on small length scales can be noticed even on relatively large angular scales in the CMB.

As the effect of DDM on the CMB may be partially degenerate with other cosmological parameters, we have explored a parameter space that allows us to constrain m_χ and $(\mathcal{D}, \mathcal{M})$ after marginalizing over other cosmological parameters. We consider flat ΛCDM models and our chosen parameter space is the dark-matter density $\Omega_m h^2$, the baryon density $\Omega_b h^2$, the Hubble parameter h in units of 100 km sec^{-1} Mpc^{-1}, the optical depth $\tau_{\rm CMB}$ to the last-scattering surface, and the primordial spectral index n. We have employed the Markov chain Monte Carlo technique (see, e.g., Ref. [4·42]) to efficiently explore this parameter space, taking the most recent results from SDSS [4·43], WMAP [4·44], CBI [4·45], VSA [4·46], and SNe Ia [4·47] as our data. Note that

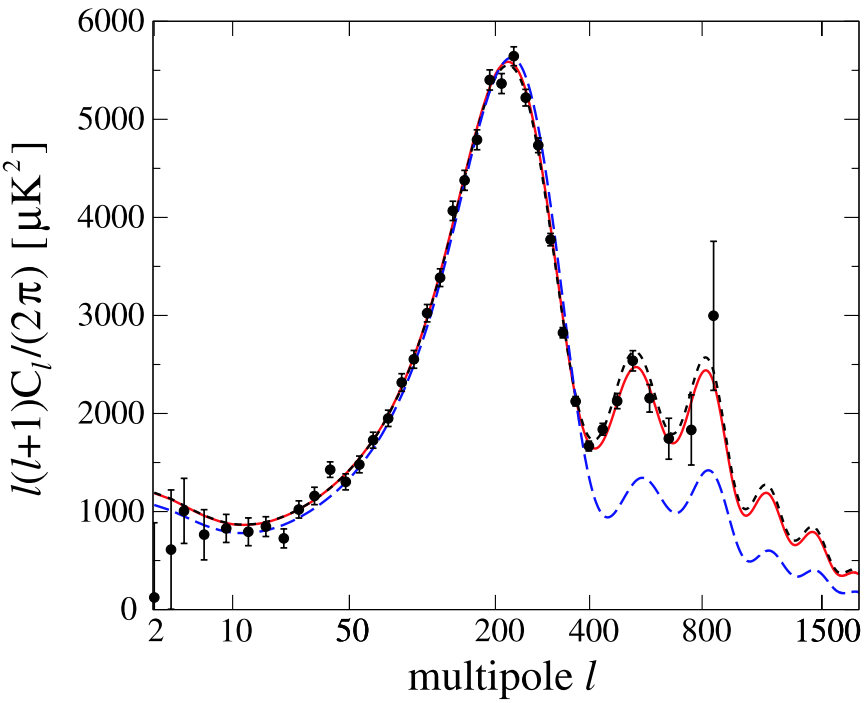

Figure 4.13: CMB power spectra including DDM-baryon drag. The labeling of the curves is the same as in Fig. 4.12, and the data points are those from WMAP [4·44]

although DDM has no effect on observations of Type Ia supernovae, we include these data because the other parameters we allow to vary are constrained by these observations. We conclude using a relative-likelihood test that cosmological measurements lead to the bound shown in Fig. 4.1. The numerical calculations confirm the qualitative behavior discussed above. Dipole moments as large as $(\mathcal{D}^2 + \mathcal{M}^2)^{1/2} \sim 10^{-17}$ e cm, near the upper end of our allowed parameter space, are thus cosmologically viable.

4.7 Gamma Rays

DDM particles in the Galactic halo can annihilate to two photons through the diagrams shown in Fig. 4.3. Since halo particles move with velocities $v \simeq 300$ km sec$^{-1} \ll c$, the photons produced will be very nearly monoenergetic with energies equal to the DDM-particle rest mass. The intensity at Earth of such gamma rays is obtained by integrating the emissivity, $n_\chi^2 \langle \sigma_{\bar{\chi}\chi \to 2\gamma} v \rangle$, where n_χ is the DDM number density, along the given line of sight. The intensity is largest toward the Galactic center, where the dark-matter density is largest. In this direction, the gamma-ray intensity is then [4·4],

$$\frac{d\mathcal{F}}{d\Omega} = 1.0 \times 10^{-10} \frac{\sigma_{\bar{\chi}\chi \to 2\gamma} v}{10^{-30}\,\mathrm{cm}^3\ \mathrm{sec}^{-1}} m_{\mathrm{GeV}}^{-2} I\ \mathrm{cm}^{-2}\ \mathrm{sec}^{-1}\ \mathrm{sr}^{-1}, \tag{4.28}$$

where I is a scaled integral of n_χ^2 along a line of sight toward the Galactic center. The numerical coefficient is one-half that from Ref. [4·4] since we have here particle-antiparticle annihilation rather than Majorana annihilation. Roughly speaking, $I \simeq 3 - 30$ for cored-isothermal-sphere models of the Galactic halo, while I can extend up to ~ 300 for Navarro-Frenk-White profiles [4·48]; i.e., uncertainty in the dark-matter distribution in the inner Galaxy leads to an uncertainty of two orders of magnitude in the predicted flux. We thus expect

$$\frac{d\mathcal{F}}{d\Omega} = (3 \times 10^{-13} - 3 \times 10^{-11})(\mathcal{D}_{17}^4 + \mathcal{M}_{17}^4)\,\mathrm{cm}^{-2}\ \mathrm{sec}^{-1}\ \mathrm{sr}^{-1}. \tag{4.29}$$

To constrain dipole moments from non-observation of a gamma-ray line, we choose to use the most conservative estimate, $I \simeq 3$ for the dimensionless line integral. Moreover, we are not aware of any EGRET analysis that places limits in particular to a line flux. We thus obtain very conservative limits by using the binned continuum fluxes for the *total* gamma-ray flux listed in Table 1 of Ref. [4·48] and noting that a line flux in that bin cannot exceed the measured continuum flux. The EGRET limits apply for masses $0.1 \lesssim (m_\chi/\mathrm{GeV}) \lesssim 10$, and range from $\mathcal{D}_{17} \lesssim 180$ for $m_\chi \lesssim$ GeV to $\mathcal{D}_{17} \lesssim 100$ for $m_\chi \simeq 10$ GeV, as shown in Fig. 4.1.

Again, a few caveats are in order. First of all, our limit is quantitatively conservative—we chose the halo model that produces the lowest flux, and a detailed EGRET analysis would probably yield a line-flux limit lower than what we have assumed. On the other hand, the strong dependence $\propto \mathcal{D}^4$ of the predicted flux on the dipole moment guarantees that the upper limit to the acceptable dipole moment will not depend quite so strongly on these details. Second, if $\mathcal{D}_{17} \gtrsim 5$ in the mass range 100 MeV to 1 GeV, then the correct cosmological abundance most likely requires a particle-antiparticle asymmetry. If so, then the annihilation rate in the halo could be reduced far below the values we have obtained above. We conclude by noting that with the increased sensitivity of the Gamma-Ray Large Area Space Telescope (GLAST), a detailed search for a line flux, and the possibility that the actual halo model provides a more generous annihilation rate, an observable GLAST signature may exist for masses $0.1 - 1$ GeV and dipole moments as low as $\mathcal{D}_{17} \sim 10$.

4.8 Discussion

In this chapter we have considered the cosmology and phenomenology of dark-matter particles with a nonzero magnetic or electric dipole moment. We have found that information from precision tests of the standard model, direct dark-matter searches, gamma-ray experiments, and the CMB and large-scale structure restrict the dipole moment to be $(\mathcal{D},\, \mathcal{M}) \lesssim 3 \times 10^{-16}\, e$ cm for masses $m_\chi \lesssim$ few GeV and $(\mathcal{D},\, \mathcal{M}) \lesssim 10^{-24}\, e$ cm for larger masses. (This improves on an earlier limit on WIMP electromagnetic dipole moments based on direct detection [4·49]). Some of the allowed regions of

parameter space may soon be probed with GLAST and with future more sensitive direct-detection experiments. The electromagnetic interactions of these particles with nuclei are coherent. Moreover, these particles cannot annihilate directly to neutrinos. Therefore, searches for energetic neutrinos from decays of the products of $\bar{\chi}\chi$ annihilation in the Sun or Earth are thus likely to provide less sensitive probes than direct searches [4·50]. Moreover, if there is a particle-antiparticle asymmetry, then the energetic-neutrino flux could be reduced without altering the direct-detection rate.

We have restricted our attention to particles with masses $m_\chi \gtrsim$ MeV, with the notion that lower-mass particles will violate BBN limit, as discussed toward the end of Section 4.2. We also consider masses $m_\chi \gtrsim$ MeV, as particles of lower mass will almost certainly undergo relativistic freezeout and thus lead to unacceptable dark-matter candidates. However, as also noted above that if an $m_\chi \lesssim$ MeV particle has a dipole moment $(\mathcal{D}, \mathcal{M}) \lesssim 10^{-22}\, e$ cm and no other interactions with ordinary matter, then it might still be consistent with BBN. Of course, such a particle will, assuming standard freezeout, have a mass density many orders of magnitude larger than the dark-matter density. But suppose we were to surmise that the dark-matter density was fixed by some other mechanism, e.g., suppose the dipole was sufficiently weak that it never came into equilibrium. In this case, an additional constraint to the dark-matter dipoles can be obtained from energy-loss arguments applied to stars in globular clusters. Such arguments eliminate dipole moments $(\mathcal{D}, \mathcal{M}) \lesssim 6 \times 10^{-23}\, e$ cm for masses $m_\chi \lesssim 5$ keV [4·51]. We have also considered constraints from astrophysical phenomena such as the stability of the Galactic disk, lifetime of compact objects, and annihilations in the solar neighborhood [4·6], and find that these constraints on the mass and interaction strengths are not competitive with those presented here.

It would be of interest to attempt to embed this scenario in a consistent particle physics model. We might find links between baryonic and non-baryonic matter abundances, the dark matter electric dipole moment and the CP violation needed for baryogenesis, and the magnetic moments of dark matter and baryons (e.g., Ref. [4·52] considered a fermionic technibaryon with electromagnetic dipole interactions as a dark matter candidate). However, such model building is beyond the scope of the present study. Our approach throughout has been entirely phenomenological, as we have been motivated by the desire to answer the question, "How dark is 'dark'?"

Acknowledgments

We thank S. Golwala, J. Albert, M. Zaldarriaga, and R. Shrock for useful discussions. KS acknowledges the support of a Canadian NSERC Postgraduate Scholarship. This work was supported in part by NASA NAG5-9821 and DoE DE-FG03-92-ER40701 (at Caltech) and NSF PHY-0099543 (at Dartmouth).

Appendix 4.A: The Baryon-Dark Matter Collision Term

To determine how the cosmological perturbation equations for baryons and dark matter are altered when we bestow the dark matter with a magnetic or electric dipole moment, we must formally evaluate the collision operator of the general-relativistic Boltzmann equation in a given gauge [4·14, 4·35, 4·36] for the dipole interaction of Eq. (4.1). We have completed this calculation in detail, and find that the dipole interaction produces a drag force proportional to the relative velocity $V = v_\chi - v_b$ of the dark-matter and baryon fluids. As the relative velocity is gauge invariant in linear perturbation theory and all scatterings are local processes, we may thus take a simpler, more physically transparent approach and just evaluate this drag force using nonrelativistic statistical mechanics. It is this approach we now present.

We wish to calculate the drag force per unit mass, or deceleration, due to collisions with protons to the dark-matter fluid as it passes through the baryon-photon fluid. Comoving scales $\lambda \gtrsim$ Mpc enter the horizon when the cosmological temperature is $T \lesssim 10$ eV, when the DDM particles (which are restricted to $m_\chi \gtrsim$ MeV) are nonrelativistic. We may thus consider thermal velocity distributions for nonrelativistic baryons and dark matter. Since the drag force can only depend on the dark-matter–baryon relative velocity, we take the baryon fluid to be at rest and the dark-matter fluid to have a velocity of magnitude V in the $\hat{x}$ direction. Then, the proton phase-space distribution is

$$f_p(\vec{v}_p) = \frac{n_p}{(2\pi)^{3/2}\bar{v}_p^3} e^{-v_p^2/2\bar{v}_p^2}, \tag{4.30}$$

where $\bar{v}_p = (kT_p/m_p)^{1/2}$ is the proton velocity dispersion and n_p the proton number density, and

$$f_\chi(\vec{v}_\chi) = \frac{n_\chi}{(2\pi)^{3/2}\bar{v}_\chi^3} \exp\left[-\frac{(\vec{v}_\chi - V\hat{x})^2}{2\bar{v}_\chi^2}\right], \tag{4.31}$$

is the dark-matter phase-space distribution, with $\bar{v}_\chi = (kT/m_\chi)^{1/2}$. Recall also that we expect $V \ll \bar{v}_p$, as discussed above.

The drag force per unit mass is obtained by integrating the momentum transfer per collision over all collisions between protons and dark-matter particles. From the symmetry of the problem, the deceleration of the dark-matter fluid will be in the $\hat{x}$ direction, and it will have a magnitude,

$$a_x = \frac{1}{n_\chi}\int d^3v_\chi f_\chi(\vec{v}_\chi)\int d^3v_p f_p(\vec{v}_p)|\vec{v}_p - \vec{v}_\chi|\int d\Omega \frac{d\sigma}{d\Omega}(v_{\chi xf} - v_{\chi xi}). \tag{4.32}$$

Here $\Omega = (\theta, \phi)$ is the scattering angle in the center-of-mass frame, and $v_{\chi xf} - v_{\chi xi}$ is the difference between the final and initial x component of the dark-matter–particle velocity; the difference is the same in the center-of-mass and laboratory frames. The differential cross section $d\sigma/d\Omega$ is that given in Eq. (4.5).

Consider an individual scattering event. Let α be the angle that $\vec{v}_p - \vec{v}_\chi$ makes with the $\hat{x}$ direction; this is then the angle that $\vec{v}_\chi$ makes with the $\hat{x}$ axis in the center-of-mass frame, and the magnitude of the initial and final dark-matter velocities in the center-of-mass frame is $v_\chi^{\rm cm} = m_p v/(m_p + m_\chi)$, where $v \equiv |\vec{v}_p - \vec{v}_\chi|$ is the relative velocity. The initial $\hat{x}$ component of the dark-matter velocity in the center-of-mass frame is then $v_{\chi xi} = v_\chi^{\rm cm} \cos\alpha$. The scattering angles θ and ϕ are then the polar and azimuthal angles that the scattered dark-matter velocity make with the initial velocity in the center-of-mass frame. By rotating this coordinate system by an angle α about the $\hat{z}$ axis to align it with the laboratory $\hat{x}$ axis, we find $v_{\chi xf} = v_\chi^{\rm cm}(\cos\alpha\cos\theta - \sin\alpha\sin\theta\sin\phi)$. Thus,

$$\int d\Omega \frac{d\sigma}{d\Omega}(v_{\chi xf} - v_{\chi xi}) = -\frac{m_p Z^2 e^2 (\mathcal{D}^2 + \mathcal{M}^2)\cos\alpha}{2\pi(m_p + m_\chi)v}. \tag{4.33}$$

Completing the integral in Eq. (4.32) in the limit $V \ll (\bar{v}_p, \bar{v}_\chi)$, we find

$$a_x = \frac{2(Ze)^2(\mathcal{D}^2 + \mathcal{M}^2) m_p n_p}{3\pi(2\pi)^{1/2}(m_p + m_\chi)} \frac{V}{\sqrt{\bar{v}_p^2 + \bar{v}_\chi^2}}. \tag{4.34}$$

Taking into account the definition $\langle v_p \rangle = \sqrt{8T_p/(\pi m_p)}$, this drag force leads to the drag-force term in Eqs. (4.17), (4.20), and (4.21), when including the simple corrections for a mass fraction Y of helium.

Bibliography

[4·1] F. Zwicky, Helv. Phys. Acta **6**, 110 (1933).

[4·2] S. Tremaine and J. E. Gunn, Phys. Rev. Lett. **42**, 407 (1979).

[4·3] S. Eidelman *et al.* [Particle Data Group], Phys. Lett. B **592**, 1 (2004).

[4·4] G. Jungman, M. Kamionkowski and K. Griest, Phys. Rept. **267**, 195 (1996).

[4·5] M. S. Turner, Phys. Rept. **197**, 67 (1990); G. G. Raffelt, Phys. Rept. **198**, 1 (1990); L. J. Rosenberg and K. A. van Bibber, Phys. Rept. **325**, 1 (2000).

[4·6] G. D. Starkman, A. Gould, R. Esmailzadeh and S. Dimopoulos, Phys. Rev. D **41**, 3594 (1990).

[4·7] E. D. Carlson, M. E. Machacek and L. J. Hall, Astrophys. J. **398**, 43 (1992).

[4·8] D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84**, 3760 (2000).

[4·9] A. Gould, B. T. Draine, R. W. Romani and S. Nussinov, Phys. Lett. B **238**, 337 (1990).

[4·10] S. Davidson, S. Hannestad and G. Raffelt, JHEP **0005**, 003 (2000).

[4·11] S. L. Dubovsky, D. S. Gorbunov and G. I. Rubtsov, JETP Lett. **79**, 1 (2004) [Pisma Zh. Eksp. Teor. Fiz. **79**, 3 (2004)].

[4·12] E. Fermi and E. Teller, Phys. Rev. **72**, 399 (1947).

[4·13] S. Burles, K. M. Nollett, J. N. Truran and M. S. Turner, Phys. Rev. Lett. **82**, 4176 (1999).

[4·14] E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, 1990).

[4·15] P. de Bernardis *et al.* [Boomerang Collaboration], Nature **404**, 955 (2000); S. Hanany *et al.*, Astrophys. J. **545**, L5 (2000); N. W. Halverson *et al.*, Astrophys. J. **568**, 38 (2002); B. S. Mason *et al.*, Astrophys. J. **591**, 540 (2003); A. Benoit *et al.* [the Archeops Collaboration], Astron. Astrophys. **399**, L25 (2003); D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148**, 175 (2003);

[4·16] K. Griest and M. Kamionkowski, Phys. Rev. Lett. **64**, 615 (1990).

[4·17] M. Kamionkowski and M. S. Turner, Phys. Rev. D **42**, 3310 (1990).

[4·18] L. L. Foldy, Phys. Rev. **83**, 688 (1951).

[4·19] D. S. Akerib *et al.* [CDMS Collaboration], Phys. Rev. Lett. **93**, 211301 (2004).

[4·20] D. S. Akerib *et al.* [CDMS Collaboration], Phys. Rev. D **68**, 082002 (2003).

[4·21] G. Angloher *et al.*, Astropart. Phys. **18**, 43 (2002).

[4·22] P. C. McGuire and P. J. Steinhardt, arXiv:astro-ph/0105567.

[4·23] J. Rich, R. Rocchia and M. Spiro, Phys. Lett. B **194**, 173 (1987).

[4·24] D. McCammon *et al.*, Astrophys. J. **576**, 188 (2002).

[4·25] G. W. Bennett *et al.* [Muon g-2 Collaboration], Phys. Rev. Lett. **89**, 101804 (2002) [Erratum-ibid. **89**, 129903 (2002)].

[4·26] K. Hagiwara *et al.* [Particle Data Group], Phys. Rev. D **66**, 010001 (2002).

[4·27] C. Bird, P. Jackson, R. Kowalewski and M. Pospelov, Phys. Rev. Lett. **93**, 201803 (2004).

[4·28] B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0304020.

[4·29] T. E. Browder *et al.* [CLEO Collaboration], Phys. Rev. Lett. **86**, 2950 (2001).

[4·30] K. Abe *et al.* [BELLE Collaboration], Phys. Rev. Lett. **88**, 021801 (2002).

[4·31] B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0107026.

[4·32] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **88**, 041801 (2002); P. Achard *et al.* [L3 Collaboration], Phys. Lett. B **587**, 16 (2004); D. Acosta [CDF Collaboration], Phys. Rev. Lett. **92**, 121802 (2004).

[4·33] S. Davidson, B. Campbell and D. C. Bailey, Phys. Rev. D **43**, 2314 (1991).

[4·34] A. A. Prinz *et al.*, Phys. Rev. Lett. **81**, 1175 (1998).

[4·35] C. P. Ma and E. Bertschinger, Astrophys. J. **455**, 7 (1995).

[4·36] S. Dodelson, *Modern Cosmology* (Academic Press, Amsterdam, 2003).

[4·37] X. l. Chen, S. Hannestad and R. J. Scherrer, Phys. Rev. D **65**, 123515 (2002).

[4·38] K. Sigurdson and M. Kamionkowski, Phys. Rev. Lett. **92**, 171302 (2004).

[4·39] F. E. Low, Phys. Rev. **96**, 1428 (1954).

[4·40] M. Gell-Mann and M. L. Goldberger, Phys. Rev. **96**, 1433 (1954).

[4·41] P. J. E. Peebles and J. T. Yu, Astrophys. J. **162**, 815 (1970).

[4·42] M. Doran and C. M. Mueller, JCAP **0409**, 003 (2004).

[4·43] M. Tegmark *et al.* [SDSS Collaboration], Astrophys. J. **606**, 702 (2004).

[4·44] G. Hinshaw *et al.*, Astrophys. J. Suppl. **148**, 135 (2003).

[4·45] A. C. S. Readhead *et al.*, Astrophys. J. **609**, 498 (2004).

[4·46] C. Dickinson *et al.*, Mon. Not. Roy. Astron. Soc. **353**, 732 (2004).

[4·47] J. L. Tonry *et al.* [Supernova Search Team Collaboration], Astrophys. J. **594**, 1 (2003); R. A. Knop *et al.* [The Supernova Cosmology Project Collaboration], Astrophys. J. **598**, 102 (2003); B. J. Barris *et al.*, Astrophys. J. **602**, 571 (2004); A. G. Riess *et al.* [Supernova Search Team Collaboration], Astrophys. J. **607**, 665 (2004).

[4·48] L. Bergstrom, P. Ullio and J. H. Buckley, Astropart. Phys. **9**, 137 (1998).

[4·49] M. Pospelov and T. ter Veldhuis, Phys. Lett. B **480**, 181 (2000).

[4·50] M. Kamionkowski, K. Griest, G. Jungman and B. Sadoulet, Phys. Rev. Lett. **74**, 5174 (1995).

[4·51] G. G. Raffelt, *Stars as Laboratories for Fundamental Physics* (University of Chicago Press, Chicago, 1996).

[4·52] J. Bagnasco, M. Dine and S. Thomas, Phys. Lett. B **320**, 99 (1994).

Chapter 5

A Running Spectral Index in Supersymmetric Dark-Matter Models with Quasi-Stable Charged Particles

We show that charged-particles decaying in the early Universe can induce a scale-dependent or 'running' spectral index in the small-scale linear and nonlinear matter power spectrum and discuss examples of this effect in minimal supersymmetric models in which the lightest neutralino is a viable cold-dark-matter candidate. We find configurations in which the neutralino relic density is set by coannihilations with a long-lived stau, and the late decay of staus partially suppresses the linear matter power spectrum. Nonlinear evolution on small scales then causes the modified linear power spectrum to evolve to a nonlinear power spectrum similar (but different in detail) to models parametrized by a constant running $\alpha_s = dn_s/d\ln k$ by redshifts of 2 to 4. Thus, Lyman-α forest observations, which probe the matter power spectrum at these redshifts, might not discriminate between the two effects. However, a measurement of the angular power spectrum of primordial 21-cm radiation from redshift $z \approx 30$–200 might distinguish between this charged-decay model and a primordial running spectral index. The direct production of a long-lived charged particle at future colliders is a dramatic prediction of this model.

Originally published as S. Profumo, K. Sigurdson, P. Ullio and M. Kamionkowski, *Phys. Rev. D* **71**, 023518 (2005).

5.1 Introduction

While recent cosmological observations provide convincing evidence that nonbaryonic dark matter exists [5·1], we do not know the detailed particle properties of the dark matter, nor the particle spectrum of the dark sector. There has been considerable phenomenological effort towards placing

model-independent limits on the possible interactions of the lightest dark-matter particle (LDP)[1] in an attempt to try and identify candidates within detailed particle theories or rule out particular candidate theories. For instance, models with stable charged dark matter have been ruled out [5·2], while significant constraints have been made to dark-matter models with strong interactions [5·3], self-interactions [5·4, 5·5], and a millicharge [5·6, 5·7]. Recently, it was shown that a neutral dark-matter particle with a relatively large electric or magnetic dipole moment remains a phenomenologically viable candidate [5·8].

Concurrent with this phenomenological effort, theorists have taken to considering physics beyond the standard model in search of a consistent framework for viable dark-matter candidates. The leading candidates, those that produce the correct relic abundance and appear in minimal extensions of the standard model (often for independent reasons), are the axion [5·9] and weakly interacting massive particles (WIMPs), such as the neutralino, the lightest mass eigenstate from the superposition of the supersymmetric partners of the $U(1)$ and $SU(2)$ neutral gauge bosons and of the neutral Higgs bosons [5·10, 5·11]. However, other viable candidates have also been considered recently, such as gravitinos or Kaluza-Klein gravitons produced through the late decay of WIMPs [5·12, 5·13]. These latter candidates are an interesting possibility because the constraints to the interactions of the LDP do not apply to the next-to-lightest dark-matter particle (NLDP), and the decay of the NLDP to the LDP at early times may produce interesting cosmological effects; for instance, the reprocessing of the light-element abundances formed during big-bang nucleosynthesis [5·14–5·16], or if the NLDP is charged, the suppression of the matter power spectrum on small scales and thus a reduction in the expected number of dwarf galaxies [5·17].

In this chapter we describe another effect charged NLDPs could have on the matter power spectrum. If all of the present-day dark matter is produced through the late decay of charged NLDPs, then, as discussed in Ref. [5·17], the effect is to essentially cut off the matter power spectrum on scales that enter the horizon before the NLDP decays. However, if only a fraction f_ϕ of the present-day dark matter is produced through the late decay of charged NLDPs, the matter power spectrum is suppressed on small scales only by a factor $(1 - f_\phi)^2$. This induces a scale-dependent spectral index for wavenumbers that enter the horizon when the age of the Universe is equal to the lifetime of the charged particles. What we show below is that, for certain combinations of f_ϕ and of the lifetime of the charged particle τ, this suppression modifies the nonlinear power spectrum in a way similar (but different in detail) to the effect of a constant $\alpha_s \equiv dn_s/d\ln k \neq 0$. Although these effects are different, constraints based on observations that probe the nonlinear power spectrum at redshifts of 2 to 4, such as measurements of the Lyman-α forest, might confuse a running index with the effect we describe here even if parametrized in terms of a constant α_s. This has significant

[1] In supersymmetric models the LDP is the lightest supersymmetric particle (LSP), but we adopt this more general notation unless we are speaking about a specific supersymmetric model.

implications for the interpretation of the detection of a large running of the spectral index as a constraint on simple single-field inflationary models. The detection of a unexpectedly large spectral running in future observations could instead be revealing properties of the dark-matter particle spectrum in conjunction with a more conventional model of inflation. We note that the Sloan Digital Sky Survey Lyman-α data [5·18] has significantly improved the limits to constant-α_s models compared to previous measurements alone [5·1,5·19,5·20]. A detailed study of the (τ,f_ϕ) parameter space using these and other cosmological data would also provide interesting limits to the models we discuss here.

While, even with future Lyman-α data, it may be difficult to discriminate the effect of a constant running of the spectral index from a scale-dependent spectral index due to a charged NLDP, other observations may nevertheless discriminate between the two scenarios. Future measurements of the power spectrum of neutral hydrogen through the 21cm-line might probe the linear matter power spectrum in exquisite detail over the redshift range $z \approx 30 - 200$ at comoving scales less than 1 Mpc and perhaps as small as 0.01 Mpc [5·21]; such a measurement could distinguish between the charged-particle decay scenario we describe here and other modifications to the primordial power spectrum. If, as in some models we discuss below, the mass of these particles is in reach of future particle colliders the signature of this scenario would be spectacular and unmistakable—the production of very long-lived charged particles that slowly decay to stable dark matter.

Although we describe the cosmological side of our calculations in a model-independent manner, remarkably, there are configurations in the minimal supersymmetric extension of the standard model (MSSM) with the right properties for the effect we discuss here. In particular, we find that if the LSP is a neutralino quasi-degenerate in mass with the lightest stau, we can naturally obtain, at the same time, LDPs providing the correct dark matter abundance $\Omega_\chi h^2 = 0.113$ [5·1] and NLDPs with the long lifetimes and the sizable densities in the early Universe needed in the proposed scenario. Such configurations arise even in minimal supersymmetric schemes, such as the minimal supergravity (mSUGRA) scenario [5·22] and the minimal anomaly-mediated supersymmetry-breaking (mAMSB) model [5·23]. This implies that a detailed study of the (τ,f_ϕ) parameter space using current and future cosmological data may constrain regions of the MSSM parameter space that are otherwise viable. Furthermore, we are able to make quantitative statements about testing the scenario we propose in future particle colliders or dark matter detection experiments.

The chapter is organized as follows: We first review in Section 5.2 how the standard calculation of linear perturbations in an expanding universe must be modified to account for the effects of a decaying charged species, calculate the linear matter power spectrum and discuss the constraints to this model from big bang nucleosynthesis (BBN) and the spectrum of the cosmic microwave background (CMB). In Section 5.3 we briefly discuss how we estimate the nonlinear power spectrum from the linear power spectrum and present several examples. In Section 5.4 we discuss how mea-

surements of the angular power spectra of the primordial 21-cm radiation can be used to distinguish this effect from other modifications to the primordial power spectrum. In Section 5.5 we describe how this scenario can be embedded in a particle physics model, concluding that the most appealing scheme is one where long lifetimes are obtained by considering nearly degenerate LDP and NLDP masses. In Section 5.6 we compute the lifetimes of charged next-to-lightest supersymmetric particles (NLSPs) decaying into neutralino LSPs, and show that, in the MSSM, the role of a NLDP with a long lifetime can be played by a stau only. In Section 5.7 we estimate what fraction of charged to neutral dark matter is expected in this case, while in Section 5.8 we describe how consistent realizations of this scenario can be found within the parameter space of mSUGRA and mAMSB. Finally, in Section 5.9 we discuss the expected signatures of this scenario at future particle colliders, such as the large hadron collider (LHC), and prospects for detection in experiments searching for WIMP dark matter. We conclude with a brief summary of our results in Section 5.10.

5.2 Charged-Particle Decay

In this section we discuss how the decay of a charged particle ϕ (the NLDP) to a neutral particle χ (the LDP) results in the suppression of the linear matter power spectrum on small scales.

As the ϕ particles decay to χ particles, their comoving energy density decays exponentially as

$$\rho_\phi a^3 = m_\phi n_{\phi_0} e^{-t/\tau}\,, \tag{5.1}$$

increasing the comoving energy density of χ particles as

$$\rho_\chi a^3 = m_\chi n_{\chi_0}(1 - f_\phi e^{-t/\tau})\,. \tag{5.2}$$

Here $n_{\chi_0} = \Omega_\chi \rho_{crit}/m_\chi$ is the comoving number density of dark matter, and $n_{\phi_0} = f_\phi n_{\chi_0}$ is the comoving number density of dark matter produced through the decay of ϕ particles, a is the scale factor, and t is the cosmic time.

Since the ϕ particles are charged, they are tightly coupled to the ordinary baryons (the protons, helium nuclei, and electrons) through Coulomb scattering. It is therefore possible describe the combined ϕ and baryon fluids as a generalized baryon-like component β as far as perturbation dynamics is concerned. We thus denote by $\rho_\beta = \rho_b + \rho_\phi$ the total charged-particle energy density at any given time. At late times, after nearly all ϕ particles have decayed, $\rho_\beta \simeq \rho_b$.

The relevant species whose perturbation dynamics are modified from the standard case are the stable dark matter (subscript χ), the charged species (subscript β), and the photons (subscript γ). By imposing covariant conservation of the total stress-energy tensor, accounting for the Compton scattering between the electrons and the photons, and linearizing about a Friedmann-Robertson-Walker

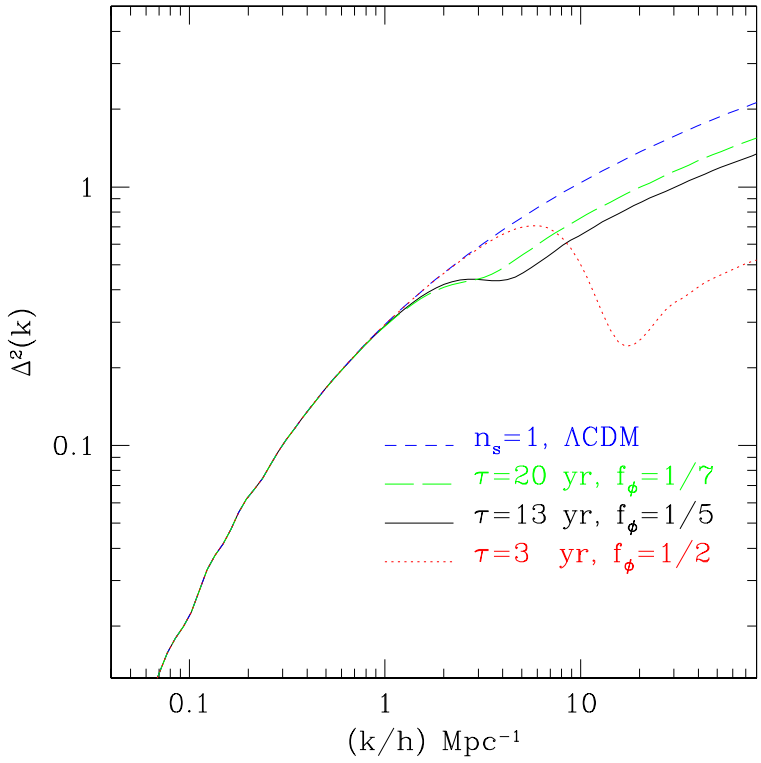

Figure 5.1: Shown is $\Delta^2(k) = k^3P(k)/2\pi^2$, the dimensionless matter power spectrum per logarithmic interval for the canonical $n_s = 1$ ΛCDM model (dashed), for a model with $\tau = 20$ yr and $f_\phi = 1/7$ (long-dashed), for a model with $\tau = 13$ yr and $f_\phi = 1/5$ (solid), and for a model with $\tau = 3$ yr and $f_\phi = 1/2$ (dotted). Modes that enter the horizon be before the ϕ particles decay are suppressed by a factor of $(1 - f_\phi)^2$.

(FRW) Universe we arrive at the equations describing the evolution of linear fluid perturbations of these components in an expanding Universe. In the synchronous gauge, for the 'β' component, the perturbation evolution equations are

$$\dot{\delta}_\beta = -\theta_\beta - \frac{1}{2}\dot{h}\,, \tag{5.3}$$

and

$$\dot{\theta}_\beta = -\frac{\dot{a}}{a}\theta_\beta + c_s^2 k^2 \delta_\beta + \frac{4\rho_\gamma}{3\rho_\beta} a n_e \sigma_T (\theta_\gamma - \theta_\beta)\,. \tag{5.4}$$

Here and in what follows, $\delta_X = \delta\rho_X/\rho_X$ is the fractional overdensity and $\theta_X = ikV_X$ is the divergence of the bulk velocity in Fourier space of a given species X. An overdot represents a derivative with respect to the conformal time. The number density of electrons is n_e, while σ_T is the Thomson

cross section. Because the ϕ particles and the baryons share a common bulk velocity and overdensity,[2] these equations are identical to the standard perturbation equations for the baryons with the replacement $b \to \beta$ (see, for example, Ref. [5·24]). For the dark matter we find that

$$\dot{\delta}_\chi = -\theta_\chi - \frac{1}{2}\dot{h} + \lambda_m \frac{\rho_\phi}{\rho_\chi}\frac{a}{\tau}(\delta_\beta - \delta_\chi)\,, \tag{5.5}$$

and

$$\dot{\theta}_\chi = -\frac{\dot{a}}{a}\theta_\chi + \lambda_m \frac{\rho_\phi}{\rho_\chi}\frac{a}{\tau}(\theta_\beta - \theta_\chi)\,, \tag{5.6}$$

where $\lambda_m \equiv m_\chi/m_\phi = (1 + \Delta m/m_\chi)^{-1}$. The modifications to the photon perturbation evolution are negligibly small because ϕ decays during the radiation-dominated epoch when $\rho_\phi \ll \rho_\gamma$ and, as discussed below, for viable models $\lambda_m \simeq 1$ to prevent unreasonably large spectral distortions to the CMB.

Combining these equations with the (unmodified) equations for the neutrino perturbations we can solve for the linear power spectrum of matter fluctuations in this model. We have solved these equations using a modified version of `CMBFAST` [5·25]. In Fig. 5.1 we show the linear matter power spectrum in this model for several values of the ϕ lifetime τ and fraction f_ϕ. As shown in this figure, the small-scale density modes that enter the horizon before the ϕ particles decay (when the age of the Universe is less than τ) are suppressed relative to the standard case by a factor of $(1 - f_\phi)^2$.

Since the decaying particles are charged, the production of the LDP will always be accompanied by an electromagnetic cascade. The latter could in principle reprocess the light elements produced during BBN, or induce unreasonably large spectral distortions to the CMB. We show here that in fact these effects are small for the models discussed in this chapter.

The energy density released by the decay of ϕ particles can be parametrized as

$$\zeta_{EM} = \varepsilon_{EM} f_\phi Y_\chi\,, \tag{5.7}$$

where ε_{EM} is the average electromagnetic energy released in a ϕ decay and $Y_\chi \equiv n_\chi/n_\gamma$ is the dark-matter to photon ratio. In the specific models we discuss below, $\varepsilon_{EM} \approx \Delta m/3$, and

$$Y_\chi = \frac{\Omega_\chi \rho_c}{m_\chi n_\gamma} = 3 \times 10^{-12} \left(\frac{\mathrm{TeV}}{m_\chi}\right)\left(\frac{\Omega_\chi}{0.23}\right)\,. \tag{5.8}$$

[2]This assumes adiabatic initial conditions. Note that the perturbation $S_{\phi\beta} \equiv \delta_\phi - \delta_\beta$ will generally evolve away from zero in an arbitrary gauge, even when starting with the adiabatic initial condition $S_{\phi\beta}(0) = 0$, due to gradients in the proper time. It is a special simplifying property of the synchronous gauge that $S_{\phi\beta} = 0$ for all time for adiabatic initial conditions.

This yields

$$\zeta_{EM} \approx f_\phi \frac{\Delta m}{m_\chi} \left(\frac{\Omega_\chi}{0.23}\right) \text{ eV}\,. \tag{5.9}$$

In the models we discuss below, $\Delta m/m_\chi \sim 10^{-4}$, and $f_\phi \leq 1/2$, giving $\zeta_{EM} \lesssim 5\times10^{-5}$ eV. The limit derived from too much reprocessing of the BBN light element abundances is $\zeta_{EM} \lesssim 3.8\ \tau_{\rm yr}^{1/4}\times10^{-3}$ eV [5·26, 5·27] where $\tau_{\rm yr} = \tau/(1\text{ yr})$, so we are safely below this bound.

For $\tau_{\rm yr} \lesssim 300$, electromagnetic energy injection will result in a chemical-potential distortion to the CMB of [5·28]

$$\mu = 4.5\ \tau_{\rm yr}^{1/2} \times 10^{-3} \left(\frac{\zeta_{EM}}{\rm eV}\right) e^{-0.128\tau_{\rm yr}^{-5/4}}\,, \tag{5.10}$$

and so we expect $\mu \lesssim 2.0\times10^{-7}$ — 2.3×10^{-6} for lifetimes between $\tau_{\rm yr} = 1$ — 100, below the current limit of $\mu < 9\times10^{-5}$ [5·29].

5.3 The Nonlinear Power Spectrum

As density perturbations grow under the influence of gravity, linear evolution ceases to describe their growth and nonlinear effects must be taken into account. On large scales, where density perturbations have had insufficient time to become nonlinear, the linear matter power spectrum describes the statistics of density fluctuations. However, on small scales the full nonlinear matter power spectrum is required.

In order to calculate the nonlinear power spectrum for a given model we have used the recently devised `HALOFIT` method [5·30]. This method uses higher-order perturbation theory in conjunction with the halo model of large-scale structure to determine the nonlinear power spectrum given a linear power spectrum. It has been shown to accurately reproduce the nonlinear power spectra of standard N-body simulations and, unlike the earlier mappings such as the Peacock and Dodds formula [5·31], it is applicable in cases (like we consider here) when $\Delta^2(k) = k^3P(k)/2\pi^2$ is not a monotonic function. In particular we have checked that it approximately reproduces the shapes of the nonlinear power spectra determined in Ref. [5·32] through N-body simulations in models where the linear power spectrum is completely cut off on small scales. As we are discussing here less drastic alterations to the linear power spectrum, we believe the `HALOFIT` procedure provides an estimate of the nonlinear power spectrum adequate for illustrating the effect we describe in this chapter. Any detailed study would require a full N-body simulation.

In Figs. 5.2–5.5 we show both the linear and nonlinear matter power spectra at redshift $z = 4$ (data from measurements the Lyman-α forest probe redshifts 2–4 at wavenumbers $k/h \sim 0.1$–$10\,\text{Mpc}^{-1}$) for a charged-decay model, and for a model with a running spectral index. Although these models have different linear power spectra, nonlinear gravitational evolution causes these models to

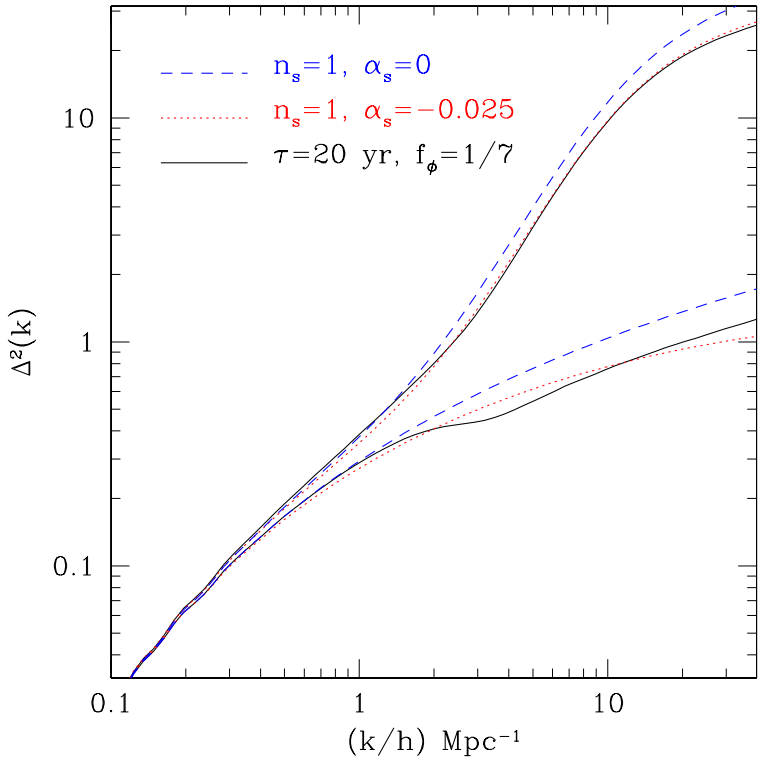

Figure 5.2: Shown at a redshift $z = 4$ are $\Delta^2(k) = k^3P(k)/2\pi^2$, the nonlinear (upper curves) and linear (lower curves) dimensionless matter power spectra per logarithmic interval for the canonical $n_s = 1$ ΛCDM model (dashed), for an $n_s = 1$ model with $\tau = 20$ yr and $f_\phi = 1/7$ (solid), and for a running-index model with $n_s = 1.00$ and $\alpha_s = -0.025$ (dotted). Although the linear power spectra differ significantly in these latter two models, nonlinear evolution causes them to have nearly degenerate nonlinear power spectra for $k/h \gtrsim 1.5$ Mpc^{-1}.

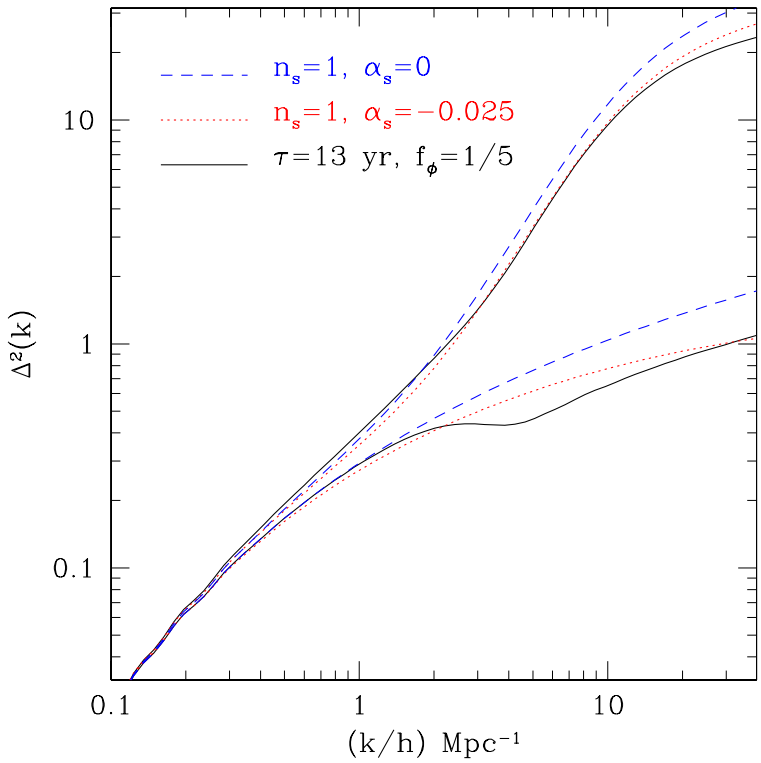

Figure 5.3: Shown at a redshift $z = 4$ are $\Delta^2(k) = k^3P(k)/2\pi^2$, the nonlinear (upper curves) and linear (lower curves) dimensionless matter power spectra per logarithmic interval for the canonical $n_s = 1$ ΛCDM model (dashed), for an $n_s = 1$ model with $\tau = 13$ yr and $f_\phi = 1/5$ (solid), and for a running-index model with $n_s = 1.00$ and $\alpha_s = -0.025$ (dotted). Nonlinear evolution causes these two models to have overlapping nonlinear power spectra for 2 Mpc$^{-1} \lesssim k/h \lesssim 10$ Mpc^{-1}.

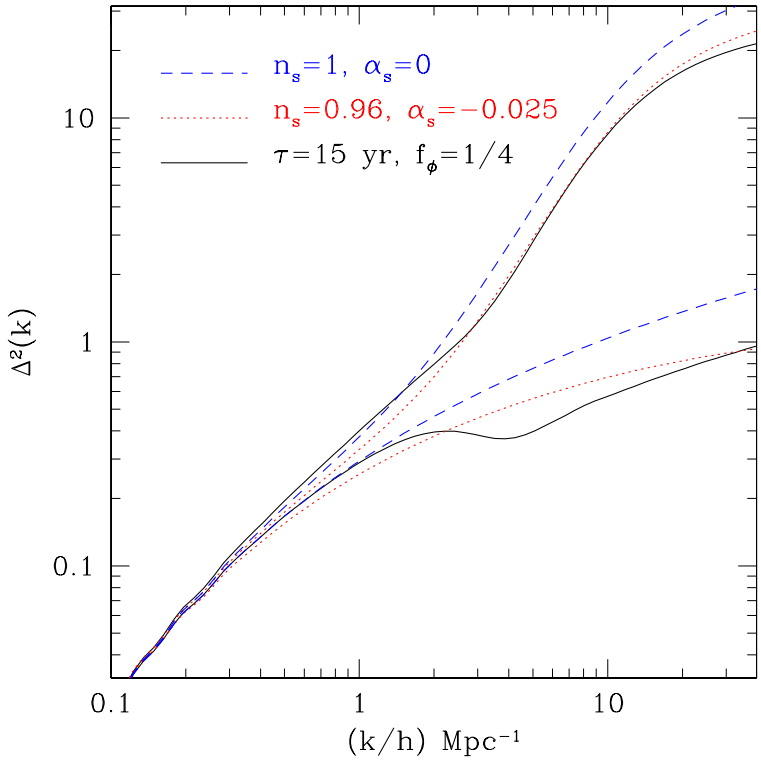

Figure 5.4: Shown at a redshift $z = 4$ are $\Delta^2(k) = k^3P(k)/2\pi^2$, the nonlinear (upper curves) and linear (lower curves) dimensionless matter power spectra per logarithmic interval for the canonical $n_s = 1$ ΛCDM model (dashed), for an $n_s = 1$ model with $\tau = 15$ yr and $f_\phi = 1/4$ (solid), and for a running-index model with $n_s = 0.96$ and $\alpha_s = -0.025$ (dotted). Nonlinear evolution causes the charged-decay model to match the canonical ΛCDM model for $k/h \lesssim 1.5$ Mpc^{-1} and the running-index model for $k/h \gtrsim 1.5$ Mpc^{-1}.

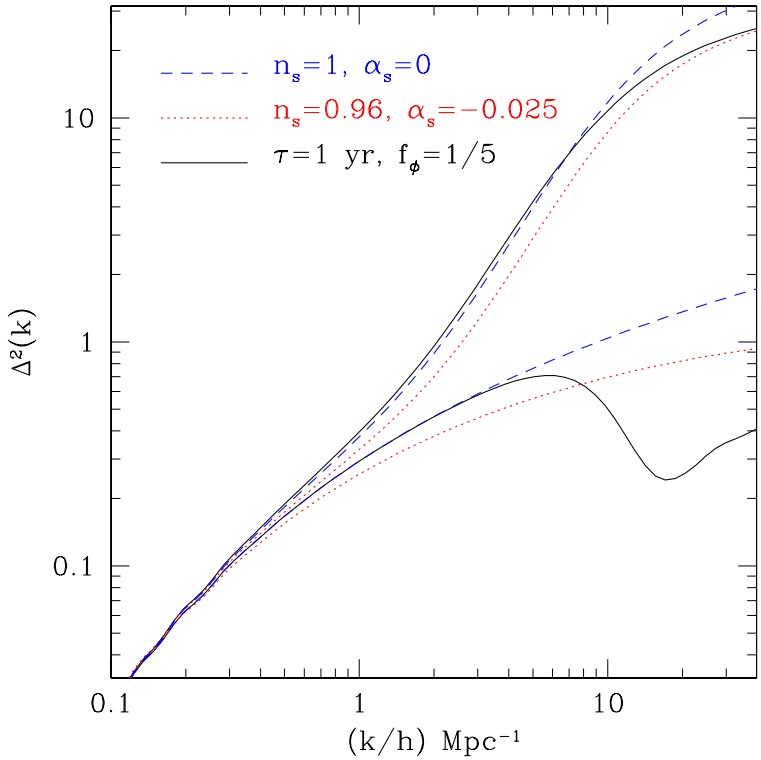

Figure 5.5: Shown at a redshift $z = 4$ are $\Delta^2(k) = k^3P(k)/2\pi^2$, the nonlinear (upper curves) and linear (lower curves) dimensionless matter power spectra per logarithmic interval for the canonical $n_s = 1$ ΛCDM model (dashed), for an $n_s = 1$ model with $\tau = 1$ yr and $f_\phi = 1/2$ (solid), and for a running-index model with $n_s = 0.96$ and $\alpha_s = -0.025$ (dotted). Despite the drastic change in the linear power spectrum nonlinear evolution causes the charged-decay model to match the canonical ΛCDM model for $k/h \lesssim 8$ Mpc^{-1} and the running-index model for $k/h \gtrsim 8$ Mpc^{-1}.

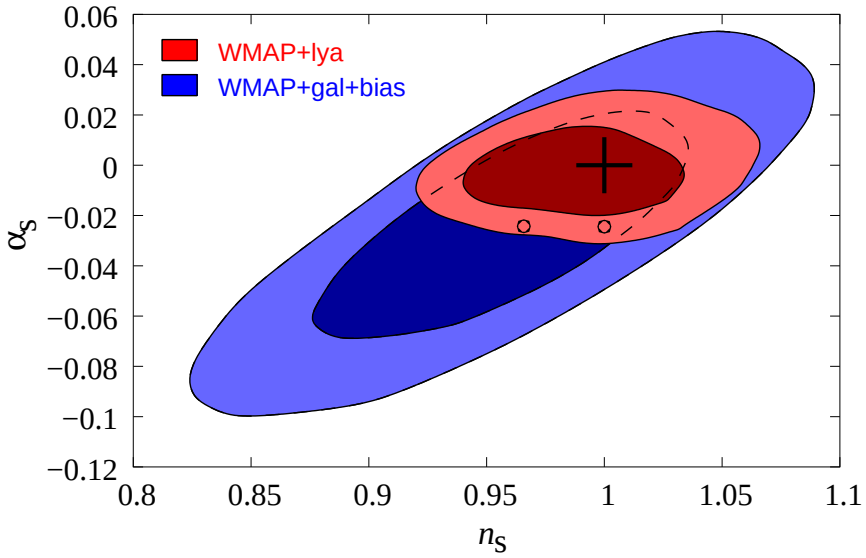

Figure 5.6: We reproduce here (with permission) Fig. 3 from Ref. [5·18], which illustrates the current constraints on the parameter α_s. The charged-decay models produce changes to the nonlinear power spectrum similar to constant $\alpha_s \neq 0$ models within the WMAP+Lya contour shown here and the examples we show in Figs. 5.2–5.5 are denoted by the black circles (ours). However, the charged-decay models leave the CMB angular power spectra unaltered (corresponding to models with $\alpha_s = 0$ on a line passing through the black cross). Matching to allowed constant-α_s models is conservative in this respect, and we thus expect the charged-decay models we have discussed in Figs. 5.2–5.5 to be consistent with current data.

have nearly identical nonlinear power spectra over interesting ranges of wavenumbers. The lifetimes shown were chosen because they produce effects on scales that can be probed by measurements of the Lyman-α forest, while the values $f_\phi = (1/2, 1/4, 1/5, 1/7)$ were chosen because they arise in the supersymmetric models we discuss below.

In Fig. 5.6 we show the current constraints on the (n_s,α_s) parameter space from the Wilkinson Microwave Anisotropy Probe (WMAP) satellite and SDSS Lyman-α forest data. The constant-α_s models we have compared our charged-decay model to lie within the WMAP+lya contours shown in the Figure. Since, in each case, the charged-decay model tends to interpolate between the standard Λ-cold-dark-matter (ΛCDM) model on larger scales (in particular on the scales probed by the CMB) and a constant-α_s running-index model on smaller scales, and both are allowed by these data, we expect the charged-decay models we have considered in Figs. 5.2–5.5 to be consistent with current data as well. We leave for future work the task of using a combined analysis of Lyman-α forest and other cosmological measurements to put limits directly on the (f_ϕ,τ) parameter space (and thus on the parameter space of the MSSM models we discuss below).

5.4 The 21-cm Power Spectrum

After recombination and the formation of neutral hydrogen, the gas in the Universe cools with respect to the CMB temperature $T_{\rm CMB}$ starting at at a redshift $z \sim 200$. The spin temperature T_s of the gas, which measures the relative populations of the hyperfine levels of the ground state of hydrogen separated by the 21-cm spin-flip transition, remains collisionally coupled to the temperature T_b of the baryons until a redshift $z \sim 30$ when collisions become inefficient and the spin temperature rises to $T_{\rm CMB}$. There is thus a window between $z \sim 30$–200 in which neutral hydrogen absorbs the CMB at a wavelength of 21 cm. It has recently been suggested that the angular fluctuations in the brightness temperature of the 21-cm transition within this window may be measured with future observations and used to constrain the matter power spectrum at these very high redshifts [5·21].

At redshifts $z \gtrsim 30$, the matter power spectrum on the scales $k/h \sim 1 - 100$ Mpc^{-1} of interest here are still in the linear regime. As discussed in Ref. [5·21], due to the unprecedented wealth of potential information contained in these 21-cm measurements, models with a running index or other small-scale modifications of the matter power spectrum can in principle be distinguished from each other. Even if the charged-particle lifetime is so small that no significant modifications to the power spectrum occur on scales probed by other cosmological observations (like the model shown in Fig. 5.5), 21-cm observations might detect or constrain such effects on the linear matter power spectrum.

5.5 The Long-Lived Charged Next-to-Lightest Dark-Matter Particle

From the particle-physics point of view, the setup we have introduced may seem *ad hoc*. We need a pair of particles that share a conserved quantum number and such that the lightest, the LDP, is neutral and stable, while the other, the NLDP, is coupled to the photon and quasi-stable, in order to significantly contribute to the cosmological energy density at an intermediate stage in the structure-formation process. Such a picture requires three ingredients: (1) the relic abundance of the LDP must be compatible with the CDM component; (2) the abundance of the NLDP must be at the correct level (namely, $\sim 1/5$ the total dark-matter density); and (3) the NLDP must have the proper lifetime (i.e., $\tau \sim 10$ yr).

Let us start with the last requirement. One way to get the required lifetime is to introduce a framework with strongly-suppressed couplings. One such possibility is, for instance, to assume that the LDP is a stable super-weakly–interacting dark-matter particle, such as a gravitino LSP in R-parity–conserving supersymmetric theories or the Kaluza-Klein first excitation G^1 of the graviton in the universal extra-dimension scenario [5·33]. The NLDP can have non-zero electric charge, but

at the same time a super-weak decay rate into the LDP (with the latter being the only allowed decay mode). In models of gauge-mediated supersymmetry breaking, this might indeed be the case with a stau NLSP decaying into a gravitino LSP. In this specific example, we have checked that, to retrieve the very long lifetimes we introduced in our discussion, we would need to impose a small mass splitting between the NLSP and LSP, as well as to raise the mass scale of the LSP up to about 100 TeV. This value is most often considered uncomfortably large for a supersymmetric (SUSY) setup, and it also makes thermal production of LDPs or NLDPs unlikely, being near a scale at which the unitary bound [5·34] gets violated. Without thermal production one would need to invoke first a mechanism to wipe out the thermal components and then provide a viable non-thermal production scheme that fixes the right portion of LDPs versus NLDPs. Finally, such heavy and extremely weakly-interacting objects would evade any dark-matter detection experiment, and would certainly not be produced at the forthcoming CERN Large Hadron Collider (LHC). This would then be a scheme that satisfies the three ingredients mentioned above, but that cannot be tested in any other way apart from cosmological observations.

An alternative (and to us, more appealing) scenario is one where long lifetimes are obtained by considering nearly degenerate LDP and NLDP masses. In this case, decay rates become small, without suppressed couplings, simply because the phase space allowed in the decay process gets sharply reduced. Sizeable couplings imply that, in the early Universe, the LDP and NLDP efficiently transform into each other through scattering from background particles. The small mass splitting, in turn, guarantees that the thermal-equilibrium number densities of the two species are comparable. To describe the process of decoupling and find the thermal relic abundance of these species, the number densities of the two have to be traced simultaneously, with the NLDP, being charged, playing the major role. This phenomenon is usually dubbed as *coannihilation* [5·35] and has been studied at length, being ubiquitous in many frameworks embedding thermal-relic candidates for dark matter, including common SUSY schemes.

We will show below that sufficiently long lifetimes may be indeed obtained in the minimal supersymmetric standard model (MSSM), when the role of the NLDP is played by a stau nearly degenerate in mass with the lightest neutralino (with the former being the stable LSP and the thermal-relic CDM candidate we will focus on in the remainder of our discussion). A setup of this kind appears naturally, e.g., in minimal supergravity (mSUGRA) [5·37]. This is the SUSY framework with the smallest possible parameter space, defined by only four continuous entries plus one sign, and hence also one of the most severely constrained by the requirement that the neutralino relic density matches the value from cosmological observations. Neutralino-stau coannihilations determine one of the allowed regions, on the border with the region where the stau, which in the mSUGRA scheme is most often the lightest scalar SUSY particle, becomes lighter than the neutralino. Although a stau-neutralino mass degeneracy is not "generic" in such models, this scenario is economical in that

the requirements of a long NLDP lifetime and of comparable LDP and NLDP relic abundances are both consequences of the mass degeneracy. In this sense, evidence for a running spectral index or any of the other observational features we discuss would simply help us sort out which configuration (if any) nature has chosen for SUSY dark matter.

5.6 Lifetimes of Charged NLSPs in the MSSM

We refer to a MSSM setup in which the lightest neutralino χ_1^0 is the lightest SUSY particle. The charged particles that could play the role of the NLDP include: (1) scalar quarks, (2) scalar charged leptons, and (3) charginos. We now discriminate among these cases by the number of particles in the final states for the decay of NLSPs to neutralino LSPs.

Scalar quarks and leptons have as their dominant decay mode a prompt two-body final state; i.e., $\tilde{S} \to \chi_1^0 S$, where we have labeled $\tilde{S}$ the SUSY scalar partner of the standard-model fermion S. A typical decay width for this process is $\mathcal{O}(1)$ GeV, corresponding to a lifetime $\mathcal{O}(10^{-24})$ s. This holds whenever this final state is kinematically allowed; i.e., if $m_{\tilde{S}} > m_S + m_{\chi_1^0}$. If it is *kinematically forbidden*, there are two possibilities: squarks may either decay through Cabibbo-Kobayashi-Maskawa (CKM) suppressed flavor-changing processes or through four-body decays. For instance, the stop decay may proceed through $\tilde{t} \to \chi_1^0 c$ or $\tilde{t} \to \chi_1^0 b f \bar{f}'$. On the other hand, within the same minimal–flavor-violation framework, scalar leptons are not allowed to decay in flavor-changing two-body final states, and only the four-body decay option remains.

The case for the chargino is different because this NLSP decay has a three-body final state, either with two quarks bound in a meson state (i.e., $\chi_1^+ \to \chi_1^0 \pi^+$) or with a leptonic three body channel (i.e., $\chi_1^+ \to \chi_1^0 l^+ \nu_l$). The latter final state becomes dominant, in particular, for electron-type leptons, $l = e$, if the mass splitting between NLSP and LSP becomes small.

We have listed all decay topologies as these are especially relevant when discussing the limit in which we force a reduction of the allowed decay phase-space volume; i.e., the limit in which the NLSP and LSP are quasi-degenerate in mass. Here we can also safely assume that the masses of the final-state particles, apart from the neutralino, are much smaller than the mass of the decaying particle. We can then consider the limit of a particle of mass $m_{\chi_1^0} + \Delta m$ decaying into a χ_1^0 and $n-1$ massless final states, and derive an analytical approximation to the behavior of the final-state phase space $\mathrm{d}\phi^{(n)}$ and of the decay width $\Gamma^{(n)}$ as functions of Δm. In the case of two-body decays, the phase space reads

$$\mathrm{d}\phi^{(2)} = \frac{\mathrm{d}\Omega}{32\pi^2}\left(1 - \left(\frac{m_{\chi_1^0}}{m_{\chi_1^0} + \Delta m}\right)^2\right) \propto \Delta m. \tag{5.11}$$

On the other hand, a recursive relation between $\mathrm{d}\phi^{(n)}$ and $\mathrm{d}\phi^{(n-1)}$ based on the invariant mass of

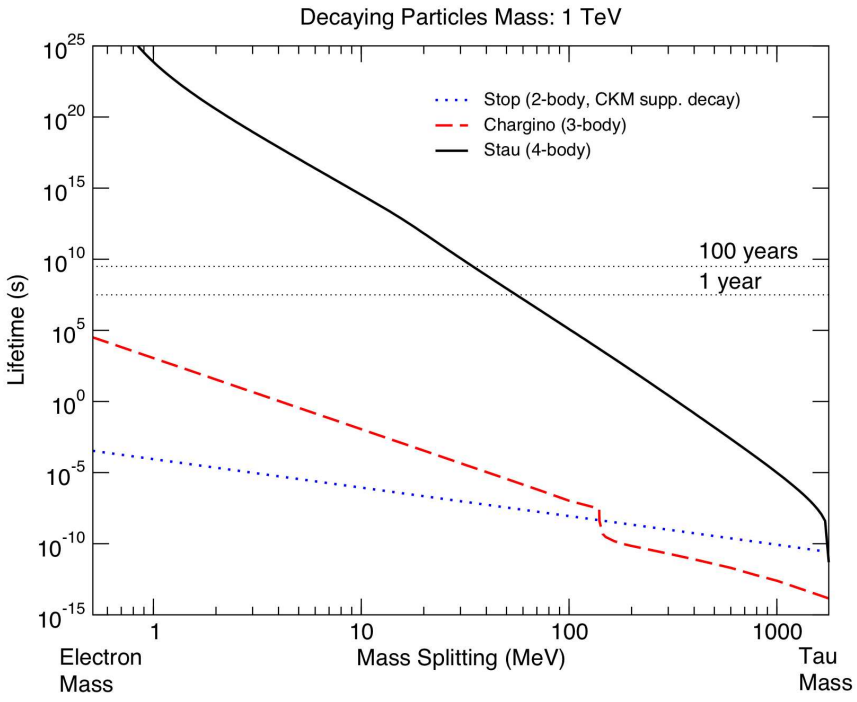

Figure 5.7: The lifetime of a 1 TeV stop (blue dotted line), a chargino (red dashed line) and a stau (black solid line), as a function of their mass splitting with the lightest SUSY particle, the lightest neutralino.

couples of final states yields

$$\begin{aligned} \mathrm{d}\phi^{(n)}(\Delta m) &\propto \mathrm{d}\phi^{(n-1)}(\Delta m) \times \int^{\Delta m} \mathrm{d}\mu(\mathrm{d}\phi^{(2)})(\mu) \\ &\propto (\Delta m)^{2(n-2)+1}. \end{aligned} \tag{5.12}$$

The dependence of the decay width $\Gamma^{(n)}$ on Δm must, however, take into account not only the phase-space dependence, but also the behavior of the amplitude squared $\mathcal{M}^{(n)}$ of the processes as a function of Δm. The occurrence of a massless final state yields, in the amplitude squared, a factor that scales linearly with the momenta circulating in the Feynman diagram. One therefore has the further factor

$$\mathcal{M}^{(n)} \propto (\Delta m)^{n-1}. \tag{5.13}$$

Finally, we have

$$\Gamma^{(n)} \propto \mathcal{M}^{(n)} \times \mathrm{d}\phi^{(n)} \propto (\Delta m)^{3n-4}, \tag{5.14}$$

i.e., the lifetime to decay to a two-body final state scales like $\tau^{(2)} \propto (\Delta m)^{-2}$, while for a four-body decay we have $\tau^{(4)} \propto (\Delta m)^{-8}$.

Reducing the NLSP mass splitting Δm one may hope to obtain "cosmologically relevant" NLSP lifetimes. For scalar quarks, this is not quite the case, as two-body final states always dominate, and

even for amplitudes that are CKM-suppressed, the scaling with Δm is too shallow, and the resulting lifetimes are rather short, even for very small mass splittings.

The situation is slightly more favorable for charginos, whose three-body decay width is approximately equal to

$$\frac{G_F^2}{(2\pi)^3}\frac{16}{15}(\Delta m)^5. \tag{5.15}$$

At tree level and in the limit of pure Higgsino-like or Wino-like states, the lightest neutralino and lightest chargino are perfectly degenerate in mass. However, when one takes into account loop corrections to the masses, Δm usually turns out to be larger than a few tens of MeV; see, e.g., [5·38, 5·39]. This translates into an absolute upper limit on the chargino lifetime of about $10^{-(2-3)}$ s.

Finally, in the case of sleptons $\tilde{l}$, the interesting regime is when $\Delta m < m_l$ and only four-body decays are allowed. For example, for the lightest stau $\widetilde{\tau}_1$, the processes

$$\widetilde{\tau}_1 \;\rightarrow\; \chi_1^0\,\nu_\tau\,f\,\bar{f}', \tag{5.16}$$

with

$$(\bar{f}', f) = (\bar{\nu}_\mu, \mu),\ (\bar{\nu}_e, e),\ (\bar{u}, d),\ (\bar{c}, s), \tag{5.17}$$

where, depending on Δm, only final states above the kinematic threshold are included. In this case, Δm can be safely taken as a free parameter of the theory: in most scenarios the gaugino mass parameter setting the neutralino mass for bino-like neutralinos, and the scalar soft mass parameter setting the stau mass, are usually assumed to be independent. Analogously, there are models in which the μ parameter setting the mass for a Higgsino-like neutralino, and scalar soft mass parameter are unrelated. A similar picture applies to the smuon, though lifetimes start to be enhanced at much smaller mass splittings (m_μ instead of m_τ), and it is theoretically difficult to figure out a scenario in which the lightest smuon is lighter than the lightest stau.

The scalings we have sketched are summarized in Fig. 5.7, where we plot lifetimes for a stop (CKM-suppressed), a chargino, and a stau as a function of Δm. The decaying-particle masses have been set to 1 TeV, and the Δm range is between the electron and the tau mass. The lifetimes of the stop and of the chargino have been computed with the code SDECAY [5·40]. Some details on the computation of the stau lifetime are given below. Notice that the scaling of Eq. (5.14) is accurately reproduced for all cases. The bottom line is that indeed *the lightest stau can play the role of the NLDP with a cosmologically relevant lifetime*, and that the stau is the only particle in the MSSM for which this can be guaranteed by adjusting only the LDP-NLDP mass splitting.

5.7 The Relative Abundance of the Charged NLSP

The next step is to determine the relic density of the NLSP and LSP in the early stages of the evolution of the Universe. As we have already mentioned, we are going to consider thermal production. We briefly review here how to compute the evolution of number densities with coannihilation [5·35,5·36].

Consider a setup with N supersymmetric particles χ_1, χ_2, ... χ_N, each with mass m_i and number of internal degrees of freedom g_i. The ordering is such that $m_1 \leq m_2 \leq \cdots \leq m_N$. In the evolution equations, the processes that change the number density of SUSY particles i are of three kinds:

$$\begin{aligned} &\text{(a)}\ \ \chi_i\,\chi_j \leftrightarrow X_a^f\,, && \forall\, j\,,\\ &\text{(b)}\ \ \chi_i\,X_b^i \leftrightarrow \chi_j\,X_b^f\,, && \forall\, j\neq i\,,\\ &\text{(c)}\ \ \chi_j \leftrightarrow \chi_i\,X_c^f\,, && \forall\, j> i\,, \end{aligned} \tag{5.18}$$

where X_a, X_b^i, X_b^f, and X_c^f are (sets of) standard-model (SM) particles. In practice, the relevant processes one should include are those for SM particles that are in thermal equilibrium. Assuming the distribution function for each particle k is the same as for the equilibrium distribution function,

$$f_k(E_k) \propto f_k^{\rm eq}(E_k) = \frac{1}{\exp(E_k/T)\pm 1}\,, \tag{5.19}$$

and invoking the principle of detailed balance, the Boltzmann equation for the evolution of the number density of SUSY particle i, $n_i = g_i/(2\pi)^3\int d^3p f_i(E)$, normalized to the entropy density of the Universe, $Y_i = n_i/s$, as a function of the variable $x\equiv m_1/T$ (with T the Universe temperature; this is equivalent to describing the evolution in time) is given by

$$\begin{aligned} \frac{x}{\hat g(x)\,Y_i^{\rm eq}}\frac{dY_i}{dx} = &-\sum_j \frac{\langle\sigma_{ij}v_{ij}\rangle\,n_j^{\rm eq}}{H}\left(\frac{Y_i\,Y_j}{Y_i^{\rm eq}\,Y_j^{\rm eq}}-1\right)\\ &-\sum_{j\neq i}\frac{[\sum_X\langle\sigma_{iX\to j}v_{iX\to j}\rangle\,n_X^{\rm eq}]}{H}\left(\frac{Y_i}{Y_i^{\rm eq}}-\frac{Y_j}{Y_j^{\rm eq}}\right)\\ &+\sum_{j>i}\frac{\Gamma_{j\to i}}{H}\left(\frac{Y_j}{Y_i^{\rm eq}}-\frac{Y_i\,Y_j^{\rm eq}}{(Y_i^{\rm eq})^2}\right)-\sum_{j<i}\frac{\Gamma_{i\to j}}{H}\left(\frac{Y_i}{Y_i^{\rm eq}}-\frac{Y_j}{Y_j^{\rm eq}}\right)\,. \end{aligned} \tag{5.20}$$

In this equation, analogously to Y_k, we have defined $Y_k^{\rm eq}\equiv n_k^{\rm eq}/s$, the ratio of the equilibrium number density of species k (at temperature x) to the entropy density. On the left-hand side, we introduced

$$\hat g(x) \equiv \left[1+\frac{T}{3\,g_{\rm eff}}\,dg_{\rm eff}/dT\right]^{-1}\,, \tag{5.21}$$

with $g_{\rm eff}(T)$ being the effective degrees of freedom in the entropy density. The function $\hat g(x)$ is close to 1 except for temperatures at which a background particle becomes nonrelativistic. On the right-hand side, the last two terms contain factors in $\Gamma_{k\to l}$ that label the partial decay width of a

particle k in any final state containing the particle l; in our discussion they play a role just at late times when NLSPs decay into LSPs, giving the scaling we have used in Eq. (5.1). The first two terms refer, respectively, to processes of the kind (a) and (b) in Eq. (5.18), including all possible SM final and initial states. The symbol $\langle\sigma_{ab}v_{ab}\rangle$ indicates a thermal average of the cross section $\sigma_{ab}v_{ab}$; i.e.,

$$\langle\sigma_{ab}v_{ab}\rangle = \frac{1}{n_a^{\rm eq}\, n_b^{\rm eq}} \int d^3p_a d^3p_b f_a^{\rm eq}(E_a) f_b^{\rm eq}(E_b) \sigma_{ab}v_{ab}\,. \tag{5.22}$$

As is evident from the form we wrote the Boltzmann equation, interaction rates have to be compared with the expansion rate H of the Universe. In general, over a large range of intermediate temperatures, the $\langle\sigma_{iX\to j}v_{iX\to j}\rangle\, n_X^{\rm eq}$ terms will be larger than the $\langle\sigma_{ij}v_{ij}\rangle\, n_j^{\rm eq}$ terms, since we expect the cross sections to be of the same order in the two cases, but the scattering rates will be more efficient as long as they involve light background particles X with relativistic equilibrium densities $n_X^{\rm eq}$ that are much larger than the nonrelativistic Maxwell-Boltzmann–suppressed equilibrium densities $n_j^{\rm eq}$ for the more massive particles j. This implies that collision processes go out of equilibrium at a smaller temperature, or later time, than pair-annihilation processes. Writing explicitly the expression for $d/dx(Y_i/Y_i^{\rm eq} - Y_k/Y_k^{\rm eq})$, in which Maxwell-Boltzmann–suppressed terms and terms in mass splitting over mass scale can be neglected, one finds explicitly that in the limit that collisional rates are much larger than the expansion rate, for any i and k, $Y_i(x)/Y_i^{\rm eq}(x) = Y_k(x)/Y_k^{\rm eq}(x)$, or equivalently,

$$Y_i(x) = \frac{Y(x)}{Y^{\rm eq}(x)} Y_i^{\rm eq}(x)\,, \tag{5.23}$$

with $Y(x) \equiv \sum_k Y_k(x)$ and $Y^{\rm eq}(x) \equiv \sum_k Y_k^{\rm eq}(x)$. At T_{cfo}, when $\sum_X \langle\sigma_{iX\to k}v_{iX\to k}\rangle\, n_X^{\rm eq} \simeq H$, collision processes decouple and the relative number densities become frozen to about

$$\frac{n_i(T)}{n_k(T)} = \frac{n_i^{\rm eq}(T_{cfo})}{n_k^{\rm eq}(T_{cfo})} \simeq \frac{g_i}{g_k}\left(\frac{m_i}{m_k}\right)^{3/2} \exp\left(\frac{m_k - m_i}{T_{cfo}}\right), \tag{5.24}$$

up to the time (temperature) at which heavier particles decay into lighter ones.

The sum $Y(x)$ of the number densities has instead decoupled long before. Eq. (5.24) is the relation that is implemented to find the usual Boltzmann equation [5·35, 5·36] for the sum over number densities of all species compared to the sum of equilibrium number densities, and that shows that the decoupling for Y occurs when the total effective annihilation rate becomes smaller than the expansion rate, at a temperature T_{afo} that, as mentioned above, is much larger than T_{cfo}.

To get an estimate for T_{cfo}, we can take, whenever a channel is kinematically allowed, the (very) rough s-wave limit,

$$\sigma_{i\to k}v_{i\to k} \sim \sigma_{ik}v_{ik} \sim \langle\sigma_{ik}v_{ik}\rangle \sim \frac{3\cdot 10^{-27}\,{\rm cm}^3\,{\rm s}^{-1}}{\Omega_\chi h^2}\,, \tag{5.25}$$

where an approximate relation between annihilation rate and relic density has been used [5·10]. There are then two possibilities depending on whether *(i)* the background particle X enforcing colli-

sional equilibrium has a mass much larger than the mass splitting Δm between the SUSY particles involved, or *(ii)* the opposite regime holds. In the first case we find that $T_{cfo} \sim m_X/(10-15)$, roughly the temperature at which X itself (except for neutrinos) decouples from equilibrium. From Eq. (5.24) we find that $n_i/n_k \simeq g_i/g_k$; i.e., they have comparable abundances. In the opposite case, we find instead that $T_{cfo} \sim \Delta m/(10-15)$ and hence $n_i/n_k \simeq g_i/g_k \left(\frac{m_i}{m_k}\right)^{3/2} \exp\left[(10-15)\mathrm{sign}(m_k - m_i)\right]$; i.e., the abundance of the heavier particle is totally negligible compared to the lighter one.

Long-lived stau NLSPs are kept in collisional equilibrium with neutralinos by scattering on background $\tau^\pm$ and emission of a photon. In this case we are clearly in the limit *(i)*, as $m_{\tilde{\tau_1}} - m_{\chi_1^0} \ll m_\tau$. Since the number of internal degrees of freedom for both staus and neutralinos is 2, we find that the fraction of charged dark matter in this model is

$$f_\phi = \frac{g_{\tilde{\tau_1}}}{g_{\chi_1^0} + g_{\tilde{\tau_1}}} = \frac{1}{2}\,. \tag{5.26}$$

More carefully, though, this is not a strict prediction of our MSSM setup for a stau NLDP, and should be interpreted just as an upper limit. In fact, if we have other SUSY particles that are quasi-degenerate in mass with the LSP, and if they then coannihilate and decouple from the neutralino at a later time than the stau, either through mode *(i)* and then immediately decaying into neutralinos (in all explicit examples, the decay into staus is strongly suppressed compared to the decay into neutralinos), or through mode *(ii)*, then f_ϕ is reduced to

$$f_\phi = \frac{g_{\tilde{\tau_1}}}{\sum_{i=1}^{n} g_i}\,, \tag{5.27}$$

where the sum in the denominator involves the neutralino, the stau and all SUSY particles with T_{cfo} lower than the T_{cfo} for staus.

In particular, if the lightest neutralino is a nearly-pure higgsino, then the next-to-lightest neutralino will also be a higgsino very nearly degenerate in mass, and the lightest chargino will also be nearly degenerate in mass, with mass splittings possibly smaller than m_τ. In this case, charginos and neutralinos will be kept in collisional equilibrium through scatterings on (ν_l, l) pairs. At the same time, the collisional decoupling of staus might be slightly delayed because of chargino-stau conversions through the emission of a photon and absorption of a tau neutrino; however this second process has a Yukawa suppression (as we are considering Higgsino-like charginos) compared to the first, and hence it is still guaranteed that the stau decoupling temperature is larger than the chargino T_{cfo} temperature. Since the mass splitting between neutralino and chargino and that between lightest neutralino and next-to-lightest neutralino cannot be smaller than few tens of MeV (due to loop corrections to the masses), decoupling will always happen in mode *(ii)* and we do not have to worry about possible stau production in their decays. Since $g_{\chi^+} = 4$, applying the formula in Eq. (5.27) we

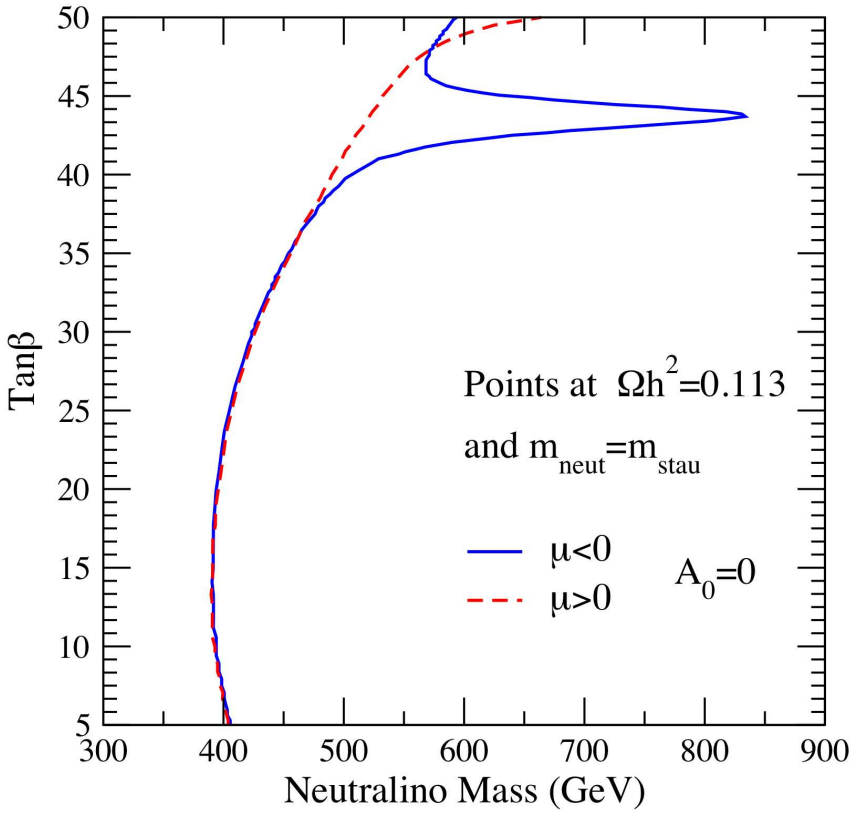

Figure 5.8: Points in the mSUGRA parameter space, at null trilinear scalar coupling A_0, featuring $\Omega_{\chi_1^0} = 0.11$ and $m_{\chi_1^0} = m_{\tilde{\tau}_1}$, for both negative (solid line) and positive (dashed line) sign of the μ parameter.

find $f_\phi = 1/5$. In the case of a wino-like lightest neutralino, instead, the only extra coannihilating partner would be the lightest chargino, yielding $f_\phi = 1/4$. Finally, adding to this picture, e.g., a quasi-degenerate smuon and selectron, $f_\phi = 1/7$ could be obtained.

5.8 Long-Lived Stau NLSPs in Sample Minimal Models

We provide a few examples of well motivated theoretical scenarios where neutralino-stau degeneracy occurs, possibly in connection with further coannihilating partners driving low values of f_ϕ. In surveying the possible models, the criterion we take here is that of the *composition* of the lightest neutralino in terms of its dominating gauge eigenstate components. We thus outline *bino*-, *higgsino*- and *wino*-like lightest-neutralino benchmark scenarios.

5.8.1 A Case with $f_\phi = 1/2$: Binos in the mSUGRA Model

As we mentioned above, in the framework of minimal supergravity (mSUGRA) [5·22] one of the few cosmologically allowed regions of parameter space is the tail where the neutralino and the stau

are quasi-degenerate. In this case, coannihilations reduce the exceedingly large bino-like neutralino relic abundance to cosmologically acceptable values for neutralino masses up to around 600 GeV. Coannihilation effects depend on the relative mass splitting between the two coannihilating species. Requiring a mass splitting as small as those found above amounts, as far as the neutralino relic density is concerned, to effectively setting $m_{\chi_1^0} = m_{\tilde{\tau}_1}$. This, in turns, sets the mass of the neutralino-stau system once a particular value of the relic abundance is required. We plot in Fig. 5.8 points fulfilling at once $m_{\chi_1^0} = m_{\tilde{\tau}_1}$ and $\Omega_{\chi_1^0} = \Omega_{\rm CDM} \simeq 0.113$, the latter being the central value as determined from the analysis of CMB data [5·1]. Result are shown in the ($m_{\chi_1^0}$-$\tan\beta$) plane, with $\tan\beta$ the ratio of the vacuum expectation values of the two neutral components of the $SU(2)$ Higgs doublets, and at a fixed value of the trilinear coupling $A_0 = 0$ (this latter quantity is, however, not crucial here). The solid line corresponds to negative values of the Higgsino mass parameter μ, while the dashed line corresponds to positive values of μ. Notice that at $\mu < 0$ the accidental overlap of the heavy Higgs resonance with the coannihilation strip, around $\tan\beta = 43$, shifts the neutralino masses to larger values. We point out that the two requirements of mass degeneracy and of the correct relic density determine, at a given value of $\tan\beta$, the required mass of the neutralino-stau system, thus solving the residual mSUGRA parameter space degeneracy, and making the present framework testable and predictive.

In the minimal configuration we have considered, only the lightest neutralino and the lightest stau are playing a role, hence $f_\phi = 1/2$. However, since the mass splitting between the lightest stau and lightest smuon and selectron is rather small, assuming a slight departure from universality in the scalar sector, two additional quasi-degenerate scalar particles can be obtained and the fraction of charged dark matter reduced to $f_\phi = 1/4$.

5.8.2 A Case with $f_\phi = 1/5$: Higgsino-like Neutralinos

When the μ term is lighter than the gaugino masses M_1 and M_2, the lightest neutralino gets dominated by the higgsino component. This situation occurs, again within the mSUGRA model, in the so-called hyperbolic branch/focus point (HB/FP) region [5·41, 5·42], where large values of the common soft breaking scalar mass m_0 drive μ to low values. In this region, scalars are naturally heavy, at least in the minimal setup; however, the occurrence of non-universalities in the scalar sector [5·43] may significantly affect the sfermion mass pattern. In particular, in a SUSY Grand Unified Theory (GUT) scenario, soft breaking sfermion masses get contributions from D-terms whenever the GUT gauge group is spontaneously broken with a reduction of rank [5·44]. Light staus may naturally occur, for instance when the weak hypercharge D-term dominates and features negative values. In this case the hierarchy between diagonal entries in the soft supersymmetry-breaking scalar mass matrices is $m_E^2 \ll m_{U,D,Q,L}^2$. The m_L^2 term may also be lowered in presence of additional D-terms originating from the breaking of further $U(1)$ symmetries.

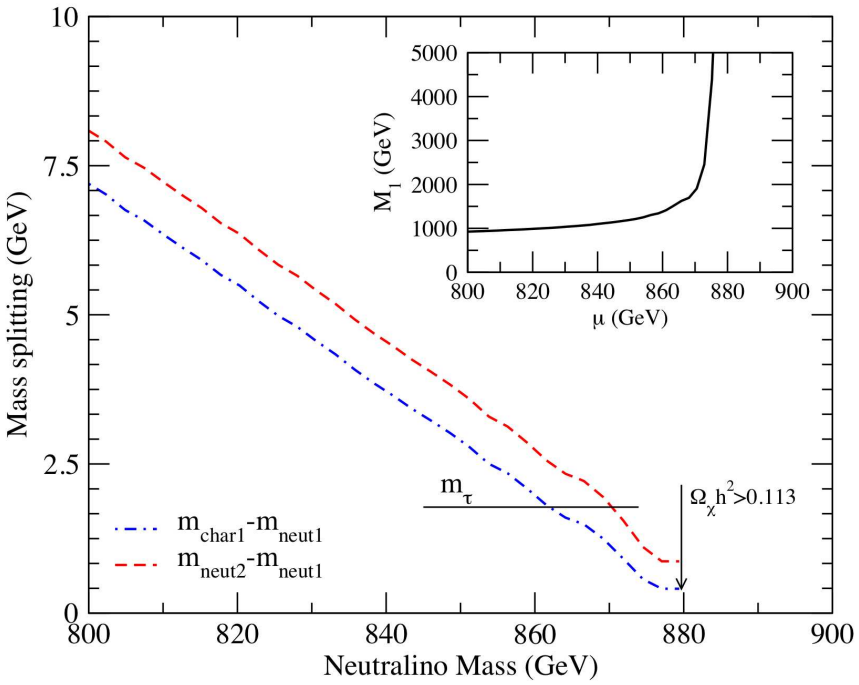

Figure 5.9: A low-energy parameterization of the higgsino-like neutralino case, at $\tan\beta = 50$. In the smaller frame we indicate the points, on the (μ, M_1) plane, that give $\Omega h^2 \simeq 0.113$. The corresponding models are reproduced in the larger frame, where we plot the relevant mass splittings of the chargino-neutralino system.

The relic neutralino abundance is fixed by the interplay of multiple chargino-stau-neutralino coannihilations. In the limit of pure higgsino, the dynamics of these processes fixes the value of μ yielding a given relic neutralino abundance. On the other hand, a mixing with the bino component along the borders of the HB/FP region may entail a larger spread in the allowed mass range, affecting the χ_1^0 higgsino fraction. We sketch the situation in Fig. 5.9, where we resort, for computational ease, to a low-energy parameterization of the above outlined scenario. The smaller frame shows the points on the (μ, M_1) plane that produce the required amount of relic neutralinos. The larger frame reproduces the values of the chargino-neutralino and neutralino–next-to-lightest-neutralino mass splitting; the lines end in the pure-higgsino regime. Suitable models, in the present framework, must also fulfill the mass splitting requirement $m_{\widetilde{\chi}_2, \widetilde{\chi}_1^+} - m_{\chi_1^0} < m_\tau$. This enforces the allowed neutralino mass range, at $\tan\beta = 50$, between 870 and 880 GeV. Had we lowered the value of $\tan\beta$, the corresponding $m_{\chi_1^0}$ range would only have shifted to masses just a few tens of GeV lighter.

As we have already mentioned, since we are dealing with a case with two neutralinos, a chargino and a stau quasi-degenerate in mass, we find $f_\phi = 1/5$. Again, a smuon and a selectron can be added to this to shift the charged particle fraction to $f_\phi = 1/7$.

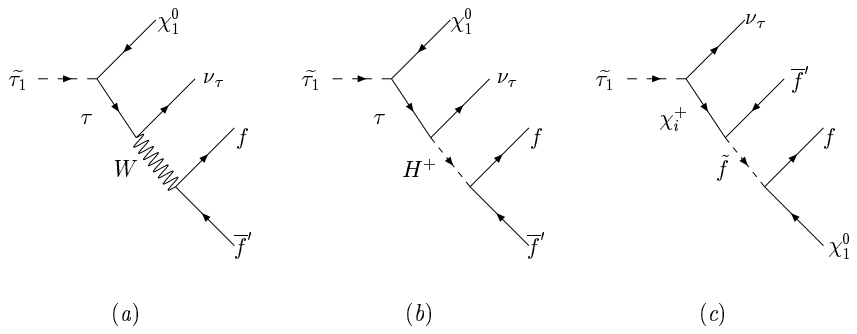

Figure 5.10: A few Feynman diagrams for four-body final states for the process $\widetilde{\tau}_1 \to \chi_1^0 \nu_\tau f \bar{f}'$,. Diagram (a) is the dominant diagram; diagrams of type (b) are sub-dominant, and diagrams of type (c) are sub-sub-dominant.

5.8.3 A Case with $f_\phi = 1/4$: Wino-like Neutralinos

A benchmark case where the lightest neutralino is wino-like is instead provided by the *minimal Anomaly mediated SUSY breaking* (mAMSB) scenario [5·23]. In this framework, tachyonic sfermion masses are cured by postulating a common scalar mass term, m_0. The lightest sfermion turns out to correspond to the lightest stau $\widetilde{\tau}_1$ again. The latter, at suitably low m_0 values, may be degenerate in mass with the lightest neutralino. We performed a scan of the mAMSB parameter space, requiring $m_{\chi_1^0} \simeq m_{\widetilde{\tau}_1}$, and found that the correct relic abundance requires the neutralino masses to lie in the range $1250 \lesssim m_{\chi_1^0} \lesssim 1600$ GeV, with the lower bound holding for small values of $\tan\beta$, while the higher for larger ones. The chargino-neutralino mass splitting is always below the τ mass, since χ_1^0 in mAMSB is always a very pure wino, and it has large masses. We finally remark that here, as in the case of higgsinos, the occurrence of stau coannihilations *raises* the neutralino relic abundance, contrary to the standard result with a bino-like LSP.

Since, in this case, we have one neutralino, a chargino and a stau quasi-degenerate in mass, we find $f_\phi = 1/4$.

5.8.4 The Stau Lifetime

The $\widetilde{\tau}_1$ four-body decay proceeds through diagrams of the types sketched in Fig. 5.10. They come in three sets: those with W exchange, those with $H^\pm$ exchange, and those with a sfermion exchange. However, since $\Delta m \equiv m_{\widetilde{\tau}_1} - m_{\chi_1^0}$ is much smaller than any supersymmetric-particle mass, the virtuality of all diagrams except those featuring a τ exchange (diagrams (a) and (b) in the Figure) is extremely large. Hence all diagrams but those with a τ exchange will be suppressed by a factor $(m_\tau/m_{\rm SUSY})^4 \sim 10^{-8}$, and the interferences with the dominating diagrams by a factor $(m_\tau/m_{\rm SUSY})^2 \sim 10^{-4}$. Of the two diagrams with a τ exchange, however, the one with the $H^\pm$ exchange has a Yukawa suppressed $H^\pm f \bar{f}'$ vertex, which gives a suppression, with respect to the

W-exchange diagram,

$$\sim (m_\mu \tan\beta/m_W)^2 (m_\tau \tan\beta/m_W)^2 \sim 10^{-7} - 10^{-3}$$
$$\text{for} \quad \tan\beta = 5 - 50 \tag{5.28}$$

in the most favorable muonic final channel. Notice that the chirality structure of the couplings entails that no interference between these two diagrams is present. Diagrams with a charged Higgs or a sfermion exchange are moreover further suppressed with respect to those with a W exchange by a factor $(m_W/m_{\rm SUSY})^4$ which, depending on the SUSY spectrum, can also be relevant.

In this respect, a very good approximation to the resulting stau lifetime is obtained by considering only the first diagram, whose squared amplitude reads

$$|\mathcal{M}|^2 = \sum_{\text{final states}} \frac{16(g_2/\sqrt{2})^4 (p_{\nu_\tau} \cdot p_f)}{\left((p_\tau)^2 - m_\tau^2\right)^2 \left((p_W)^2 - m_W^2\right)^2} \Big[|V_R|^2 \left(2(p_{\chi_1^0} \cdot p_\tau)(p_\tau \cdot p_{\bar{f}'}) - (p_\tau)^2 (p_{\chi_1^0} \cdot p_{\bar{f}'}) \right) +$$
$$+ m_\tau^2 |V_L|^2 (p_{\chi_1^0} \cdot p_{\bar{f}'}) - 2 m_{\chi_1^0} m_\tau \mathrm{Re}(V_L^* V_R)(p_\tau \cdot p_{\bar{f}'}) \Big], \tag{5.29}$$

with $V_{L,R}$ the left- and right-handed coupling of the $\widetilde{\tau}_1$ in the $\widetilde{\tau}_1 \tau \chi_1^0$ vertex. The sum is extended over the final states of Eq. (5.17). For numerical purposes, the four-particles phase space is integrated with the use of the Monte Carlo routine Rambo, with final-state finite-mass corrections for all final states.

Fig. 5.11 shows the stau lifetime for a sample of the supersymmetric scenarios outlined in the preceding sections. We fully account for threshold effects in the phase space, with the four contributions from electronic, muonic, and first- and second-generation quarks. The quark masses have been set to their central experimental values [5·45]. For the case of higgsinos, we reproduce the two extreme regimes when the m_L^2 term is large (no Left-Right mixing) and when $m_L^2 \simeq m_R^2$ (maximal Left-Right mixing). The differences in the lifetimes are traced back to overall mass effects and to the values of the V_L and V_R couplings (for instance, $|V_L| \gg |V_R|$ in the wino case, while the opposite regime holds for the case of higgsinos and no LR-mixing). In any case, we conclude that lifetimes of the order of $1 - 100$ years are obtained with a mass splitting $\Delta m = 20 - 70$ MeV.

5.9 Dark Matter Searches and Collider Signatures

Unlike other charged long-lived NLDP scenarios, the framework we outlined above has the merit of being, in principle, detectable at dark-matter–detection and collider experiments.

We show in Fig. 5.12 the spin-independent neutralino-proton scattering cross section for the mSUGRA parameter-space points of Fig. 5.8, and for the higgsino and wino (mAMSB) cases, as discussed in Sec. 5.8.2 and 5.8.3, together with the current exclusion limits from the Cryogenic Dark

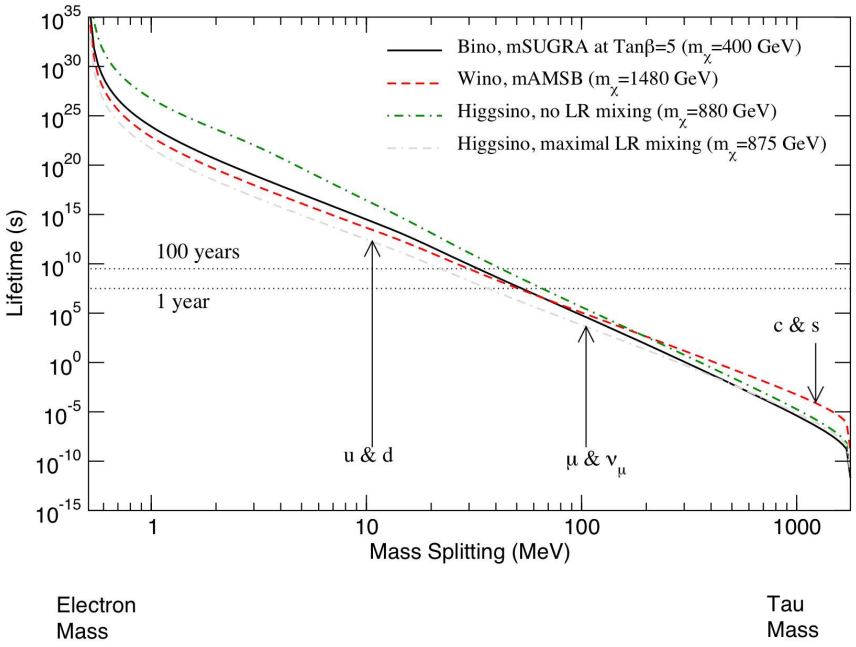

Figure 5.11: The stau lifetime, as a function of the mass splitting with the lightest neutralino. The parameter space points are defined by the two requirements $m_{\chi_1^0} \simeq m_{\tilde{\tau}_1}$ and $\Omega_{\chi_1^0} = 0.11$. For the wino and higgsino cases $\tan\beta = 50$, while in all cases $A_0 = 0$ and $\mu > 0$.

Matter Search (CDMS) experiment [5·46], and the future reach of the XENON 1-ton facility [5·47]. For binos, at $\mu > 0$ most points lie within less than one order of magnitude with respect to the future projected sensitivity, making it conceivable that this scenario may be tested in next-generation facilities. For negative μ, destructive interference among the lightest- and heavy-Higgs contributions lead instead to cancellations in $\sigma_{\chi_1^0 P}$. Finally, higgsino and wino detection rates respectively lie one and two orders of magnitude below the future expected sensitivity.

Indirect-detection experiments look less promising, even for winos and higgsinos, which always feature uncomfortably large neutralino masses. We checked, for instance, that the expected muon flux from the Sun, generated by neutralino pair annihilations, is at most $10^{-(4-5)}$ muons per km^2 per year, far below the sensitivity of future neutrino telescopes like IceCube [5·48].

Turning to collider experiments, considering as the searching tool the usual missing transverse-energy channels, dedicated studies have shown that the mSUGRA coannihilation strip will be within LHC reach, mainly through in a mass range that extends up to $m_{\chi_1^0} \lesssim 550$ GeV along the coannihilation strip, quite independently of $\tan\beta$ [5·49]. Concerning higgsinos and winos, instead, the relevant mass range we study here appears to be beyond standard LHC searches [5·58].

Even at high-energy colliders, the peculiar and distinctive feature of this scenario is however represented by the long-lived stau. The production of what have sometimes been dubbed long-lived *charged massive particles* (CHAMP's) [5·50] has been repeatedly addressed. Exclusion limits were

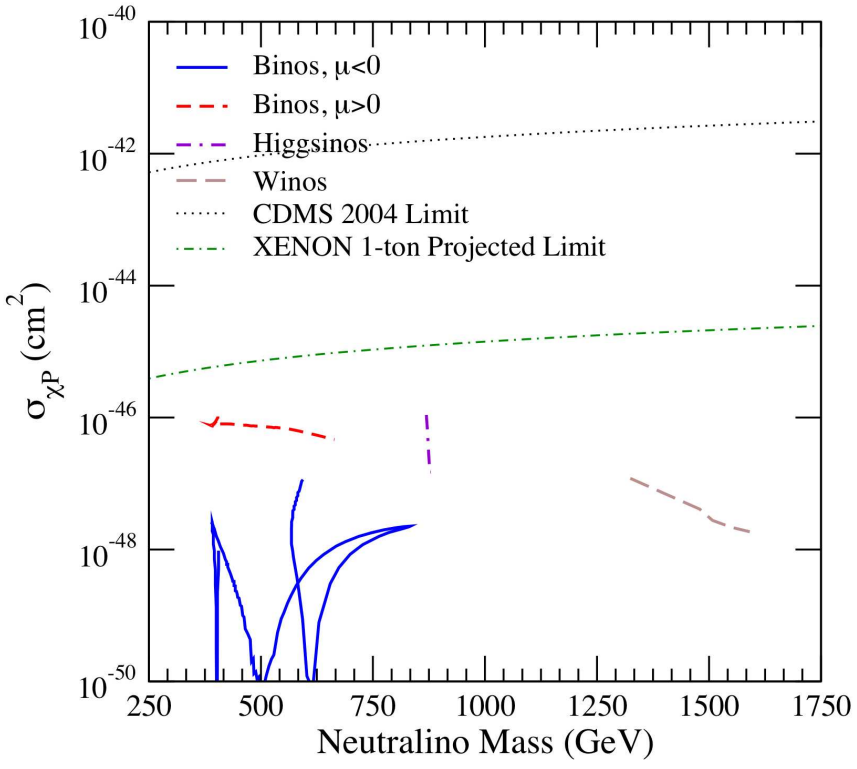

Figure 5.12: The spin-independent neutralino-proton scattering cross section, as a function of the neutralino mass, along the parameter-space points outlined in Fig. 5.8. We also indicate the current exclusion limits from the CDMS experiment [5·46], and the expected reach of the XENON 1-ton facility [5·47].

determined by the CDF Collaboration [5·50, 5·51], and the future reach of the LHC and of a future Linear Collider has been also assessed for this broad class of exotic new particles [5·52, 5·53]. The case of a stable[3] stau has been, in particular, considered several times, since it occurs in the context of various SUSY-breaking models, like gauge-mediated SUSY-breaking (GMSB) scenarios [5·54]; see also, e.g., Ref. [5·55, 5·56]. We remark that, contrary to GMSB, in the present framework the production of long lived staus at accelerators is expected to come along together with the production of neutralinos, thus making the two scenarios distinguishable, at least in principle.

A long-lived stau would behave as a highly penetrating particle, appearing in the tracking and muon chambers of collider detectors with a small energy deposit in calorimeters. Staus, depending on their velocities, would produce either highly-ionizing tracks in the low-β regime, or, if quite relativistic, they would appear similar to energetic muons. In this latter case, one could look at excesses of dimuon or multi-lepton events as a result of superparticle production; for example, considering the ratio $\sigma(\mu^+\mu^-)/\sigma(e^+e^-)$. In case additional particles have masses close to the neutralino-stau system, the total production cross section of superparticles would be greatly enhanced. Long-lived charginos may also give interesting accelerator signals [5·57].

While the reach of the Tevatron appears to be insufficient to probe the parameter space of the models we are considering here [5·55], the discovery of this kind of scenarios at the LHC, though challenging, looks quite conceivable, particularly in the case of binos. In fact, even in the less promising case in which the particle spectrum does not feature any particle close in mass with the $\chi_1^0 - \widetilde{\tau}_1$ system, the highly-ionizing–track channel should cover a mass range widely overlapping that indicated in Fig. 5.8. On the other hand, excess dimuon events could provide an independent confirmation, although the 5-σ LHC reach for CHAMP's in this channel alone has been assessed to lie around 300 GeV [5·55]. Moreover, if additional coannihilating particles (charginos, smuons, or selectrons) are present, the discovery at the LHC would certainly look even more promising [5·56].

A further recently proposed detection technique for long-lived staus is represented by trapping these particles into large water tanks placed outside the LHC detectors [5·59]. Following the results of Ref. [5·59], a 10-kton water tank may be capable of trapping more than 10 staus per year, if the mass $m_{\widetilde{\tau}_1} \simeq 400$ GeV. The subsequent decays could then be studied in a background-free environment. More than twice as many sleptons would also get trapped in the LHC detectors, although in this case a study of the stau properties would look more challenging [5·59]. Other methods of trapping and detecting staus have also been proposed [5·60].

[3]Here *stable* simply means that the decay length is much larger than the detector size.

5.10 Discussion

We have examined a scenario in which a fraction f_ϕ of the cold-dark-matter component in the Universe is generated in the decay of long-lived charged particles and shown that a scale-dependent (or 'running') spectral index is induced. The power spectrum on scales smaller than the horizon size when the age of the Universe is equal to the charged-particle lifetime gets suppressed by a factor $(1-f_\phi)^2$. Such a feature might be singled out unambiguously by future measurements of the power spectrum of neutral hydrogen through the 21-cm line, obtaining direct information on the charged particle lifetime and f_ϕ.

On the contrary, current and future tests for departures from a scale-invariant power spectrum based on Lyman-α data may fail to uniquely identify the scenario we propose. In fact, we have estimated the modifications to the non-linear power spectrum at the redshifts and wavenumbers currently probed by Lyman-α forest data, and shown that these resemble (but are different in detail) those in models with a constant running of the spectral index α_s. We expect, based on this resemblance, models with f_ϕ in the range $1/2-1/7$ (as predicted in some explicit models we have constructed) to be compatible with current cosmological data for lifetimes in the range $1-20$ yr. We have also verified that constraints from the primordial light-element abundances and distortions to the CMB spectrum are not violated.

From the particle-physics point of view, we have shown that the proposed scenario fits nicely in a picture in which the lightest neutralino in SUSY extensions of the standard model appears as the cold-dark-matter candidate, and a stau nearly degenerate in mass with the neutralino as the long-lived charged counterpart. A small mass splitting forces the stau to be quasi-stable, since the phase space allowed in the its decay process gets sharply reduced. At the same time, it implies that neutralino and stau are strongly linked in the process of thermal decoupling, with the charged species playing the major role. Owing to these coannihilation effects, the current neutralino thermal relic abundance is compatible with the value inferred from cosmological observatations and, at early times, the stau thermal relic component is at the correct level.

We have described several explicit realizations of this idea in minimal supersymmetric frameworks, including the minimal supergravity scenario, namely the supersymmetric extension of the standard model with smallest possible parameter space. We have pointed out that charged dark matter fraction from 1/2 to 1/7 (or even lower) can be obtained and that stau lifetimes larger than 1 yr are feasible. We have also shown that some of the models we have considered may be detected in future WIMP direct searches, and discussed the prospects of testing the most dramatic feature of the model we propose, i.e., the production of long-lived staus at future high-energy particle colliders.

Acknowledgments

We thank R. R. Caldwell and U. Seljak for useful discussions. KS acknowledges the support of a Canadian NSERC Postgraduate Scholarship. This work was supported in part by NASA NAG5-9821 and DoE DE-FG03-92-ER40701.

Bibliography

[5·1] P. de Bernardis *et al.* [Boomerang Collaboration], Nature **404**, 955 (2000); S. Hanany *et al.*, Astrophys. J. **545**, L5 (2000); N. W. Halverson *et al.*, Astrophys. J. **568**, 38 (2002); B. S. Mason *et al.*, Astrophys. J. **591**, 540 (2003); A. Benoit *et al.* [the Archeops Collaboration], Astron. Astrophys. **399**, L25 (2003); D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148**, 175 (2003);

[5·2] A. Gould, B. T. Draine, R. W. Romani and S. Nussinov, Phys. Lett. B **238**, 337 (1990).

[5·3] G. D. Starkman, A. Gould, R. Esmailzadeh, and S. Dimopoulos, Phys. Rev. D **41**, 3594 (1990).

[5·4] E. D. Carlson, M. E. Machacek, and L. J. Hall, Astrophys. J. **398**, 43 (1992).

[5·5] D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84**, 3760 (2000).

[5·6] S. Davidson, S. Hannestad, and G. Raffelt, JHEP **0005**, 003 (2000).

[5·7] S. L. Dubovsky, D. S. Gorbunov, and G. I. Rubtsov, JETP Lett. **79**, 1 (2004) [Pisma Zh. Eksp. Teor. Fiz. **79**, 3 (2004)].

[5·8] K. Sigurdson, M. Doran, A. Kurylov, R. R. Caldwell, and M. Kamionkowski, Phys. Rev. D **70**, 083501 (2004).

[5·9] M. S. Turner, Phys. Rept. **197**, 67 (1990); G. G. Raffelt, Phys. Rept. **198**, 1 (1990); L. J. Rosenberg and K. A. van Bibber, Phys. Rept. **325**, 1 (2000).

[5·10] G. Jungman, M. Kamionkowski, and K. Griest, Phys. Rept. **267**, 195 (1996).

[5·11] L. Bergstrom, Rept. Prog. Phys. **63**, 793 (2000).

[5·12] J. L. Feng, A. Rajaraman, and F. Takayama, Phys. Rev. Lett. **91**, 011302 (2003).

[5·13] J. L. Feng, S. f. Su and F. Takayama, Phys. Rev. D **70**, 063514 (2004).

[5·14] J. R. Ellis, K. A. Olive, Y. Santoso and V. C. Spanos, Phys. Lett. B **588**, 7 (2004).

[5·15] K. Jedamzik, Phys. Rev. D **70**, 083510 (2004).

[5·16] J. L. Feng, A. Rajaraman and F. Takayama, Phys. Rev. D **68**, 063504 (2003).

[5·17] K. Sigurdson and M. Kamionkowski, Phys. Rev. Lett. **92**, 171302 (2004).

[5·18] U. Seljak *et al.*, arXiv:astro-ph/0407372.

[5·19] U. Seljak, P. McDonald and A. Makarov, Mon. Not. Roy. Astron. Soc. **342**, L79 (2003).

[5·20] M. Tegmark *et al.* [SDSS Collaboration], Phys. Rev. D **69**, 103501 (2004).

[5·21] A. Loeb and M. Zaldarriaga, Phys. Rev. Lett. **92**, 211301 (2004).

[5·22] A. H. Chamseddine, R. Arnowitt and P. Nath, Phys. Rev. Lett. **49**, 970 (1982); R. Barbieri, S. Ferrara and C. A. Savoy, Phys. Lett. B **119**, 343 (1982); L. J. Hall, J. Lykken and S. Weinberg; Phys. Rev. D **27**, 2359 (1983); P. Nath, R. Arnowitt and A. H. Chamseddine, Nucl. Phys. B **227**, 121 (1983).

[5·23] L. Randall and R. Sundrum, Nucl. Phys. B **557**, 79 (1999); G. F. Giudice, M. A. Luty, H. Murayama and R. Rattazzi, JHEP **9812**, 027 (1998); T. Gherghetta, G. F. Giudice and J. D. Wells, Nucl. Phys. B **559**, 27 (1999).

[5·24] C. P. Ma and E. Bertschinger, Astrophys. J. **455**, 7 (1995).

[5·25] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996).

[5·26] R. H. Cyburt, J. R. Ellis, B. D. Fields, and K. A. Olive, Phys. Rev. D **67**, 103521 (2003).

[5·27] K. Jedamzik, Phys. Rev. Lett. **84**, 3248 (2000).

[5·28] W. Hu and J. Silk, Phys. Rev. Lett. **70**, 2661 (1993).

[5·29] J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, Astrophys. J. **473**, 576 (1996).

[5·30] R. E. Smith *et al.* [The Virgo Consortium Collaboration], Mon. Not. Roy. Astron. Soc. **341**, 1311 (2003).

[5·31] J. A. Peacock and S. J. Dodds, Mon. Not. Roy. Astron. Soc. **280**, L19 (1996).

[5·32] M. J. White and R. A. C. Croft, Astrophys. J. **539**, 497 (2000).

[5·33] G. Servant and T. M. P. Tait, Nucl. Phys. B **650**, 391 (2003).

[5·34] K. Griest and M. Kamionkowski, Phys. Rev. Lett. **64**, 615 (1990).

[5·35] K. Griest and D. Seckel, Phys. Rev. D **43**, 3191 (1991).

[5·36] J. Edsjo and P. Gondolo, Phys. Rev. D **56**, 1879 (1997).

[5·37] J. R. Ellis, K. A. Olive, Y. Santoso, and V. C. Spanos, Phys. Lett. B **565**, 176 (2003).

[5·38] D. M. Pierce, J. A. Bagger, K. T. Matchev, and R. j. Zhang, Nucl. Phys. B **491**, 3 (1997).

[5·39] H. C. Cheng, B. A. Dobrescu, and K. T. Matchev, Nucl. Phys. B **543**, 47 (1999).

[5·40] M. Muhlleitner, A. Djouadi, and Y. Mambrini, hep-ph/0311167.

[5·41] K. L. Chan, U. Chattopadhyay, and P. Nath, Phys. Rev. D **58**, 096004 (1998).

[5·42] J. L. Feng, K. T. Matchev, and T. Moroi, Phys. Rev. D **61**, 075005 (2000).

[5·43] S. Profumo, Phys. Rev. D **68**, 015006 (2003).

[5·44] M. Drees, Phys. Lett. B **181**, 279 (1986).

[5·45] K. Hagiwara *et al.* [Particle Data Group Collaboration], Phys. Rev. D **66**, 010001 (2002).

[5·46] D. S. Akerib *et al.* [CDMS Collaboration], Phys. Rev. Lett. **93**, 211301 (2004).

[5·47] E. Aprile *et al.*, arXiv:astro-ph/0207670.

[5·48] J. Edsjö, Internal Amanda/IceCube Report, (2000).

[5·49] H. Baer, C. Balazs, A. Belyaev, T. Krupovnickas, and X. Tata, JHEP **0306**, 054 (2003).

[5·50] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **90**, 131801 (2003).

[5·51] F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **46**, 1889 (1992).

[5·52] A. Nisati, S. Petrarca, and G. Salvini, Mod. Phys. Lett. A **12**, 2213 (1997).

[5·53] P. G. Mercadante, J. K. Mizukoshi, and H. Yamamoto, Phys. Rev. D **64**, 015005 (2001).

[5·54] G. F. Giudice and R. Rattazzi, Phys. Rept. **322**, 419 (1999).

[5·55] J. L. Feng and T. Moroi, Phys. Rev. D **58**, 035001 (1998).

[5·56] H. Baer, P. G. Mercadante, X. Tata, and Y. L. Wang, Phys. Rev. D **62**, 095007 (2000).

[5·57] J. L. Feng, T. Moroi, L. Randall, M. Strassler, and S. f. Su, Phys. Rev. Lett. **83**, 1731 (1999).

[5·58] H. Baer, J. K. Mizukoshi and X. Tata, Phys. Lett. B **488**, 367 (2000); A. J. Barr, C. G. Lester, M. A. Parker, B. C. Allanach and P. Richardson, JHEP **0303**, 045 (2003).

[5·59] J. L. Feng and B. T. Smith, Phys. Rev. D **71**, 015004 (2005) [Erratum-ibid. D **71**, 0109904 (2005)].

[5·60] K. Hamaguchi, Y. Kuno, T. Nakaya and M. M. Nojiri, Phys. Rev. D **70**, 115007 (2004).

Chapter 6

Cosmic 21-cm Delensing of Microwave Background Polarization and the Minimum Detectable Energy Scale of Inflation

We propose a new method for removing the effect of gravitational lensing from maps of cosmic microwave background (CMB) polarization anisotropies. Using observations of anisotropies or structures in the cosmic 21-cm radiation, which was emitted or absorbed by neutral hydrogen atoms that underwent a spin-flip transition at redshifts 10 to 200, the CMB can be delensed. We find that this method could allow CMB experiments to have increased sensitivity to a primordial background of inflationary gravitational waves (IGWs) compared to methods which rely on CMB observations alone — reducing the minimum detectable energy scale of inflation below 10^{15} GeV. While the detection of cosmic 21-cm anisotropies at high resolution is a challenging endeavor, the detection of these fluctuations is already being pursued as a probe of the Universe at or before the epoch of reionization. A combined study with a relatively low-resolution (but high-sensitivity) CMB polarization experiment may constrain alternative models of inflation which were heretofore considered to have undetectable IGW amplitudes. The ultimate theoretical limit to the detectable inflationary energy scale via this method may be as low as 3×10^{14} GeV.

Originally available online as K. Sigurdson and A. Cooray, arXiv:astro-ph/0502549. Submitted to *Phys. Rev. Lett.*

6.1 Introduction

The curl (B) modes of cosmic microwave background (CMB) polarization anisotropies are a unique probe of the primordial background of cosmological gravitational waves [6·1]. At these long wave-

lengths, inflation [6·2] is the only known mechanism to causally generate such a background of gravitational waves [6·3] . Since the amplitude of these inflationary gravitational waves (IGWs) is proportional to $\mathcal{V}$, the value of the inflaton potential $V(\varphi)$ during inflation, the amplitude of gravitational-wave induced B-mode polarization anisotropies directly constrains the energy scale of inflation $\mathcal{V}^{1/4}$ (see, for example, Ref. [6·4]). While the experimental sensitivity to B-mode polarization can be improved, the expected signal is contaminated by foreground effects [6·5]. The main confusion to the detection of B-mode polarization anisotropies generated by IGWs at recombination is the mixing of gradient-mode (E-mode) and B-mode anisotropies via gravitational lensing [6·6].

In this chapter we propose a new method for separating lensing-induced B modes from the IGW signal using observations of anisotropies in the cosmic 21-cm radiation emitted or absorbed by neutral hydrogen atoms at redshifts 10 to 200. While the detection of cosmic 21-cm anisotropies at high resolution is challenging, a combined study with CMB polarization data could probe inflationary energy scales well below the Grand Unified Theory (GUT) scale of 10^{16} GeV — constraining inflationary models with energy scales below 10^{15} GeV. The ultimate theoretical limit to the minimum detectable energy scale of inflation via this method could reach as low as 3×10^{14} GeV.

6.2 Gravitational Lensing

Lensing induces a remapping of the polarization field at the last-scattering surface ${}_{\pm}X(\hat{\mathbf{n}})$ such that ${}_{\pm}\tilde{X}(\hat{\mathbf{n}}) = {}_{\pm}X[\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})]$ is the observed polarization field, where ${}_{\pm}X = Q \pm iU$ are linear combinations of the Stokes parameters Q and U and $\alpha(\hat{\mathbf{n}}) = \nabla\phi(\hat{\mathbf{n}})$ is the lensing deflection angle. Here,

$$\phi(\hat{\mathbf{n}}; z_s) \quad = \quad -2\int_0^{r(z_s)} dr' \frac{r - r'}{r'r}\Phi(\hat{\mathbf{n}}, r') \tag{6.1}$$

is the deflection potential, a line-of-sight projection of the gravitational potential Φ to redshift z_s. The total lensing potential $\phi(\hat{\mathbf{n}}) \equiv \phi(\hat{\mathbf{n}}; z_{\rm CMB})$ is this quantity evaluated at $z_s \to z_{\rm CMB} \approx 1100$.

Using the flat-sky approximation and the E-mode/B-mode decomposition [6·1], the lensed B-mode polarization power spectrum, in the relevant limit $C_l^{BB} << C_l^{EE}$, is

$$\tilde{C}_l^{BB} = C_l^{BB} + \int \frac{d^2\mathbf{l}'}{(2\pi)^2} \left[\mathbf{l}'' \cdot \mathbf{l}'\right]^2 \sin^2(2\theta_l') C_{l''}^{\phi\phi} C_{l'}^{EE}\,, \tag{6.2}$$

where $\mathbf{l}'' = \mathbf{l} - \mathbf{l}'$. The second term in this expression is the lensing confusion in the B-mode map which must be separated from C_l^{BB} — the IGW signal. Here, $C_l^{\phi\phi}$ is the angular power spectrum of the total deflection potential and is simply related to a weighted projection of the matter power spectrum [6·6]; $C_l^{\phi\phi}(z_s)$ is the incomplete power spectrum out to source redshift $z_s < z_{\rm CMB}$.

Unlike the B modes generated by tensor perturbations (the IGWs), C_l^{EE} is dominated by larger

amplitude scalar perturbations. The expected few-percent conversion of E modes creates a large signal in the B-mode power spectrum [6·6]. For tensor-to-scalar ratios $\mathcal{T}/\mathcal{S}$ below 2.6×10^{-4} or, since $\mathcal{V}^{1/4} = 3.0 \times 10^{-3}(\mathcal{T}/S)^{1/4} m_{Pl}$ [6·7], $\mathcal{V}^{1/4}$ below 4.6×10^{15} GeV the IGW signal is completely confused by the lensing contaminant [6·8]. To bypass this limit one must separate the lensing induced B-modes from those due to IGWs. Clearly, the lensing confusion could be exactly removed if one knew the three-dimensional distribution of mass out to the CMB last-scattering surface. However, as our goal is a measurement of C_l^{BB}, knowledge of the projected quantity $\phi(\hat{\mathbf{n}})$ is sufficient. One way to estimate $\phi(\hat{\mathbf{n}}; z_s)$ is by using quadratic estimators or maximum likelihood methods to statistically infer the deflection-angle field given some lensed random field $\tilde{\chi}(\hat{\mathbf{n}})$ at source redshift z_s. For instance, arcminute resolution CMB temperature and polarization maps could be used to make such an estimate [6·9–6·11]. Another way to estimate $\phi(\hat{\mathbf{n}}; z_s)$ is by observing the weak-lensing distortions of the shapes of objects of a known average shape at source redshift z_s. We note here that observations of the weak lensing of galaxies, which have $z_s \sim 1$–2, can not be used to delense CMB maps because a large fraction of the lensing contamination (55% at $l = 1000$) comes from structure at $z > 3$. Higher source redshifts are required for effective delensing. We first review the potential observational signatures of the cosmic 21-cm radiation, and then discuss the methods for delensing the CMB outlined above.

6.3 Cosmic 21-cm Radiation

Neutral atoms kinetically decouple from the thermal bath of CMB photons at $z \sim 200$ and cool adiabatically as $T_{\rm g} \propto (1+z)^2$ [6·12]. Since the spin temperature of the hydrogen atoms remains collisionally coupled to T_g these atoms resonantly absorb CMB photons at $\lambda_{21} = 21.1$ cm — the hyperfine transition of the ground state of hydrogen. The cosmic 21-cm radiation is thus first observable in absorption by low-frequency radio telescopes which could detect brightness-temperature fluctuations at wavelength $\lambda = \lambda_{21}(1+z)$ [6·13–6·15]. During reionization, the neutral gas distribution is likely to be complex due to the first luminous sources [6·16] and cosmic 21-cm signatures shift to emission [6·17]. Yet, even before reionization, it is possible the 21-cm sky is brightened by emission from neutral hydrogen gas contained in minihalos with masses $\sim 10^3$–10^7 $\mathrm{M}_\odot$ [6·18].

Like the CMB, the statistics of the high-z absorption fluctuations are expected to be Gaussian and quadratic estimators of the lensing potential, described below, could be straightforwardly adapted to reconstruct the deflection field. If the statistics of the 21-cm fluctuations during reionization can be understood the lensing of the 21-cm emission from that era might also provide a useful probe of the lensing potential. The most promising (but futuristic) possibility involves using the shape statistics of high-redshift minihalos to infer the lensing potential.

6.4 Quadratic Estimators

Quadratic estimators can be used to extract lensing information from the gravitationally-lensed field $\tilde{\chi}(\hat{\mathbf{n}})$ of some intrinsic field $\chi(\hat{\mathbf{n}})$ at redshift z_s. The quadratic form $\nabla \cdot [\chi(\hat{\mathbf{n}})\nabla\chi(\hat{\mathbf{n}})]$ provides an estimate of the deflection angle at position $\hat{\mathbf{n}}$ on the sky given the $\tilde{\chi}$ anisotropy map. For the CMB, the quantity $\tilde{\chi}$ could be the temperature anisotropies [6·10], the polarization anisotropies, or some combination of both [6·11]. The brightness temperature fluctuations in the 21-cm transition of neutral hydrogen from redshifts 10 to 200 could similarly be used [6·19].

In Fourier space, the quadratic estimator for the deflection potential is

$$\hat{\phi}(\mathbf{l}; z_s) = \mathcal{Q}_l(z_s) \int \frac{d^2\mathbf{l}'}{(2\pi)^2} \left(\mathbf{l}\cdot\mathbf{l}' C_{l'}^{\chi\chi} + \mathbf{l}\cdot\mathbf{l}'' C_{l''}^{\chi\chi}\right) \frac{\chi(l')\chi(l'')}{2T_{l'}^{\chi\chi}T_{l''}^{\chi\chi}}\,, \tag{6.3}$$

where $C_l^{\chi\chi}$ is the unlensed power spectrum and $T_l^{\chi\chi} = \tilde{C}_l^{\chi\chi} + N_l^{\chi\chi}$ is to total power spectrum, including lensing corrections and a noise power spectrum $N_l^{\chi\chi}$. The expectation value of the deflection-potential estimator $\langle\hat{\phi}(\mathbf{l}; z_s)\rangle$ (the ensemble average over realizations of the random field χ) is just $\phi(\mathbf{l}; z_s)$. Here,

$$[\mathcal{Q}_l(z_s)]^{-1} = \int \frac{d^2\mathbf{l}'}{(2\pi)^2} \frac{\left(\mathbf{l}\cdot\mathbf{l}' C_{l'}^{\chi\chi} + \mathbf{l}\cdot\mathbf{l}'' C_{l'}^{\chi\chi}\right)^2}{2T_{l'}^{\chi\chi}T_{l''}^{\chi\chi}} \tag{6.4}$$

is the noise power spectrum associated with a quadratic reconstruction of $C_l^{\phi\phi}(z_s)$ using the field $\tilde{\chi}$ [6·10].

6.5 Partial Delensing Bias

An estimate of $\phi(\hat{\mathbf{n}})$ can be used to delense the CMB B-mode polarization map. In the limit $z_s \to z_{\rm CMB}$ (conventional CMB delensing) the extraction of C_l^{BB} from the delensed map is limited by the noise introduced during delensing. The residual contamination of the B-modes is given by the second term of Eq. (6.2) with the replacement $C_l^{\phi\phi} \to \mathcal{Q}_l$. However if $z_s < z_{\rm CMB}$ this noise is not necessarily the factor limiting a measurement of the IGW signal. Using $\hat{\phi}(\hat{\mathbf{n}}; z_s)$ as a proxy for $\phi(\hat{\mathbf{n}})$ to delense the map leaves a residual lensing contamination not due to noise. Accounting for this partial delensing bias $\mathcal{B}_l(z_s) \equiv C_l^{\phi\phi} - C_l^{\phi\phi}(z_s)$ (due to the difference in source redshift between the lensed field $\tilde{\chi}(\hat{\mathbf{n}})$ and the CMB) the residual contamination of the B-mode power spectrum is instead the second term of Eq. (6.2) with $C_l^{\phi\phi} \to \mathcal{B}_l(z_s) + \mathcal{N}_l(z_s)$. This is true whether $\hat{\phi}(\hat{\mathbf{n}}; z_s)$ is estimated using quadratic estimators or by some other method. Here, $\mathcal{N}_l(z_s)$ is the residual noise power spectrum of the deflection potential due to noise associated with the delensing process — for quadratic reconstruction $\mathcal{N}_l(z_s) = \mathcal{Q}_l(z_s)$. If the deflection potential is reconstructed from a line, as is the case for cosmic 21-cm radiation, the source redshift is exactly know and $\mathcal{B}_l(z_s)$ can be reliably estimated.

6.6 Quadratic Reconstruction

Unlike the CMB anisotropies, which lack power on angular scales below a few arcminutes due to Silk damping, the cosmic 21-cm anisotropies extend to much higher values of l (limited by the Jeans wavelength of the gas) and peak in amplitude at higher values of l [6·14, 6·15]. Additionally, measurements of cosmic 21-cm anisotropies in different frequency bins provide several estimates of essentially the same deflection field.

As shown in Fig. 6.1, we estimate a 21-cm experiment centered around $z_s \sim 30$ with a 20 MHz coverage in frequency space capable of observing anisotropies out to $l \sim 5000$ would have an $\mathcal{N}_l(z_s)$ higher than the planned CMBpol mission. In this case residual confusion arises from noise rather than bias. A quadratic reconstruction of the deflection field using this type of measurement, in conjunction with a CMB polarization experiment, could detect $\mathcal{T}/\mathcal{S} \gtrsim 2.5 \times 10^{-5}$ or $\mathcal{V}^{1/4} > 2.6 \times 10^{15}$ GeV. A next-generation 21-cm experiment capable of observing anisotropies out to $l \sim 10^5$ with the same central redshift and bandwidth would have a $\mathcal{N}_l(z_s)$ an order of magnitude below CMBpol and now be limited largely by the bias $\mathcal{B}_l(z_s)$. Paired with CMB polarization observations, this type of measurement could detect $\mathcal{T}/\mathcal{S} \gtrsim 1.0 \times 10^{-6}$ or $\mathcal{V}^{1/4} > 1.1 \times 10^{15}$ GeV — comparable to very high sensitivity and resolution future CMB observations alone [6·20]. However, as the CMB data would not need to be used to reconstruct the deflection field a much lower resolution CMB experiment would suffice. Furthermore, if lensing information need not be extracted from the CMB observations, the optimal observing strategy is to integrate over a few square degree patch of the sky as proposed in Ref. [6·22]. Such a CMB experiment could thus be ground based.

6.7 Other Methods

If minihalos bright in 21-cm emission exist in the early Universe, just as galaxy shapes are sheared by weak gravitational lensing so will be the shapes of these minihalos. Ellipticity information obtained from such 21-cm minihalos could be used to reconstruct the projected potential out to high z [6·23]. Based on the dark-matter halo mass function, we expect roughly a surface density of 10^{11}/sr of such minihalos at $z \sim 30$ for a bandwidth of 1 MHz with masses between 10^5 and 10^7 $M_\odot$. A typical halo of mass 10^6 $M_\odot$ has a characteristic projected angular size of ~ 60 milliarcseconds. If resolved, then techniques currently applied to measure shear in background galaxies in the low-z Universe could be adapted for this application. Regardless of the exact method, a deflection potential reconstructed from high-redshift 21-cm observations could then be used to delense CMB B-mode maps.

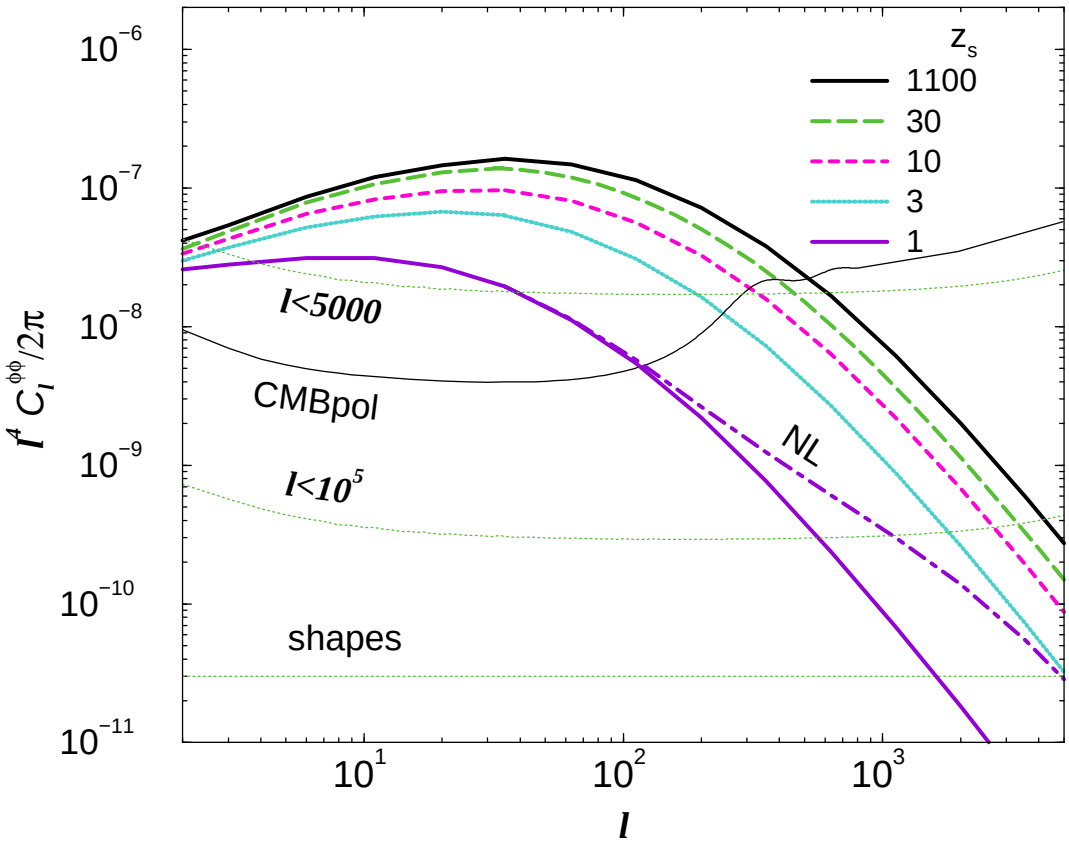

Figure 6.1: Shown is the angular power spectrum of the deflection potential as a function of source redshift z_s. The curves labeled '$l < 5000$' and '$l < 10^5$' are the estimated noise levels for quadratic reconstruction using 21-cm anisotropies in 40 0.5 MHz bins centered around $z_s \approx 30$. We assume a noise power spectrum with $T_{sys} = 3000K$ at 46 MHz, that $l_{\rm max} f_{\rm cov} \approx 15$, and a year of integration [6·17]. The curve labeled shapes shows the residual noise curve in a scenario where shear is directly measured using resolved minihalos in a 1 MHz bandwidth about $z_s \approx 30$. Also shown is the noise levels for a CMB reconstruction of deflections with the planned CMBpol mission assuming a 3 arcminute beam, a noise level of 1μK $\sqrt{\rm sec}$, and a year of integration.

6.8 Bias-Limited Delensing

To understand to what extent bias-limited reconstructions, where the residual lensing contamination is dominated by $\mathcal{B}_l(z_s)$, would result in the removal of lensing confusion we have calculated the residual B-mode power spectrum after correcting for the modified lensing kernel when $z_s < z_{\rm CMB}$. We have adapted the formalism of Ref. [6·21] to estimate the smallest detectable background of IGWs and the resulting limits are summarized in Fig. 2. While knowing the projected mass distribution out to $z_s = 1$ does not allow the confusion to be reduced significantly, if it is known to $z_s = 10$ the confusion is reduced by an order of magnitude and the minimum detectable energy scale of inflation is reduced below the limit derived using quadratic CMB statistics [6·21]. A lensing-source redshift $z_s \gtrsim 30$ would be required to improve beyond the practical 1.1×10^{15} GeV limit of the more

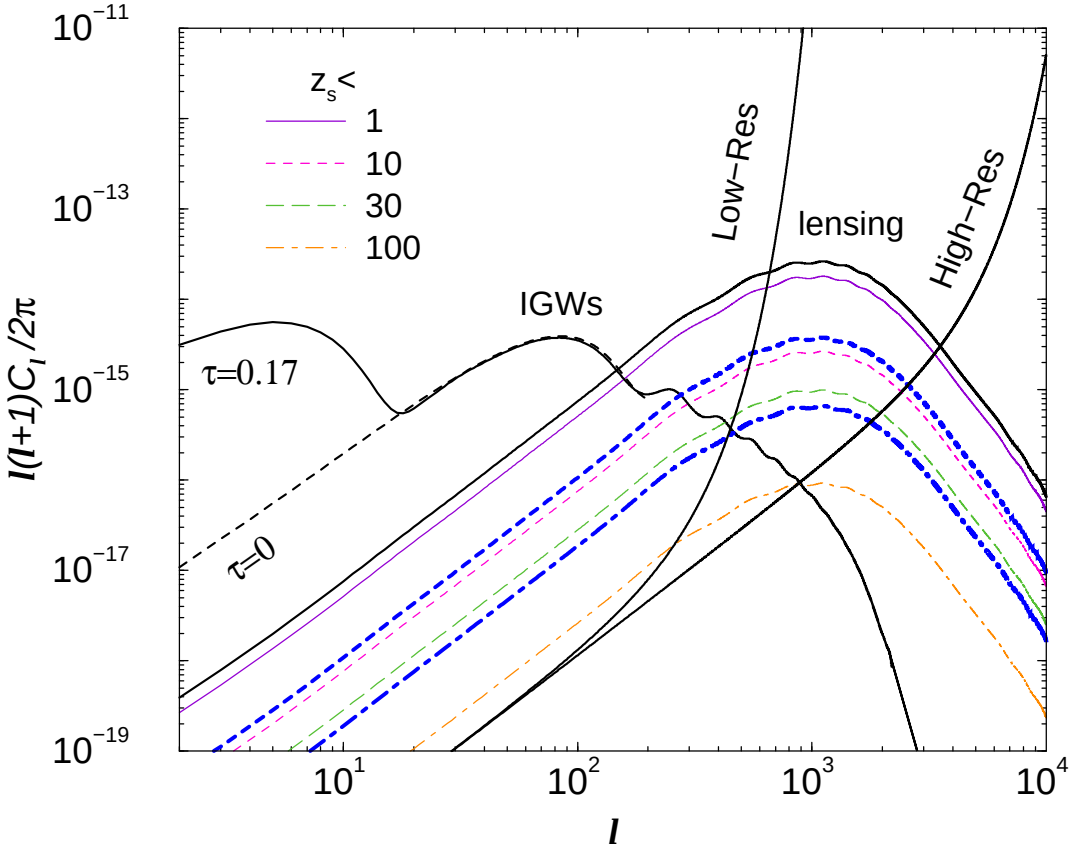

Figure 6.2: Shown are power spectra of CMB B-mode polarization. The curve labeled 'IGWs' is the IGW contribution to B-modes assuming a tensor-to-scalar ratio of 0.1 with (solid line; $\tau = 0.17$) and without (dashed line) reionization. The curve labeled 'lensing' is the total lensing confusion to B-modes. Thin lines show the residual B-mode lensing contamination for bias-limited delensing out to z_s. Thick lines show previous estimates of the residual confusion from CMB experiments alone using quadratic estimators with an ideal noise-free experiment (dashed line), and likelihood methods using a high resolution/sensitivity experiment (dot-dashed). The noise curve of latter this experiment, with 2 arcminute beams and a pixel noise of 0.25 μK-arcminute, is the curve labeled 'High-Res'. Bias-limited lensing information to $z_s \gtrsim 10$ improves upon the limit to the IGW amplitude based on quadratic statistics, and the likelihood level can be reached with $z_s \sim 30$. If $z_s \sim 100$ an additional order-of-magnitude in the IGW amplitude could be probed. The curve labeled 'Low-Res' is the noise curve for a lower resolution CMB polarization experiment with 30 arcminute beams sufficiently sensitive to detect IGWs when paired with a cosmic 21-cm lensing reconstruction. For very efficient delensing other foregrounds, such as patchy reionization [6·28], might dominate confusion.

sophisticated maximum likelihood method with high resolution and sensitivity CMB polarization observations [6·20]. However we emphasize here that in the case where a bias-limited reconstruction of the deflection field exists for $z_s \sim 30$, a very high-resolution CMB polarization experiment is not necessary, and a similarly sensitive experiment with a lower resolution could do the same job. Thus, it is conceivable that the Inflationary Probe of NASA's Beyond Einstein Program can be designed in combination with a cosmic 21-cm radiation experiment. Such a cosmic 21-cm radiation experiment could be extremely exciting in its own right [6·14, 6·24].

For a bias-limited reconstruction out to a $z_s \sim 100$, the limit on the tensor-to-scalar ratio is 7.0×10^{-8} or $\mathcal{V}^{1/4} > 6.0 \times 10^{14}$ GeV. For $z_s \sim 200$, the maximum redshift where 21-cm fluctuations are expected to be nonzero, one could probe down to $\mathcal{V}^{1/4} > 3 \times 10^{14}$ GeV.

6.9 Discussion

While obtaining bias-limited measurements out to $z_s \approx 100$ is a daunting task, with many experimental and theoretical obstacles to overcome, there is great interest in detecting the fluctuations in the cosmic 21-cm radiation at $z \sim 10 - 200$ for their own sake. The observational study of 21-cm fluctuations, especially during and prior to the era of reionization, is now being pursued by a variety of low frequency radio interferometers such as the Primeval Structure Telescope (PAST [6·25]), the Mileura Widefield Array (MWA), and the Low Frequency Array (LOFAR [6·26]). Planned interferometers such as the Square Kilometer Array (SKA) will improve both sensitivity and low-frequency coverage.

While certain models of inflation, those related to Grand Unified Theories (GUTs), are expected to have an energy scale $\mathcal{V}^{1/4}$ between 10^{15} GeV and 10^{16} GeV there are certainly other possibilities. For instance, some supersymmetric theories of inflation have energy scales of several times 10^{14} GeV [6·27]. New methods, such as the idea of using observations of the cosmic 21-cm radiation to delense the CMB B-mode polarization suggested here, are needed to push the minimum detectable energy scale of inflation below 10^{15} GeV and discriminate between between physical theories at the highest energy scales.

Acknowledgments

KS acknowledges the support of a Canadian NSERC Postgraduate Scholarship. This work was supported in part by NASA NAG5-9821 and DoE DE-FG03-92-ER40701.

Bibliography

[6·1] U. Seljak and M. Zaldarriaga, Phys. Rev. Lett. **78**, 2054 (1997); M. Kamionkowski, A. Kosowsky and A. Stebbins, Phys. Rev. Lett. **78**, 2058 (1997).

[6·2] A. H. Guth, Phys. Rev. D **23**, 347 (1981); A. D. Linde, Phys. Lett. B **108**, 389 (1982); A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).

[6·3] A. A. Starobinsky, JETP Lett. **30**, 682 (1979) [Pisma Zh. Eksp. Teor. Fiz. **30**, 719 (1979)].

[6·4] M. Kamionkowski and A. Kosowsky, Ann. Rev. Nucl. Part. Sci. **49**, 77 (1999).

[6·5] M. Tucci, E. Martinez-Gonzalez, P. Vielva and J. Delabrouille, arXiv:astro-ph/0411567.

[6·6] M. Zaldarriaga and U. Seljak, Phys. Rev. D **58**, 023003 (1998); W. Hu, Phys. Rev. D **62**, 043007 (2000).

[6·7] M. S. Turner and M. J. White, Phys. Rev. D **53**, 6822 (1996).

[6·8] A. Lewis, A. Challinor and N. Turok, Phys. Rev. D **65**, 023505 (2002).

[6·9] U. Seljak and M. Zaldarriaga, Phys. Rev. Lett. **82**, 2636 (1999); W. Hu, Astrophys. J. **557**, L79 (2001); M. H. Kesden, A. Cooray and M. Kamionkowski, Phys. Rev. D **67**, 123507 (2003); C. M. Hirata and U. Seljak, Phys. Rev. D **68**, 083002 (2003).

[6·10] W. Hu, Phys. Rev. D **64**, 083005 (2001).

[6·11] W. Hu and T. Okamoto, Astrophys. J. **574**, 566 (2002).

[6·12] P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton University Press, Princeton, 1993).

[6·13] P. A. Shaver, R. A. Windhorst, P. Madau and A. G. de Bruyn, Astron. Astrophys. **345**, 380 (1999).

[6·14] A. Loeb and M. Zaldarriaga, Phys. Rev. Lett. **92**, 211301 (2004).

[6·15] S. Bharadwaj and S. S. Ali, Mon. Not. Roy. Astron. Soc. **352**, 142 (2004).

[6·16] R. Barkana and A. Loeb, Phys. Rept. **349**, 125 (2001).

[6·17] M. Zaldarriaga, S. R. Furlanetto and L. Hernquist, Astrophys. J. **608**, 622 (2004).

[6·18] I. T. Iliev, P. R. Shapiro, A. Ferrara and H. Martel, Astrophys. J. **572**, L123 (2002); I. T. Iliev, E. Scannapieco, H. Martel and P. R. Shapiro, Mon. Not. Roy. Astron. Soc. **341**, 81 (2003).

[6·19] A. R. Cooray, New Astron. **9**, 173 (2004).

[6·20] U. Seljak and C. M. Hirata, Phys. Rev. D **69**, 043005 (2004).

[6·21] L. Knox and Y. S. Song, Phys. Rev. Lett. **89**, 011303 (2002); M. Kesden, A. Cooray and M. Kamionkowski, Phys. Rev. Lett. **89**, 011304 (2002).

[6·22] M. Kamionkowski and A. Kosowsky, Phys. Rev. D **57**, 685 (1998).

[6·23] U. L. Pen, New Astron. **9**, 417 (2004).

[6·24] S. Profumo, K. Sigurdson, P. Ullio and M. Kamionkowski, Phys. Rev. D **71**, 023518 (2005).

[6·25] U. L. Pen, X. P. Wu and J. Peterson, arXiv:astro-ph/0404083.

[6·26] H. Rottgering, New Astron. Rev. **47**, 405 (2003).

[6·27] G. G. Ross and S. Sarkar, Nucl. Phys. B **461**, 597 (1996); J. A. Adams, G. G. Ross and S. Sarkar, Phys. Lett. B **391**, 271 (1997).

[6·28] M. G. Santos, A. Cooray, Z. Haiman, L. Knox and C. P. Ma, Astrophys. J. **598**, 756 (2003).

Chapter 7

Measuring the Primordial Deuterium Abundance During the Cosmic Dark Ages

We discuss how measurements of fluctuations in the absorption of cosmic microwave background (CMB) photons by neutral gas during the cosmic dark ages, at redshifts $z \approx 7$–200, could reveal the primordial deuterium abundance of the Universe. The strength of the cross-correlation of brightness-temperature fluctuations due to resonant absorption of CMB photons in the 21-cm line of neutral hydrogen with those due to resonant absorption of CMB photons in the 92-cm line of neutral deuterium is proportional to the fossil deuterium to hydrogen ratio [D/H] fixed during big bang nucleosynthesis (BBN). Although technically challenging, this measurement could provide the cleanest possible determination of [D/H], free from contamination by structure formation processes at lower redshifts, and has the potential to improve BBN constraints to the baryon density of the Universe $\Omega_{\rm b}h^2$. We also present our results for the thermal spin-change cross-section for deuterium-hydrogen scattering, which may be useful in a more general context than we describe here.

Originally available online as K. Sigurdson and S. R. Furlaneto, arXiv:astro-ph/0505173. Submitted to *Phys. Rev. Lett.*

7.1 Introduction

After the cosmic microwave background (CMB) radiation decoupled from the baryons at a redshift $z \approx 1100$, most CMB photons propagated unfettered through the neutral primordial medium. This has allowed exquisite measurements of the temperature fluctuations in the primordial plasma at the surface of last scattering, and the statistical properties of these fluctuations have recently been used, in conjunction with other observations, to determine the cosmology of our Universe [7·1]. After the photons kinetically decoupled from the gas at $z \sim 200$, the latter cooled adiabatically with

$T_{\rm g} \propto (1+z)^2$, faster than the $T_\gamma \propto (1+z)$ cooling of the CMB. This epoch, with most of the baryons in the form of relatively cold neutral atoms and before the first stars formed, is known as the cosmic dark ages.

The reason most CMB photons propagate unimpeded through the neutral primordial gas is elementary quantum mechanics — atoms absorb non-ionizing radiation only at the discrete wavelengths determined by the differences of their atomic energy levels. One interesting example is the well-known 21-cm spin-flip transition [7·2], due to the hyperfine splitting of the ground state of the hydrogen (H) atom. At any given z, CMB photons with wavelength $\lambda_{21} = 21.1$ cm can resonantly excite this transition. By measuring brightness-temperature fluctuations due to density fluctuations in the neutral gas [7·3], radio telescopes observing at $\lambda = (1+z)\lambda_{21}$ can probe the matter power spectrum at $z \approx 30$–200 [7·4].

In this chapter we discuss another application of these measurements. Less well-known than the 21-cm transition of neutral H is the spin-flip transition of neutral deuterium (D) at $\lambda_{92} = 91.6$ cm [7·5,7·6]. We show below that cross-correlating brightness-temperature fluctuations at a wavelength $\lambda_{\rm H} = (1+z)\lambda_{21}$ with those at a wavelength $\lambda_{\rm D} = (1+z)\lambda_{92}$ allows a measurement of the primordial D abundance. In principle, this technique could constrain the primordial value of $[{\rm D/H}] \equiv n_{\rm D}/n_{\rm H}$ to better than 1%. While there is no physical obstacle to such a measurement, it would certainly be technically challenging, and require a heroic experimental effort; simply detecting neutral D during the cosmic dark ages would be a significantly easier goal.

Deuterium has long been recognized as our best 'baryometer' because its primeval relic abundance is so sensitive to the baryon-to-photon ratio $\eta = n_{\rm b}/s$. Moreover, big bang nucleosynthesis (BBN) [7·7] is the only known natural production mechanism, although mechanisms inside galaxies can destroy it [7·8]. The measurement we describe below could thus determine the true BBN abundance of D and, in principle, might improve BBN constraints to the baryon density of the Universe $\Omega_{\rm b}h^2$.

7.2 Hyperfine Structure of H and D Atoms

The $\boldsymbol{\mu}\cdot\mathbf{B}$ interaction between the magnetic moments of the electron and the nucleus splits the ground state of single-electron atoms into eigenstates of the total spin operator $\mathbf{F} = \mathbf{S} + \mathbf{I}$ with eigenvalues $F_+ = I + 1/2$ and $F_- = I - 1/2$ and $\Delta E = (16/3)F_+\mu_{\rm B}(g_{\rm N}\mu_{\rm N}/a_0^3)$ (e.g., [7·9]). Here, $\mathbf{S}$ is electron spin, $\mathbf{I}$ is nuclear spin, a_0 is the Bohr radius, $\mu_{\rm B}$ is the Bohr magneton, $\mu_{\rm N}$ is the nuclear magneton, and $g_{\rm N}$ is the nuclear g factor ($g_{\rm p} = 5.56$ for H; $g_{\rm D} = 0.857$ for D). The proton, with $I = 1/2$, splits the H ground state into a triplet with $F_+ = 1$ and a singlet with $F_- = 0$. The deuteron, with $I = 1$, splits the D ground state into a quartet with $F_+ = 3/2$ and a doublet with $F_- = 1/2$.

7.3 The Spin-Temperature

The population of atoms in the excited spin state relative to the ground state

$$(n_+/n_-) = (g_+/g_-)e^{-T_\star/T_{\rm s}} \tag{7.1}$$

can be characterized by a spin temperature $T_{\rm s}$. Here $g_+ = 2F_+ + 1$ and $g_- = 2F_- + 1$ are the spin degeneracy factors and $T_\star = \Delta E/k_{\rm B}$. For H and D, we have respectively $T_\star^{\rm H} = 0.0682$ K, $T_\star^{\rm D} = 0.0157$ K, and $(g_+^{\rm H}/g_-^{\rm H}) = 3$, $(g_+^{\rm D}/g_-^{\rm D}) = 2$.

Competition between three factors determines $T_{\rm s}$: absorption of 21-cm CMB photons, absorption and re-emission of Lyman-α photons (the Wouthuysen-Field or WF effect [7·10, 7·11]), and atomic spin-change collisions (collisions with free electrons are unimportant in these environments [7·10]). The first drives $T_{\rm s}$ toward T_γ, while the latter two drive it toward the gas temperature $T_{\rm g}$. In equilibrium the spin temperature of the a species X (either D or H) is

$$T_{\rm s}^{\rm X} = \frac{(1+\chi^{\rm X})T_{\rm g}T_\gamma}{(T_{\rm g}+\chi^{\rm X}T_\gamma)}\,, \tag{7.2}$$

where $\chi^{\rm X} \equiv \chi_c^{\rm X} + \chi_\alpha^{\rm X}$ is the sum of the equilibrium threshold parameters for spin-change collisions and for radiative coupling through the WF effect. Explicitly, $\chi_c^{\rm X} = (C_{+-}^{\rm X}T_\star^{\rm X})/(A_{+-}^{\rm X}T_\gamma)$ and $\chi_\alpha^{\rm X} = (P_{+-}^{\rm X}T_\star^{\rm X})/(A_{+-}^{\rm X}T_\gamma)$, where $C_{+-}^{\rm X}$ is the collisional de-excitation rate, $A_{+-}^{\rm X}$ is an Einstein coefficient, and $P_{+-}^{\rm X} \propto P_\alpha$, where P_α is the total Lyman-α scattering rate. At $z \gg 10$, before the first galaxies formed, P_α is tiny and the WF effect can be neglected. However, it might have interesting consequences near $z \sim 10$.

7.4 H-H and D-H Collision Rates

While the cross section for H-H spin-change collisions $\sigma_{+-}^{\rm HH}$ is well known [7·12–7·16], we were unable to locate the D-H spin-change cross section for the temperature range of interest and computed $\sigma_{+-}^{\rm DH}$ using standard methods[1]. In the elastic approximation,

$$\sigma_{+-}^{\rm DH} = \frac{\pi}{3k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2({}^t\eta_l^{\rm DH} - {}^s\eta_l^{\rm DH})\,, \tag{7.3}$$

where $k = \mu_{\rm DH}v/\hbar$, $\mu_{\rm DH} = m_{\rm D}m_{\rm H}/(m_{\rm D}+m_{\rm H})$ is the reduced mass of the D-H system, and v is the relative velocity.[2] The partial wave phase shifts in the triplet and singlet electronic potentials $V_t(R)$ and $V_{\rm s}(R)$ in which the D and H atoms scatter are ${}^t\eta_l^{\rm DH}$ and ${}^s\eta_l^{\rm DH}$ respectively. We used the

[1] It has been computed at higher temperatures by Ref. [7·18], at 1 K by Ref. [7·19], and measured at 1 K by Ref. [7·20].

[2] Compare with $\sigma_{+-}^{\rm HH} = (\pi/4k^2)\sum_{l=0}^{\infty}(2l+1)\sin^2({}^t\eta_l^{\rm HH} - {}^s\eta_l^{\rm HH})$ for H-H collisions when quantum symmetry can be neglected [7·16].

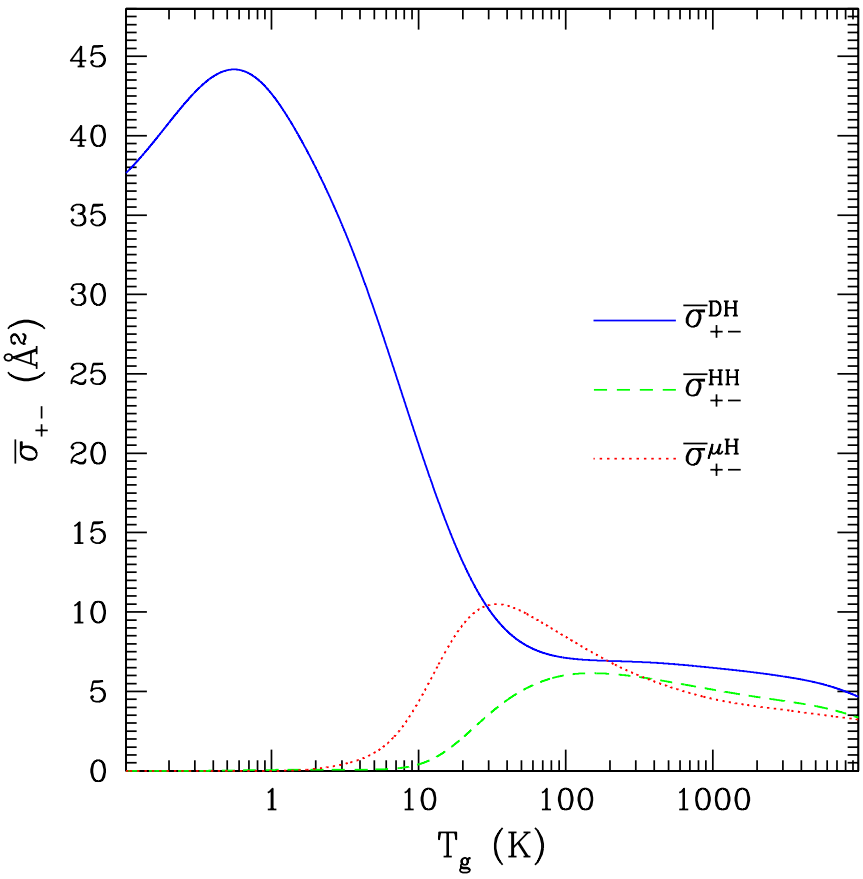

Figure 7.1: The thermal spin-change cross sections that keep T_s collisionally coupled to T_g for D (solid line), H (dashed line), and muonium (dotted line). Although the potentials are identical, the peaks differ because of the reduced masses.

variational potentials of Refs. [7·21, 7·22] for $R \leq 12\,a_0$, the smooth fit of Ref. [7·23] for $V_t(R)$ between $12\,a_0 \leq R \leq 15\,a_0$, and the asymptotic form of Ref. [7·24] for larger R. We found the phase shifts by solving the radial Schrödinger equation at each energy $E = \hbar^2 k^2 / 2\mu_{\rm DH}$ and angular momentum $\sqrt{l(l+1)}\hbar$, truncating at $l_{max} \sim 100$ to resolve the resonant structure at sufficiently high E. We verified that our results agree with Ref. [7·19] at 1 K (after accounting for our more recent potentials), with those of Refs. [7·14–7·17][3] for $\sigma^{\rm HH}_{+-}$, and with Ref. [7·17] for the muonium-H spin-change cross section $\sigma^{\mu\rm H}_{+-}$.

The collisional de-excitation rate is $C^{\rm X}_{+-} = \bar{v}_{\rm XH} \bar{\sigma}^{\rm XH}_{+-} n_{\rm H}$, where $\bar{v}_{\rm XH} = \sqrt{8 k_{\rm B} T_{\rm g} / \mu_{\rm XH} \pi}$ is the thermal velocity, $\bar{\sigma}^{\rm XH}_{+-}$ is the thermal spin-change cross-section (averaged over the Maxwell-Boltzmann distribution of relative velocities), and $n_{\rm H}$ is the number density of H atoms. In Fig. 7.1 we plot $\bar{\sigma}^{\rm HH}_{+-}$, $\bar{\sigma}^{\rm DH}_{+-}$, and $\bar{\sigma}^{\mu\rm H}_{+-}$. While $\bar{\sigma}^{\rm HH}_{+-}$ falls off for $T_{\rm g} \lesssim 100$ K, $\bar{\sigma}^{\rm DH}_{+-}$ continues to rise to a peak near $T_{\rm g} \sim 1$ K. This occurs because of low-energy s-wave and p-wave contributions to D-H scattering (a much larger

[3] We are in harmony with Ref. [7·16] which found that $\kappa(1 \to 0) = \bar{v}_{\rm HH} \bar{\sigma}^{\rm HH}_{+-}$ is 4/3 larger than previously quoted [7·13, 7·14] at high $T_{\rm g}$.

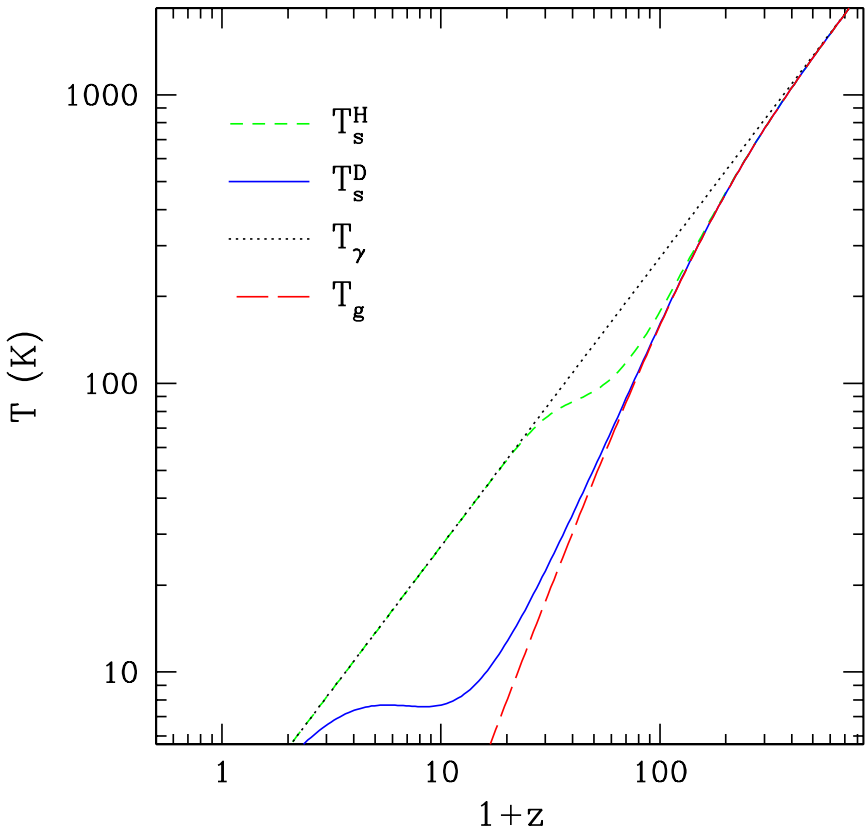

Figure 7.2: The H and D spin temperatures as a function of z. Here we assume $P_\alpha = 0$ for all z.

scattering length and a scattering resonance). These do not appear for H-H because of the different reduced mass ($\mu_{\rm HH} \approx m_{\rm H}/2$ while $\mu_{\rm DH} \approx 2m_{\rm H}/3$). The discussion of D-H spin-change in Ref. [7·10] did not account for this and incorrectly concluded that $\sigma_{+-}^{\rm DH} \sim \sigma_{+-}^{\rm HH}$.

7.5 Spin-Temperature Evolution

In Fig. 7.2 we plot T_γ, T_g (found using `recfast` [7·25]), $T_{\rm s}^{\rm H}$, and $T_{\rm s}^{\rm D}$ as a function of z. After the gas cools below T_γ, collisions keep $T_{\rm s}^{\rm H}$ and $T_{\rm s}^{\rm D}$ coupled to $T_{\rm g}$. Near $z \sim 30$ collisions become inefficient for H and $T_{\rm s}^{\rm H}$ returns to T_γ. $T_{\rm s}^{\rm D}$ remains coupled to $T_{\rm g}$ down to significantly lower redshift both because the lifetime of the excited state of D is relatively long ($A_{+-}^{\rm H}/A_{+-}^{\rm D} = 61.35$) and because $\bar{\sigma}_{+-}^{\rm DH} \gg \bar{\sigma}_{+-}^{\rm HH}$ at low temperatures.

7.6 Brightness Temperature Fluctuations

When the spin temperature of a given species is less than T_γ it will absorb CMB photons. The brightness temperature is $T_{\rm b}^{\rm x} = a\tau_{\rm x}(T_{\rm s}^{\rm x} - T_\gamma)$, where

$$\tau_{\rm x} = \frac{g_+^{\rm x} c\lambda^2 h A_{+-}^{\rm x} n_{\rm x}}{8(g_+^{\rm x} + g_-^{\rm x})\pi k_{\rm B} T_{\rm s}^{\rm x} \mathcal{H}(z)} \tag{7.4}$$

is the optical depth of the spin-flip transition in question, $a = 1/(1+z)$ and $\mathcal{H}(z)$ is the Hubble parameter. We are interested in correlations between brightness temperature fluctuations $\delta_{T_{\rm b}}^{\rm x}(\hat{\mathbf{n}}, a) \equiv \delta T_{\rm b}^{\rm x}(\hat{\mathbf{n}}, a)/T_{\rm b}^{\rm x}(a) = \beta_{T_{\rm b}}^{\rm x}(a)\delta(\hat{\mathbf{n}}, a)$ observed in a direction $\hat{\mathbf{n}}$ at wavelengths differing by a factor $\lambda_{92}/\lambda_{21}$. Here,

$$\beta_{T_{\rm b}}^{\rm x} = 1 + \frac{\chi_{\rm c}^{\rm x}}{\widehat{\chi}^{\rm x}} + \Gamma\left[\frac{T_\gamma}{T_{\rm g} - T_\gamma} + \frac{\chi_{\rm c}^{\rm x}}{\widehat{\chi}^{\rm x}}\frac{d\ln(C_{+-}^{\rm x})}{d\ln(T_{\rm g})}\right], \tag{7.5}$$

$\delta(\hat{\mathbf{n}}, a) = \delta n_{\rm H}(\hat{\mathbf{n}}, a)/n_{\rm H}(a) = \delta n_{\rm D}(\hat{\mathbf{n}}, a)/n_{\rm D}(a)$ is the density contrast, and $\widehat{\chi}^{\rm x} \equiv \chi^{\rm x}(1+\chi^{\rm x})$. At high z, when $T_{\rm g} \approx T_\gamma$, $\Gamma \to 0$ due to residual Thomson scattering with free electrons [7·26] (fluctuations are isothermal), but as the gas begins to cool adiabatically $\Gamma \to 2/3$ [7·27]. In Eq. (7.5) we have neglected the contributions to δ_{T_b} from fluctuations in the neutral fraction (likely to be small at high z) and, for simplicity, fluctuations in the gradient of the radial velocity $\delta_{\partial_r v_r}$ [7·28]. The latter will enhance our signal by a factor of $\sim$ 1–2. In Fig. 7.3 we plot $T_{\rm b}^{\rm H}$, $a\beta_{T_{\rm b}}^{\rm H} T_{\rm b}^{\rm H}$, $\widetilde{T}_{\rm b}^{\rm D}$ and $a\beta_{T_{\rm b}}^{\rm D}\widetilde{T}_{\rm b}^{d}$, where $T_{\rm b}^{\rm D} \equiv \epsilon\widetilde{T}_{\rm b}^{\rm D}$ and $\epsilon \equiv [{\rm D/H}]$. We see that $\widetilde{T}_{\rm b}^{\rm D}$ and $a\beta_{T_{\rm b}}^{\rm D}\widetilde{T}_{\rm b}^{d}$ peak at much lower z than their H counterparts because, as discussed above, $T_{\rm s}^{\rm D}$ is coupled to $T_{\rm g}$ to lower z.

7.7 D-H Cross Correlations

We now estimate the cross-correlation of brightness temperature fluctuations across frequencies related by $\lambda_{21}/\lambda_{92}$. We write the brightness temperature fluctuation due to H or D as $H(\hat{\mathbf{n}}, a) = \beta_{T_{\rm b}}^{\rm H}(a)T_{\rm b}^{\rm H}(a)\delta(\hat{\mathbf{n}}, a)$ and $\epsilon D(\hat{\mathbf{n}}, a) = \epsilon\beta_{T_{\rm b}}^{\rm D}(a)\widetilde{T}_{\rm b}^{\rm D}(a)\delta(\hat{\mathbf{n}}, a)$ respectively. A radio telescope observing at a frequency ν will measure the quantity $\mathcal{O}[\hat{\mathbf{n}}; \nu] = H(\hat{\mathbf{n}}, \nu/\nu_{21}) + \epsilon D(\hat{\mathbf{n}}, \nu/\nu_{92}) + N[\hat{\mathbf{n}}; \nu]$, where $N[\hat{\mathbf{n}}; \nu]$ is the instrumental noise. We form the product $\mathcal{O}[\hat{\mathbf{n}}; \nu_\alpha]\mathcal{O}[\hat{\mathbf{n}}; \nu_\beta]$, where $\nu_\beta \equiv (\nu_{92}/\nu_{21})\nu_\alpha$. Assuming that $\delta(\hat{\mathbf{n}}, a)$ is a zero-mean Gaussian random field and uncorrelated Gaussian noise, its expectation value is $\langle\mathcal{O}[\hat{\mathbf{n}}; \nu_\alpha]\mathcal{O}[\hat{\mathbf{n}}; \nu_\beta]\rangle = \epsilon\langle H_\alpha D_\beta\rangle$ to leading order in ϵ. Here we have introduced the shorthand $H_\alpha \equiv H(\hat{\mathbf{n}}, \nu_\alpha/\nu_{21})$, $D_\beta \equiv D(\hat{\mathbf{n}}, \nu_\beta/\nu_{92})$, and $N_\alpha = N[\hat{\mathbf{n}}; \nu_\alpha]$.

We now understand the crucial point of this chapter. The 21-cm and 92-cm fluctuations at these frequency separations must be correlated because they trace the same underlying patches of the Universe.

Note that we have neglected the relatively small intrinsic correlations of the H brightness fluc-

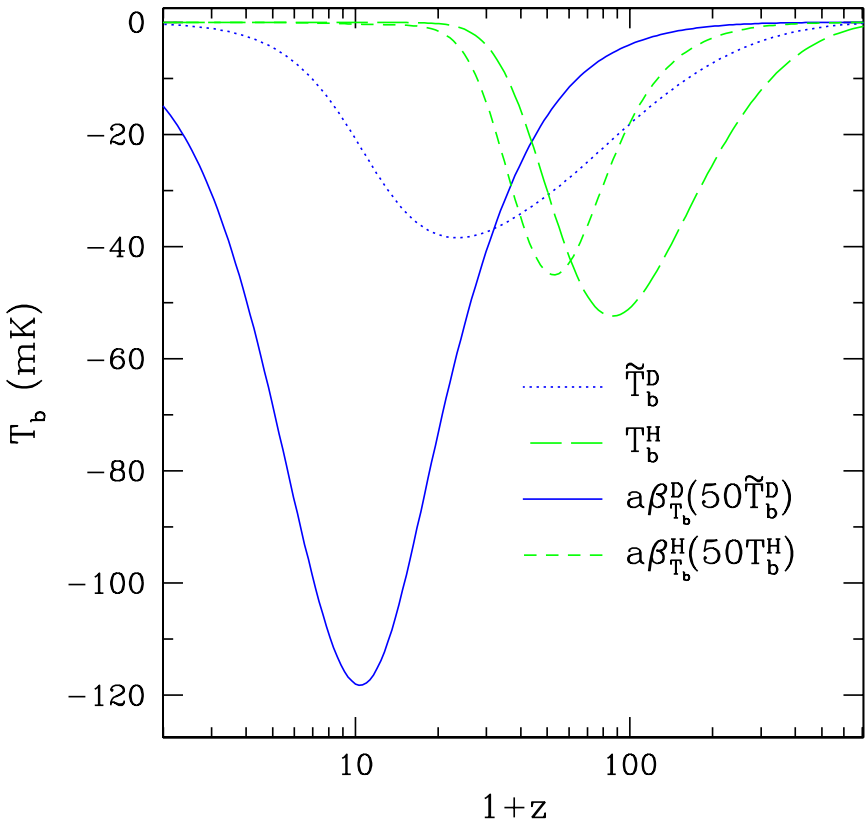

Figure 7.3: The brightness temperatures $T_{\rm b}^{\rm H}$ and $\widetilde{T}_{\rm b}^{\rm D}$ and the brightness temperature fluctuations $\beta_{T_{\rm b}}^{\rm H} T_{\rm b}^{\rm H}$ and $\beta_{T_{\rm b}}^{\rm D} \widetilde{T}_{\rm b}^{\rm D}$ (scaled to the growth rate of density perturbations $\delta \propto a$).

tuations from large-scale modes of the density field which contribute at the level of $\sim 0.1\%$ or less of the D signal. If necessary, these intrinsic H correlations could be removed through independent measurements of the matter power spectrum or by correlating the 21-cm signal with frequencies near but not equal to the corresponding D patches.

7.8 Signal Estimate

The signal-to-noise contributed by a pair of frequency bands centered around (ν_α, ν_β) for an experiment with a maximum baseline of L and frequency resolution $\Delta\nu$ is

$$\frac{\mathcal{S}}{\mathcal{N}}(\nu_\alpha) = \epsilon \frac{4}{\theta_\beta} \frac{\langle H_\alpha D_\beta \rangle}{\sqrt{(\langle H_\alpha^2 \rangle + \langle N_\alpha^2 \rangle)(\langle H_\beta^2 \rangle + \langle N_\beta^2 \rangle}}, \tag{7.6}$$

where $\theta_\beta = \lambda_\beta / L$ is the angular resolution in the D band. Here, $\langle H_\alpha D_\beta \rangle = \sigma_\delta^2 (\beta_{T_{\rm b}}^{\rm H} T_{\rm b}^{\rm H})(\beta_{T_{\rm b}}^{\rm D} \widetilde{T}_{\rm b}^{\rm D})$ and $\langle H_\alpha^2 \rangle = \sigma_\delta^2 (\beta_{T_{\rm b}}^{\rm H} T_{\rm b}^{\rm H})^2$, where

$$\sigma_\delta^2 = \frac{2}{\pi^2} \int_0^\infty dk k P(k) \int_0^k dk_z j_0^2(\xi k_z \rho) \frac{J_1^2(\sqrt{k^2 - k_z^2}\rho)}{(k^2 - k_z^2)\rho^2} \tag{7.7}$$

is the variance in the density field at scale-factor a smoothed over the coin-shaped regions of the Universe of comoving radius ρ and thickness $2\xi\rho$ that are sampled by an experiment with high spectral (radial) resolution but lower angular (transverse) resolution. We adopt a noise variance of $\langle N_\alpha^2 \rangle = T_{\rm sys}^2 / (f_{\rm cov}^2 \, \Delta\nu \, t_{\rm int})$ (e.g., [7·29]), where $T_{\rm sys} = 6500 [\nu_\alpha / (30{\rm MHz})]^{-2}$ K is the noise temperature, $f_{\rm cov}$ is the covering fraction of the array, and $t_{\rm int}$ is the integration time. Our choice for $T_{\rm sys}$ is only an estimate, and the noise (ultimately due to Galactic synchrotron radiation) varies strongly across the sky. The total signal-to-noise ratio $(\mathcal{S}/\mathcal{N})_{\rm tot}$ is the sum in quadrature over all pairs of frequency bands (ν_α, ν_β).

If collisions dominate the coupling between $T_{\rm s}^{\rm H}$ and $T_{\rm g}$ down to $z \sim 7$ then a value [D/H] $\sim 3 \times 10^{-5}$ could be detected at 1- to 2-σ by an experiment with $L \sim 7.5$ km and $\Delta\nu \sim 100$ kHz in $\sim$6 years. If, however, the first generation of stars created a flux of Lyman-α photons which coupled $T_{\rm s}^{\rm H}$ to $T_{\rm g}$ until $z \sim 7$ through the WF effect without significantly heating the gas, a similar detection might be made by a smaller experiment with $L \sim 2.5$ km and $\Delta\nu \sim 1$ kHz. Although it dramatically enhances the 21-cm signal, the WF effect does not improve the [D/H] measurement by the same margin because $\mathcal{S}/\mathcal{N}$ becomes independent of H_α once $\langle H_\alpha^2 \rangle \gg \langle N_\alpha^2 \rangle$ (it only serves to make the H fluctuations a better matched template). Finally, we note that an experiment capable of mapping 21-cm brightness-temperature fluctuations out to $l_{max} \sim 10^5$ (where it may be a powerful probe of the small-scale matter power spectrum [7·4]) could measure [D/H] to a precision as good as $\sim 1\%$ — or even $\sim 0.1\%$ if the WF effect coupling is efficient.

For these estimates we have assumed a ΛCDM cosmology with $n_{\rm s} = 1$, and that a significant fraction of the Universe remains neutral until $z \sim 7$. The largest contribution to the signal originates from $z \lesssim 10$ where the D signal peaks and the variance in the density fluctuations is largest. Varying these assumptions, or including additional sources of temperature fluctuations, could change $(\mathcal{S}/\mathcal{N})_{\rm tot}$ by factors of order unity.

7.9 Discussion

Despite the obvious technical challenges in observing this signal, we emphasize that it has the virtue of providing the cleanest possible measurement of the primordial [D/H], free from contamination by structure formation processes at lower z. Via the window of BBN, this would allow radio telescopes to peer into the first few minutes of the Universe. We believe future searches for cosmic 21-cm

fluctuations should bear this possibility in mind.

We also note that $^3He^+$ has a hyperfine transition (with $\lambda = 3.46$ cm) that can be used in a similar fashion; it has the advantage of much lower foreground contamination at higher frequencies. This line will appear during reionization and should exhibit a strong anti-correlation with the corresponding 21-cm signal. If the astrophysics of reionization can be understood well enough, the cross-correlation of this line with the 21-cm line could supplement the D-H experiment in order to probe BBN in even more detail.

Acknowledgments

We thank A. Dalgarno, J. Gair, M. Hayden, M. Kamionkowski, T. Pearson, M. Reynolds, D. Scott and B. Zygelman for discussions. KS acknowledges the support of a Canadian NSERC Postgraduate Scholarship. This work was supported in part by NASA NAG5-9821 and DoE DE-FG03-92-ER40701.

Bibliography

[7·1] D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148**, 175 (2003).

[7·2] E. Fermi, Zeits. f. Physik **60**, 320 (1930).

[7·3] C. J. Hogan and M. J. Rees, Mon. Not. Roy. Astron. Soc. **188**, 791 (1979); D. Scott and M. J. Rees, Mon. Not. Roy. Astron. Soc. **247**, 510 (1990).

[7·4] A. Loeb and M. Zaldarriaga, Phys. Rev. Lett. **92**, 211301 (2004).

[7·5] J. E. Nafe, E. B. Nelson, and I. I. Rabi, Phys. Rev. **71** 914 (1947).

[7·6] S. Weinreb, Nature **195** 367 (1962); J. M. Pasachoff and D. A. Cesarsky, Astrophys. J. **193**, 65 (1974); K. R. Anantharamaiah and V. Radhakrishnan, Astron. Astrophys. **79**, L9 (1979); L. Blitz and C. Heiles, Astrophys. J. Lett. **313**, L95 (1987); J. N. Chengalur, R. Braun and W. B. Burton, Astron. Astrophys. **318**, L35 (1997).

[7·7] R. V. Wagoner, W. A. Fowler and F. Hoyle, Astrophys. J. **148**, 3 (1967).

[7·8] H. Reeves, J. Audouze, W. A. Fowler, and D. N. Schramm, Astrophys. J. **179**, 909 (1973).

[7·9] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics, Vol. 2.* (Wiley, New York, 1977).

[7·10] G. B. Field, Proc. IRE **46**, 240 (1958).

[7·11] S. A. Wouthuysen, Astron. J. **57**, 31 (1952).

[7·12] A. Dalgarno, Proc. R. Soc. London Ser. A **262**, 132 (1961).

[7·13] F. J. Smith, Planet. Space Sci. **14**, 929 (1966).

[7·14] A. C. Allison, and A. Dalgarno, Astrophys. J. **158**, 423 (1969).

[7·15] A. C. Allison, Phys. Rev. A **5**, 2695 (1972).

[7·16] B. Zygelman, Astrophys. J. **622**, 1356 (2005).

[7·17] B. Shizgal, J. Phys. B: At. Mol. Phys. **12**, 3611 (1979).

[7·18] P. S. Krstic and D. R. Schultz, Phys. Rev. A **60**, 2118 (1999).

[7·19] M. W. Reynolds, Ph.D. thesis, University of British Columbia, 1989.

[7·20] M. E. Hayden and W. N. Hardy, J. Low. Temp. Phys. **99**, 787 (1995).

[7·21] W. Kolos, K. Szalewicz, and H. J. Monkhorst, J. Chem. Phys. **84**, 3278 (1986).

[7·22] D. Frye, G. C. Lie, and E. Clementi, J. Chem. Phys. **91**, 2366 (1989).

[7·23] M. J. Jamieson, A. Dalgarno, and J. N. Yukich, Phys. Rev. A **46**, 6956 (1992).

[7·24] W. Kolos, Int. J. Quantum Chem. **1**, 169 (1967).

[7·25] S. Seager, D. D. Sasselov and D. Scott, Astrophys. J. **523**, L1 (1999); S. Seager, D. D. Sasselov and D. Scott, Astrophys. J. Suppl. **128**, 407 (2000).

[7·26] P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton University Press, Princeton, 1993).

[7·27] S. Bharadwaj and S. S. Ali, Mon. Not. Roy. Astron. Soc. **352**, 142 (2004).

[7·28] R. Barkana and A. Loeb, arXiv:astro-ph/0409572.

[7·29] M. Zaldarriaga, S. R. Furlanetto and L. Hernquist, Astrophys. J. **608**, 622 (2004).