Polarization in the Sunyaev-Zeldovich Effect

by

Jamie Portsmouth

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2003

Author .....
Department of Physics
August 28, 2003

Certified by .....
Edmund Bertschinger
Professor
Thesis Supervisor

Certified by .....
Alan Guth
Professor
Thesis Supervisor

Accepted by .....
Thomas J. Greytak
Associate Department Head for Education

Polarization in the Sunyaev-Zeldovich Effect

by

Jamie Portsmouth

Submitted to the Department of Physics
on August 28, 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

A novel covariant formalism for the treatment of the transfer and Compton scattering of partially polarized light is presented. In this approach, the polarization state of a light beam is described by a tensor constructed from the time average of quadratic products of the electric field components in a local observer frame. This leads naturally to a covariant description which is ideal for calculations involving the boosting of polarized light beams between Lorentz frames, and is more flexible than the traditional Stokes parameter approach in which a separate set of polarization basis vectors is required for each photon.

The covariant kinetic equation for Compton scattering of partially polarized light by relativistic electrons is obtained in the tensor formalism by a heuristic semi-classical line of reasoning. The kinetic equation is derived first in the electron rest frame in the Thomson limit, and then is generalized to account for electron recoil and allow for scattering from an arbitrary distribution of electrons.

This formalism is applied to a calculation of the relativistic corrections to the spectral distortions imprinted in the intensity and polarization of the cosmic microwave background radiation (CMB) by inverse Compton scattering in clusters of galaxies (the Sunyaev-Zeldovich effects). We develop a Monte Carlo method for simulating these effects, based on the tensor formalism and kinetic equation.

We also consider the use of the polarization signal generated by scattering of the CMB from distant clusters as a probe of cosmological perturbations. Such observations allow an indirect measure of the CMB quadrupole as seen on the last scattering surfaces of observers at nonzero redshift. The statistical properties of this signal in a simple cosmological model are derived. We demonstrate that measurements of this signal would yield more information than is available from observations of the CMB anisotropies on our sky, and would potentially allow more precise measurement of cosmological parameters and the primordial power spectrum of density fluctuations.

Thesis Supervisor: Edmund Bertschinger
Title: Professor

Thesis Supervisor: Alan Guth
Title: Professor

Acknowledgments

I must first thank my research advisers Ed Bertschinger and Alan Guth. After a few years at MIT, I found myself involved in some uninspiring numerical calculations in magnetohydrodynamics. About this time, mostly due to Ed's lectures on relativity and Paul Schechter's on extragalactic astronomy, I realized that something not far short of a golden age in cosmology was underway, and I wanted to find some way to be a part of it. Ed and Alan were kind enough to agree to take me on at a relatively late stage in my career at MIT, and found for me a rather original project which has followed a somewhat tortuous (or torturous) path but ultimately been quite successful. Ed has really gone beyond the line of duty in mentoring me academically and taking a hands on approach to my project, and I thank him sincerely for it. Alan has injected at every stage an extra dose of clear thinking and rigor, which has substantially improved this thesis and given me a standard to aspire to in future work.

At MIT, I have been privileged to interact with some other great minds. In particular, it has been a pleasure and inspiration to TA for Jeffrey Goldstone and chat often about E&M problems (trivial for him) at his blackboard. I've enjoyed the lectures of Mehran Kardar on statistical field theory – I have him to thank for many a happy hour spent in my second year doing ϵ–expansions down in the Science library. Barton Zwiebach's undergraduate course on string theory was cool, and I thank him for making it an official breadth requirement so I could graduate! Thanks also to Scott Burles for reading the thesis and making good comments. Outside of MIT, I thank Greg Bryan for very kindly agreeing to provide me with data from some of his simulations and making considerable effort to explain how to read the files.

There are a few friends who've helped me along the way that I'd like to express my gratitude to. First, to Phil Singer, whose public school wit kept me in touch with my Britishness. To Gang Chen, for easing my transition into MIT as a fellow international student. To my former 8-218M "computer graveyard" buddies, Giovanni Covone, Josh Faber, John Fregeau, Nick Morgan, and Alex Shirokov, for their comradeship. Special thanks to Nick for keeping me laughing during study for the astrophysics oral exam, and to John for keeping my bench press up. I thank Keng-Hui Lin for her companionship and sympathy in my last year at MIT. Mike Muno deserves thanks for suffering through a year as my room-mate. Mike Malm has sought to civilize me throughout my time at MIT and he almost succeeded. I've enjoyed discussions with plasma physicist Mark Hess about his ideas for hard E&M problems and keeping me informed about the latest greatest Greens functions. Last but not least, I thank fellow Ashdown geek Keith Duggar for keeping me on my toes with countless thorny games of chess and heated debates over evolutionary theory, politics or whatever.

Finally, my love and gratitude to my mum, dad, and Anna, my sister, for their unconditional support, without which I would not have got as far as this.

Contents

Chapter 1

Introduction

In the last few decades, there has been increasing interest in measurements of the anisotropies in the cosmic microwave background (CMB) radiation, the relic radiation of the big bang, which is an almost perfectly isotropic blackbody radiation field at a temperature of 2.726 K. Primary fluctuations in the CMB temperature and polarization were imprinted at the last scattering surface at redshift $z \approx 1100$. Theoretical predictions of the observable fluctuations in the temperature and polarization maps of the CMB, taking into account all of the details of the dynamics of the growth of matter and radiation perturbations, have become easy with codes such as CMBFAST (Ma & Bertschinger, 1995; Seljak & Zaldarriaga, 1996; Zaldarriaga & Seljak, 1997). The agreement between these theoretical models and the observations of the temperature anisotropy from instruments such as COBE (Lubin et al., 1983) and WMAP (Bennett et al., 2003; Kogut et al., 2003) has so far been quite spectacular.

Largely due to these observations and the underlying theory (in conjunction with data from galaxy surveys, weak lensing, and the Ly-α forest), it is now widely accepted that the universe is very nearly flat, as predicted by the inflationary paradigm (Guth, 1981), and has energy density dominated by a cosmological constant. Measurements of the polarization of the CMB will yield additional information about the physics at the era of recombination, and are more sensitive than the temperature measurements to the stochastic background of gravitational waves predicted in inflationary models.

With some confidence that the background cosmology has been reliably determined, attention has turned to the study of CMB foregrounds (Peterson et al., 1999), which are usually defined as any distortion to the CMB additional to that imprinted at the era of recombination. The most interesting foregrounds from a cosmological point of view are those due to Thomson scattering of the CMB photons by the electrons in ionized plasma. At a redshift $z = 10 - 100$, the first stars and galaxies formed and re-ionized the surrounding neutral gas. This process leaves a signature in the polarization of the CMB at large and small angular scales.

At low redshift, inverse Compton scattering by the dense, hot concentrations of ionized gas in clusters of galaxies injects energy into the CMB, leading to a net increase in the (frequency integrated) CMB brightness in the direction of clusters. This is the thermal *Sunyaev-Zeldovich effect* (SZE). (Zeldovich & Sunyaev, 1969; Sunyaev & Zeldovich, 1980a). The effect leads to a deviation of the CMB from a perfect

Planck spectrum with a well known frequency dependence. At low frequencies, in the Rayleigh-Jeans limit, the CMB is actually cooler, while it is hotter in the Wien tail. In conjunction with measurements of the X-ray brightness of clusters, the thermal SZE allows measurement of the Hubble constant, although the measurement is difficult (Birkinshaw, 1999).

The bulk motion of the gas in a cluster with respect to the CMB rest frame leads to a Doppler boosting or dimming of the CMB according to whether the cluster is approaching of receding along the line of sight. This *kinematic Sunyaev-Zeldovich effect.* has a different frequency dependence from the thermal SZE (Sunyaev & Zeldovich, 1980b). Measurements of this effect (which have not yet been achieved) would allow determination of the radial bulk motions of gas in the cluster. This in turn would allow an estimate of the underlying dark matter peculiar velocity field and provide further constraints on cosmological parameters and models of structure formation. Several groups have used hydrodynamic simulations to investigate the statistical properties of the SZ contamination of the CMB power spectra, and considered how to extract information on peculiar velocities from the kinematic effect (Molnar & Birkinshaw, 2000; Seljak et al., 2001; Refregier & Teyssier, 2002; Nagai et al., 2003).

In addition to these effects which distort the CMB intensity, a CMB polarization signal can be generated via Compton scattering - we refer to this as the generation of *polarization in the Sunyaev-Zeldovich effect.* The basic process responsible for the generation of polarization is Thomson scattering of a radiation field with a quadrupole anisotropy (Sunyaev & Zeldovich, 1980b; Audit & Simmons, 1999; Sazonov & Sunyaev, 1999). There are several means by which this anisotropy may be generated, in the case of the CMB radiation incident on a galaxy cluster: the primary CMB temperature quadrupole Q_2 at the cluster (yielding polarization fraction $\sim \tau_{\rm T} Q_2$, where $\tau_{\rm T}$ is the optical depth to Thomson scattering), the kinematic quadrupole arising from the Doppler boost of the isotropic CMB into the electron rest frame ($\sim \tau_{\rm T} v^2/c^2$), and double scattering of the anisotropic radiation due to the single scattering thermal and kinematic effects ($\sim \tau_{\rm T}$ times the thermal and kinematic SZE). These effects are described in detail in Chapters 4 and 5.

For rich clusters with $\tau_{\rm T} \sim 0.01$, these polarization effects are expected to produce a 0.1μK distortion at most, and so their measurement will be challenging. However there are compelling reasons for studying them both observationally and theoretically. First of all, measurements of the kinematically generated polarization effect would allow determination of the components of the cluster bulk velocity perpendicular to the line of sight, complementing the measurement of the radial component via the unpolarized kinematic effect. Secondly, the polarization signal produced by scattering of the intrinsic CMB quadrupole anisotropy allows an *in situ* measurement of the local CMB quadrupole. As described in §1.2, this provides a means of studying directly the time evolution of the CMB anisotropies. And finally, theoretical study of the SZ polarization effects has motivated the introduction of an original formalism for dealing with the Compton scattering of polarized photons, as described in §1.1, which will no doubt have applications in other astrophysical settings.

1.1 A tensor method for polarization calculations

The computation of the polarization in the Sunyaev-Zeldovich effect requires a formalism for the treatment of the Compton scattering of a polarized radiation field. In looking at the details of these calculations it becomes apparent that the Stokes parameter formalism conventionally used in polarized radiative transfer (Chandrasekhar, 1960), and in the primary CMB calculations (Zaldarriaga & Seljak, 1997), is very cumbersome for this purpose, due to the fact that a separate set of polarization basis vectors has to be specified for every photon. Since Compton scattering involves a relativistic scattering electron in general, Lorentz transformation of the Stokes parameters is necessary, which turns out to be complicated.

To get around this difficulty, we found it convenient and illuminating to introduce a novel formalism for doing radiative transfer calculations with polarized photons, the *polarization tensor formalism*. The formalism is described in detail in Chapter 2, and its application to Compton scattering described in Chapter 3. There exist other approaches to the description of the polarization properties of radiation fields, for example the Jones calculus, Mueller matrices, and coherency matrices (see Swindell (1975)). Our approach is closest in spirit to the coherency matrices of Wolf (1959), except we go further and develop a covariant tensor formalism. With this we can study the polarization generated by relativistic Compton scattering.

The basic idea of this matrix formalism is to associate a 3×3 matrix with each photon, rather than a set of polarization basis vectors and the associated Stokes parameters. For example, a beam of partially-polarized light travelling in the z-direction is described by the Hermitian matrix (termed the *polarization matrix*):

$$\boldsymbol{I} = [I_{ij}] = \frac{1}{2}\begin{pmatrix} I+Q & U-iV & 0 \\ U+iV & I-Q & 0 \\ 0 & 0 & 0 \end{pmatrix} , \tag{1.1}$$

where I, Q, U, V are the Stokes parameters (with units of specific intensity). The trace of the matrix is the total beam intensity. For a general photon direction $\boldsymbol{n}$, the beam is described by a matrix $\boldsymbol{I}(\boldsymbol{n})$, and transversality of the polarization implies $n^i I_{ij} = 0$. In general, the polarization matrix is a function of photon frequency and direction as well as spatial position and time, $\boldsymbol{I} = \boldsymbol{I}(\nu, \boldsymbol{n}, \boldsymbol{x}, t)$. The real advantage of this description is that there is no need to perform a complicated rotation of axes when examining photons with different direction vectors (in the angular integrations needed in the radiative transfer equation for example). In addition, it is simple to extend the 3×3 matrix description to a 4×4 tensor description in which Lorentz transformation of polarized beams between frames is easy.

In the matrix approach, the radiative transfer equation for scattering of polarized radiation is much more straightforward than in the Stokes approach. There is no need for rotation of axes to define separate Stokes parameters for the incoming and outgoing beams. Both are described by a single polarization matrix. By contrast, when using Stokes parameters one has a complicated angular integral involving rotation matrices (Hansen & Lilje, 1999; Chandrasekhar, 1960). The transfer equation for Thomson

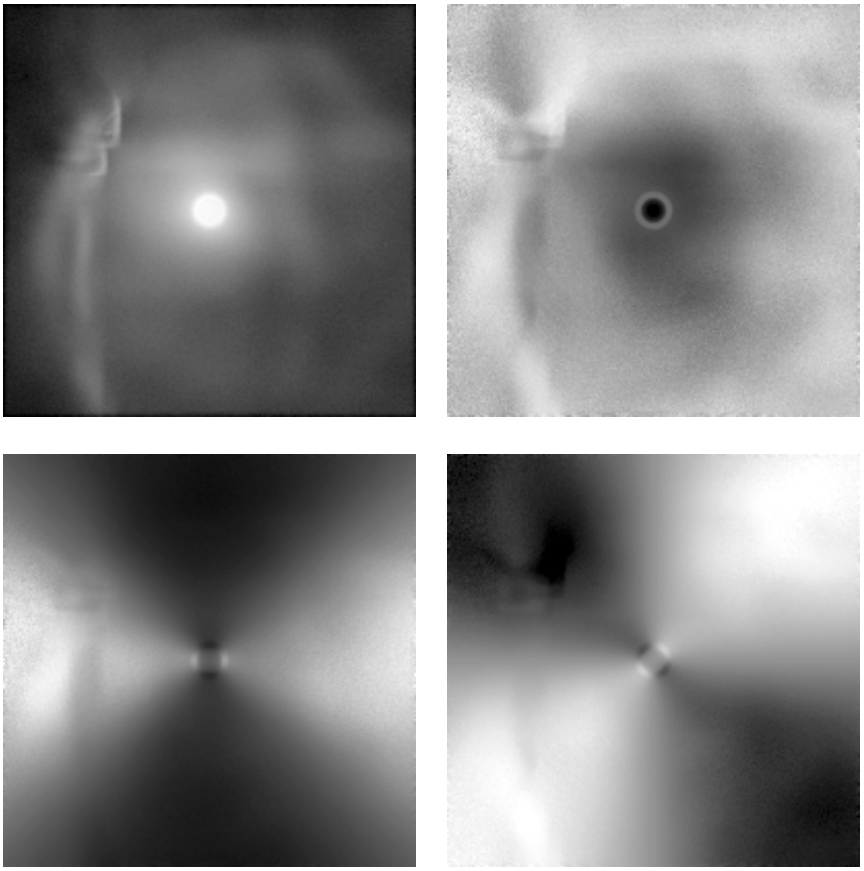


Figure 1-1: Radiation Thomson scattered into the line of sight from a source at the center of an inhomogeneous cloud of electrons. In the various panels we show: the total intensity I (upper left), the magnitude of the polarization Π (upper right), the Stokes Q parameter (lower left), the Stokes U parameter (lower right). The Stokes parameters are defined with respect to the axes of the projected face.

scattering is elegantly expressed in terms of a set of projection matrices $\boldsymbol{P}(\boldsymbol{n}_s)$ which project out of the matrix $\boldsymbol{I}(\boldsymbol{n})$ the component of polarization orthogonal to $\boldsymbol{n}_s$.

The Monte Carlo simulation of polarized radiative transfer is also easy with polarization matrices. The result of such a simulation is shown in Fig. 1-1. These are simulated observations of the intensity (log scale) and polarization (linear scale) of the Thomson scattered (monochromatic) radiation produced by a point source at the center of a cubic volume containing an inhomogeneous distribution of stationary electrons (as observed from a point in the far field). The plasma density distribution, defined on a cubic grid of 32^3 points, is taken from a cosmological simulation of a galaxy cluster by Greg Bryan. The cloudy and filamentary nature of the gas is apparent, as is the discreteness of the grid on which it is defined. In this simulation 4×10^7 unpolarized photons were generated at the source and allowed to Thomson scatter through the cloud until they hit the boundary. The opacity was adjusted so that average number of scatterings undergone by each photon before leaving the cube was 2.0. The procedure used to generate these images is described in Appendix B. To

perform such a simulation of the generation of polarization with multiple scattering would be much more difficult in the Stokes approach than with the matrix approach advocated here.

1.2 Spatial correlations of CMB anisotropies from polarization

Kamionkowski & Loeb (1997) made the interesting suggestion that measurements of the polarized SZE could enable one to get around the cosmic variance limit on the CMB quadrupole ($l = 2$) anisotropy. Part of the polarization measured from a cluster is proportional to the CMB quadrupole at the cluster itself. Since an observer in a distant cluster has a different last-scattering surface than an observer at redshift $z = 0$, this leads us to ask, what additional information is present in principle if one could measure the CMB quadrupole at other points in spacetime? This is explored in detail in Chapter 5. This has required generalizing the usual CMB treatment, which is based on the assumption that we can measure the anisotropy at only one point in space. Expanding the anisotropy in the usual way,

$$\Delta(\boldsymbol{x}, \boldsymbol{n}, \tau) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm}(\boldsymbol{x}, \tau) Y_{lm}(\boldsymbol{n}) \ , \tag{1.2}$$

where $\boldsymbol{x}$ is the comoving position vector, τ is the conformal time, and $\boldsymbol{n}$ is the photon direction vector. One must define a direction for the polar axis and prime meridian everywhere in space; this is easily done with respect to the comoving coordinate grid. Now one can generalize the usual CMB angular power spectrum to distinct points in space as well as multipoles, resulting in the generalized CMB correlation functions:

$$C_{lml'm'}(\boldsymbol{x}, \tau, \boldsymbol{x}', \tau') \equiv \langle a_{lm}(\boldsymbol{x}, \tau) a^*_{l'm'}(\boldsymbol{x}', \tau') \rangle \ . \tag{1.3}$$

When the two points are brought together, $C_{lml'm'} = C_l(\tau)\delta_{ll'}\delta_{mm'}$ is diagonal. However, it is non-diagonal when the points are distinct. The cluster polarization allows us to extract some of the information contained in the generalized correlation functions that is not present in the simple angular power spectrum C_l measured at redshift $z = 0$. In Chapter 5, we show that by combining polarization measurements from many clusters at different redshifts, the cluster polarization signal can be used as a probe of the time evolution of the angular power spectrum harmonic $C_2(\tau)$. In the now-standard ΛCDM cosmology, a late-time contribution to the CMB anisotropy on large scales is made by the time-changing gravitational potential as Ω_m drops significantly below 1 starting about $z \approx 1$. This effect, known as the integrated Sachs-Wolfe effect, could be directly measured by using clusters to provide *in situ* measurements of the quadrupole. The experimental measurement is difficult because it is based on polarization anisotropy in the SZE. However, it is worthwhile to develop any possible way to provide an independent check on the acceleration of the universe, starting at $z = 2$, inferred from type Ia supernovae.

Chapter 2

Tensor formalism for transfer of polarized light

There is an extensive literature dealing with the radiative transfer of polarized light (Wolf, 1959; Chandrasekhar, 1960; Dialetis, 1969; Acquista & Anderson, 1974; Henney, 1994; Lee et al., 1994; Code & Whitney, 1995; Hansen & Lilje, 1999; Challinor, 2000; Carozzi et al., 2000). Most treatments use the four Stokes parameters I, Q, U, V, which provide a complete description of the polarization properties of the radiation field. These parameters have dimensions of specific intensity, and are functions of time, photon propagation direction, and frequency. In the case of unpolarized photons, a complete description of the radiation field is given by the total specific intensity Stokes parameter I_ν, or equivalently the phase space density of photons. The Stokes parameters are essentially time averages of quadratic products of the electric field components of the electromagnetic field. They are usually not written in a form which is manifestly covariant, but this can be done quite simply by expressing the electric fields in terms of the Maxwell field strength tensor components in a local observer frame. This leads naturally to a classical formalism in which the four Stokes parameters are replaced by a two index complex Hermitian tensor $I^{\mu\nu}$ whose trace reduces in the unpolarized case to the usual total intensity. A tensor analogue of the phase space distribution function, $f^{\mu\nu}$, is also easily defined. These objects are collectively termed the *polarization tensor*. This formalism is similar to that introduced by Challinor (2000).

In §2.1, the Stokes formalism is reviewed and we reconsider the notion of the polarization coherency matrix. In §2.2, this notion is generalized and our tensor description of polarized light described, first in a non-covariant manner. The covariant formalism is introduced in §2.3. The properties of the polarization tensors, their evolution in the absence of scattering and in the geometrical optics limit, and their relation to the Stokes description are discussed. In §2.4, the behaviour of the polarization tensor under Lorentz transformation is discussed, and an explicit example of the computation of the polarization of a boosted beam presented.

Note that throughout this thesis, boldface quantities, e.g. $\boldsymbol{p}$, denote 3-vectors, and quantities with vector arrows, e.g. $\vec{p}$, denote 4-vectors. The indices of 3-vectors and tensors are denoted with Roman indices, and those of 4-vectors and tensors with

Greek indices. Both 3×3 and 4×4 matrices are denoted with boldface quantities. Note also that $\Re$ denotes the set of real numbers, $\mathcal{C}$ denotes the set of complex numbers, and $\Re\mathrm{e}$ denotes the operation of taking the real part.

2.1 The coherency matrix

The classical description of partially polarized light uses the well known Stokes parameters, which are defined operationally in terms of experiments with polarizing plates. Physically the Stokes parameters can be thought of as time averages of instantaneous products of electric field components. There is a close relationship between the Stokes parameters and the notion of the coherence of the two photon polarization states, which is described mathematically by the *coherency matrix* introduced by Wolf (1959), based on the work of Wiener (1930). Additional work was done by Barakat (1963) to extend the concept to a spectral coherency matrix. It is worthwhile reviewing the notion of the coherency matrix, since this leads naturally to the polarization tensor description.

We will only consider electromagnetic fields which are superpositions of plane electromagnetic waves. An idealized superposition of such waves whose wave-vectors are all perfectly aligned will be termed a *beam*. Consider first a beam propagating along the z-axis. The transverse electric field components at a specified fixed spatial point (x, y, z) are real functions of time, $E_x^{(r)}(t)$, $E_y^{(r)}(t)$. These functions can be expressed as a superposition of an infinite number of monochromatic waves with arbitrary phases, i.e. as Fourier transforms

$$E_j^{(r)}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{E}_j(\omega) e^{-i\omega t} d\omega \ . \qquad E_j^{(r)} \in \Re, \quad j \in \{x, y\} \tag{2.1}$$

We have assumed here of course that the Fourier transform exists – which is not true for all functions $E_x^{(r)}(t)$, $E_y^{(r)}(t)$, but we will gloss over this point (the existence of the Fourier transform can be assured without difficulty by working with functions which are truncated as $t \to \pm\infty$. See for example Born & Wolf (1980)). In order to ensure reality of $E_j^{(r)}(t)$, the Fourier transforms must satisfy $\tilde{E}_j(-\omega) = \tilde{E}_j^*(\omega)$. Now we split the integral above into two parts:

$$\begin{aligned} E_j^{(r)}(t) &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \tilde{E}_j(\omega) e^{-i\omega t} d\omega + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \tilde{E}_j(\omega) e^{-i\omega t} d\omega \\ &= \frac{1}{2} \left(E_j(t) + E_j^*(t) \right) = \Re\mathrm{e}\ E_j(t) \ , \end{aligned} \tag{2.2}$$

where we have defined the complex functions $E_j(t)$, conventionally called the *analytic signal* (Born & Wolf, 1980) associated with $E_j^{(r)}(t)$:

$$E_j(t) = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \tilde{E}_j(\omega) e^{-i\omega t} d\omega \ . \qquad E_j(t) \in \mathcal{C} \tag{2.3}$$

We may decompose $\tilde{E}_j(\omega)$ uniquely into a real amplitude and complex phase factor:

$$\tilde{E}_j(\omega) = a_j(\omega)e^{i\phi_j(\omega)} \ . \qquad a_j, \phi_j \in \Re \tag{2.4}$$

The analytic signal is thus

$$E_j(t) = \frac{2}{\sqrt{2\pi}} \int_0^\infty a_j(\omega)e^{i\phi_j(\omega)-i\omega t}d\omega \ . \tag{2.5}$$

How are all of these quantities related to what is measured by a real polarimeter? Generally speaking, polarimeters measure the time average of the intensity of the light beam at a fixed spatial point after it has traveled through a combination of filters (See e.g. Britton (2000) for a good general discussion of astronomical polarimetry). The two basic filter elements required to measure the polarization state are a *polarizing plate*, and a *compensator* (Stone, 1963). We shall describe how the time average is constructed from the quantities we have defined, and then consider the effect of the two types of filter on the beam.

We first make the simplifying assumption that the beam is *quasi-monochromatic*, which means that the functions $\tilde{E}_j(\omega)$ are assumed to be non-vanishing only in a narrow frequency band $\omega \in [\omega_0 - \Delta\omega/2, \omega_0 + \Delta\omega/2]$, with $\Delta\omega \ll \omega_0$. Physically this means that the beam is a wave-packet of spectral width $\Delta\omega$, centered roughly on frequency ω_0. This implies that the functions $a_j(t), \phi_j(t)$ vary slowly in comparison to $\cos(\omega_0 t)$. To see this, first note that we can always choose to write the analytic signals in the form

$$E_j(t) = a_j(t)e^{i[\phi_j(t)-\omega_0 t]} \ . \tag{2.6}$$

Then it follows from Eqn. (2.5) that

$$\begin{aligned} a_j(t)e^{i\phi_j(t)} &= \frac{2}{\sqrt{2\pi}} \int_0^\infty \tilde{E}_i(\omega)e^{-i(\omega-\omega_0)t}d\omega \\ &= \frac{2}{\sqrt{2\pi}} \int_{-\omega_0}^\infty \tilde{E}_i(\omega'+\omega_0)e^{-i\omega' t}d\omega' \ , \end{aligned} \tag{2.7}$$

Then since $\tilde{E}_i(\omega'+\omega_0)$ vanishes by assumption for $|\omega'| > \Delta\omega/2$, the left hand side is a superposition of Fourier modes of low frequency $|\omega'| < \Delta\omega/2 \ll \omega_0$.

Then the time average is defined by

$$\left\langle E_j^{(r)}(t) \right\rangle \equiv \frac{1}{2T} \int_{t-T}^{t+T} E_j^{(r)}(t)dt \ , \tag{2.8}$$

where T is chosen such that $\Delta\omega \ll \frac{2\pi}{T} \ll \omega_0$. The quantities measured by the detector will be some combination of the following time averaged real quantities (expanding using Eqn. (2.2)):

$$\left\langle E_i^{(r)}(t)E_j^{(r)}(t) \right\rangle = \frac{1}{4}\left[\langle E_i(t)E_j(t)\rangle + \langle E_i(t)E_j^*(t)\rangle + \langle E_i^*(t)E_j(t)\rangle + \langle E_i^*(t)E_j^*(t)\rangle\right] \ . \tag{2.9}$$

With the assumption of quasi-monochromaticity we may now ignore time averages which contain the rapidly varying phase factor $e^{i\omega_0 t}$ and retain only those over the slowly varying functions $a_j(t)$, $e^{i\phi_j(t)}$. Thus, for example

$$\begin{aligned}
\langle E_x(t)E_x(t)\rangle &= \left\langle a_x^2(t)e^{2i\phi_x(t)}e^{-2i\omega_0 t}\right\rangle \\
&= 0 \\
\langle E_x(t)E_x^*(t)\rangle &= \left\langle a_x^2(t)\right\rangle \\
\langle E_x(t)E_y(t)\rangle &= 0 \\
\left\langle E_x(t)E_y^*(t)\right\rangle &= \left\langle a_x(t)a_y(t)e^{i(\phi_x(t)-\phi_y(t))}\right\rangle \ . \qquad (2.10)
\end{aligned}$$

The non-vanishing elements are all of the form $J_{ij} = \langle E_i(t)E_j^*(t)\rangle$. We denote the Hermitian matrix of quantities J_{ij} the *coherency matrix*:

$$\boldsymbol{J} = \left[\begin{array}{cc} \langle a_x^2(t)\rangle & \langle a_x(t)a_y(t)e^{i(\phi_x(t)-\phi_y(t))}\rangle \\ \langle a_x(t)a_y(t)e^{i(\phi_y(t)-\phi_x(t))}\rangle & \langle a_y^2(t)\rangle \end{array} \right] \ . \qquad (2.11)$$

This matrix was introduced by Wolf (1959). Now we relate the elements of the coherency matrix to measurements with a polarimeter. With an optical element known as a compensator, a coherent phase delay between the x and y components of the beam can be introduced. After passing through this device, the resulting analytic signal has the form

$$\bar{E}_j(t) = a_j(t)e^{i[\phi_j(t)-\epsilon_j-\omega_0 t]} \ , \qquad (2.12)$$

where the phase difference $\delta \equiv \epsilon_x - \epsilon_y$ is a known constant. Taking time averages of products of these quantities yields

$$\begin{aligned}
\langle \bar{E}_i(t)\bar{E}_j^*(t)\rangle &= e^{-i(\epsilon_i-\epsilon_j)}\langle a_i(t)a_j(t)e^{i(\phi_i(t)-\phi_j(t))}\rangle \\
&= e^{-i(\epsilon_i-\epsilon_j)}J_{ij} \ . \qquad (2.13)
\end{aligned}$$

The polarization is measured by passing the beam through a further optical element, a polarizing plate oriented at angle θ to the x–direction, and measuring the total intensity of the transmitted light, $I(\theta)$. The transmitted electric field is

$$\bar{E}^{(r)}(\theta,t) = \bar{E}_x^{(r)}(t)\cos\theta + \bar{E}_y^{(r)}(t)\sin\theta \ . \qquad (2.14)$$

The measured intensity is thus

$$\begin{aligned}
I(\theta) &= 2\left\langle \bar{E}^{(r)}(\theta,t)^2\right\rangle \\
&= J_{xx}\cos^2\theta + J_{yy}\sin^2\theta + \sin(2\theta)\left[J_{xy}e^{-i\delta} + J_{yx}e^{i\delta}\right] \ . \qquad (2.15)
\end{aligned}$$

The Stokes parameters are then identified as

$$\begin{aligned} I &= J_{xx} + J_{yy} = \langle a_x^2(t)\rangle + \langle a_y^2(t)\rangle \ , \\ Q &= J_{xx} - J_{yy} = \langle a_x^2(t)\rangle - \langle a_y^2(t)\rangle \ , \\ U &= J_{xy} + J_{xy}^* = 2\langle a_x(t)a_y(t)\cos(\phi_x(t) - \phi_y(t))\rangle \ , \\ V &= -i\left[J_{xy} - J_{xy}^*\right] = 2\langle a_x(t)a_y(t)\sin(\phi_x(t) - \phi_y(t))\rangle \ . \end{aligned} \tag{2.16}$$

The measured intensity in terms of the Stokes parameters is:

$$I(\theta) = \cos^2\theta(I+Q)+\sin^2\theta(I-Q)+\frac{1}{2}\sin(2\theta)\left[(U+iV)e^{-i\delta} + (U-iV)e^{+i\delta}\right] \ . \tag{2.17}$$

The Stokes parameters can thus be determined by choosing various combinations of δ and θ and measuring $I(\theta)$. (Note that the assumption of quasi-monochromaticity is actually not necessary to define the Stokes parameters, e.g. see Wolf (1959)).

A few properties of the Stokes parameters and the associated coherency matrix $\boldsymbol{J}$ are worth noting. The Stokes Q parameter measures the amount of linear polarization in the beam in the x or y–directions. U measures the linear polarization in the directions at an angle $\pi/4$ to the x axis in the x–y plane. V measures the amount of circular polarization. If the wave is perfectly monochromatic, the amplitudes and phases of the electric field components do not vary in time. Then we may remove the time average brackets in Eqn.(2.16) and there is the following relation between the Stokes parameters:

$$I^2 = Q^2 + U^2 + V^2 \ . \tag{2.18}$$

For the general case, this constraint becomes an inequality instead:

$$I^2 \geq Q^2 + U^2 + V^2 \ . \tag{2.19}$$

The matrix $\boldsymbol{J}$ is obviously Hermitian, $J_{ij}^* = J_{ji}$. The determinant of $\boldsymbol{J}$ is:

$$\det[\boldsymbol{J}] = \frac{1}{4}\left[I^2 - (Q^2 + U^2 + V^2)\right] \geq 0 \ . \tag{2.20}$$

The *polarization magnitude* Π (or degree of polarization) is a dimensionless quantity defined by

$$\Pi^2 \equiv \frac{Q^2 + U^2 + V^2}{I^2} = 1 - 4\,\det[\boldsymbol{J}]/I^2 \ . \tag{2.21}$$

Note that most authors use the dimensionless polarization magnitude Π as defined here, but some prefer to use dimensions of specific intensity (by multiplication by the total intensity) or brightness temperature. A beam with $\Pi = 0$ is said to be *unpolarized*. A beam with $\Pi = 1$ is said to be a *pure state* (this terminology stems from the analogy [1] between coherency matrices and density matrices in quantum mechanics). By Eqn. (2.18), a perfectly monochromatic beam (as opposed to a quasi-monochromatic beam) is a pure state.

[1] See for example Simmons & Guttman (1970) and Kosowsky (1994).

If several quasi-monochromatic beams all with the same mean frequency are superimposed, and the electric fields of each beam have phases which are varying completely independently of the phases of the other beams, then the coherency matrix of the total beam is simply the sum of the coherency matrices of the separate beams. An elementary proof may be found in Rybicki & Lightman (1979) — the gist is that in the forming the time average of the quadratic products of the sum of the electric fields, the cross terms between separate beams vanish (by the assumption of the independence of the phases). Beams with electric fields with no permanent phase relations are said to be *incoherent*. We will always assume, in summing two beams with the same direction and frequency, that the beams are incoherent and thus that the coherency matrices may be summed.

A general beam can be constructed by superimposing an arbitrary number of quasi-monochromatic beams. We then obtain *spectral* Stokes parameters (Barakat, 1963) which are functions of the mean frequency ω_0, which we rename as ω for convenience. The polarization state and intensity of the beam associated with each frequency may also be a function of time. One can imagine decomposing the beam into a time series and Fourier analyzing successive segments of the time series to obtain the time dependence of each Fourier mode (this is what is actually done in polarimetric measurements of the time dependence of spectral Stokes parameters, see e.g. Costa et al. (2001).).

2.2 A tensor generalization of the coherency matrix

The polarization state and intensity of a beam of light propagating in the z–direction is characterized completely by the 2×2 Hermitian matrix J_{ij}, with $(i, j) \in \{x, y\}$. There are several papers which study a description of polarized radiation transfer using the 2×2 coherency matrix (Acquista & Anderson, 1974; Dautcourt & Rose, 1978; Bildhauer, 1989a,b, 1990; Kosowsky, 1994). An obvious generalization is to allow (i, j) to become Cartesian tensor indices and to run over all of $\{x, y, z\}$. We obtain a 3×3 matrix:

$$Q_{ij} = \left\langle E_i(t) E_j^*(t) \right\rangle , \qquad (i, j) \in \{x, y, z\} \ . \tag{2.22}$$

This matrix and its 4-dimensional generalization is one of the main tools in this thesis. It differs from the usual 2×2 coherency matrix in that it is 3×3, the extra dimension corresponding to the direction of photon propagation $\boldsymbol{n}$. Adding the extra dimension (and a fourth, when we introduce the covariant form in the next chapter) makes it much easier to handle the computation of the polarization of photons after general rotations, Lorentz boosts, and scattering.

To our knowledge, only Challinor (2000) and Carozzi et al. (2000) have systematically explored a similar approach previously. The matrix Q_{ij} is denoted the *polarization matrix* or *polarization tensor* (whether the 3-dimensional or 4-dimensional version is being talked about ought to be clear from the context). The polarization

information is contained in the normalized version of Q_{ij}, termed the *normalized polarization tensor*:

$$\phi_{ij} = \frac{Q_{ij}}{\mathrm{Tr}[\boldsymbol{Q}]} \ . \tag{2.23}$$

For a given photon direction $\boldsymbol{n}$, the polarization vector is transverse, implying

$$n^i Q_{ij} = 0 \ . \tag{2.24}$$

It is useful to define a matrix with dimensions of specific intensity, also called the polarization tensor or matrix:

$$I_{ij} = I\phi_{ij} \ , \tag{2.25}$$

where the specific intensity I and the components of ϕ_{ij} are associated with some mean frequency ω as discussed in the last section. The transition from the quasi-monochromatic case to the general polychromatic case may be taken as discussed in the previous section, and the components become functions of photon frequency. In general, the polarization matrix is a function of photon frequency (or momentum) and direction as well as space and time:

$$I_{ij} = I_{ij}(\nu, \boldsymbol{n}, \boldsymbol{x}, t) \ . \tag{2.26}$$

Other conventions are also useful – in the computation of the Sunyaev-Zeldovich effects (SZE), it will be convenient to work with polarization matrices whose trace is either the occupation number $n(\nu, \boldsymbol{n}, \boldsymbol{x}, t)$ or the phase space distribution function $f(\nu, \boldsymbol{n}, \boldsymbol{x}, t)$ (associated with a particular photon momentum state and spatial position). Since the Stokes parameters are usually taken to have dimensions of specific intensity, we usually work with I_{ij}, but it is occasionally useful to use the other forms.

Now in the usual description of polarized light, the Stokes parameters are defined with respect to a particular choice of "polarization basis". This is a pair of mutually orthogonal unit vectors $\boldsymbol{e}^{(1)}, \boldsymbol{e}^{(2)}$, both orthogonal to the beam direction. The Stokes parameters Q and U depend on the orientation of these vectors. By contrast the polarization matrix is a tensor and its components in any basis contain all the information about the polarization ellipse. Its advantage is that there is no need to rotate axes to define Stokes parameters. The Stokes parameters are given in terms of the polarization matrix and the polarization basis vectors as:

$$\frac{1}{2}\begin{pmatrix} I+Q & U+iV \\ U-iV & I-Q \end{pmatrix} \ = \ e_i^{(a)} e_j^{(b)} I_{ij} \ , \qquad (a,b) \in (1,2) \ , \tag{2.27}$$

(the sum over the Cartesian indices ij is implied) which is just the previously defined coherency matrix $\boldsymbol{J}$ of equation (2.11).

It is of interest to see how the Stokes parameters transform if we choose a rotated set of basis vectors. In the case of a beam propagating in the z-direction for example,

we have, choosing polarization basis vectors $\boldsymbol{e}^{(1)} = \boldsymbol{x},\ \boldsymbol{e}^{(2)} = \boldsymbol{y}$,

$$I_{ij} = \frac{1}{2}\begin{pmatrix} I+Q & U+iV & 0 \\ U-iV & I-Q & 0 \\ 0 & 0 & 0 \end{pmatrix} \ . \tag{2.28}$$

If the basis vectors are rotated clockwise (according to an observer looking in the direction of propagation) through an angle χ, the new set of basis vectors is

$$\begin{aligned} \boldsymbol{e}_R^{(1)} &= \cos\chi\,\boldsymbol{e}^{(1)} + \sin\chi\,\boldsymbol{e}^{(2)} \ , \\ \boldsymbol{e}_R^{(2)} &= \cos\chi\,\boldsymbol{e}^{(2)} - \sin\chi\,\boldsymbol{e}^{(1)} \ . \end{aligned} \tag{2.29}$$

Forming the matrices $e_{R,i}^{(a)} e_{R,j}^{(b)}$, with $(a,b) \in \{1,2\}$, the primed Stokes parameters according to Eqn. (2.27) are:

$$\begin{aligned} I' &= I \\ Q' &= Q\cos 2\chi + U\sin 2\chi \ , \\ U' &= U\cos 2\chi - Q\sin 2\chi \ , \\ V' &= V \ . \end{aligned} \tag{2.30}$$

These transformations are also obtained directly from Q_{ij} by forming the rotation matrix:

$$\boldsymbol{R}(\chi) = \begin{pmatrix} \cos\chi & -\sin\chi & 0 \\ \sin\chi & \cos\chi & 0 \\ 0 & 0 & 1 \end{pmatrix} \ . \tag{2.31}$$

Then

$$\frac{1}{2}\begin{pmatrix} I'+Q' & U'+iV' & 0 \\ U'-iV' & I'-Q' & 0 \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{R}(\chi)\boldsymbol{Q}\boldsymbol{R}^T(\chi) \ . \tag{2.32}$$

These factors of $\cos 2\chi$, $\sin 2\chi$ in the transformation law are well known and associated with the fact that the linear polarization is described by a "headless vector" which is invariant under a rotation through π radians.

Now, given a set of matrix elements I_{ij}, supposed to represent a beam propagating in the direction $\boldsymbol{n}$, how do we go about deciding if this matrix can represent a physical beam? Clearly the matrix must be Hermitian and satisfy $I_{ij}n^j = 0$. This yields a matrix whose elements contain four independent real quantities. In addition, the elements must satisfy some analogue of the relation between the Stokes parameters Eqns. (2.18) or (2.19). The required condition is apparent from Eqn. (2.20) — the eigenvalues of the matrix $\boldsymbol{I}$ must be non-negative.

Another obvious question to ask is, how does one construct the matrix of an unpolarized beam propagating in a general direction $\boldsymbol{n}$? The only quantities we have available to construct the matrix are the intensity I, the components of the direction

vector $\boldsymbol{n}$, and the Kronecker delta δ_{ij}. The matrix must therefore be of the form:

$$I_{ij}(\boldsymbol{n}) = A\delta_{ij} + Bn_in_j \ . \tag{2.33}$$

Now the matrix of an unpolarized beam propagating in the z–direction is obviously

$$I_{ij} = \frac{I}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{2.34}$$

Comparing this with the form of Eqn. (2.33) for the special case $n_i = \delta_{iz}$, we see that $A = -B = I/2$. Thus the matrix of an unpolarized beam in a general direction $\boldsymbol{n}$ is

$$I_{ij}(\boldsymbol{n}) = \frac{I}{2}(\delta_{ij} - n_in_j) = \frac{I}{2}P_{ij}(\boldsymbol{n}) \ , \tag{2.35}$$

where we have defined the projection matrix $\boldsymbol{P}$ which will figure prominently later.

The polarization magnitude (squared) of the beam described by a general matrix I_{ij} is given by

$$\Pi^2 = \frac{2\mathrm{Tr}[\boldsymbol{I}^2]}{\mathrm{Tr}[\boldsymbol{I}]^2} - 1 \ . \tag{2.36}$$

This is readily checked with the matrix (2.28) of a beam propagating in the $\boldsymbol{z}$ direction. To see that this relation is true for any beam, we need only note that the matrix of a beam propagating in a general direction is related to (2.28) by a similarity transformation with an orthogonal rotation matrix, which does not change the traces in Eqn. (2.36). Note also that reality of the right hand side of Eqn. (2.36) follows automatically from the Hermiticity of $\boldsymbol{I}$ (since $\boldsymbol{I}$ and $\boldsymbol{I}^2$ are Hermitian, and the trace of a Hermitian matrix is real).

In the computation of the Sunyaev-Zeldovich effect in the single scattering limit, derived in detail in §4.2, we have a situation where the scattered beam consists of an unpolarized component plus a small polarized perturbation proportional to the optical depth to scattering, τ. It is useful to compute at this point an expression for the polarization matrix of the total beam to first order in the intensity of the perturbation. From Eqn. (2.35), the beam has polarization matrix

$$I_{ij}(\boldsymbol{n}) = I^0_{ij} + \tau\Delta I_{ij}(\boldsymbol{n}) \ , \quad \text{where} \quad I^0_{ij} = \frac{I_0}{2}P_{ij}(\boldsymbol{n}) \ , \tag{2.37}$$

or in matrix notation, $\boldsymbol{I} = \boldsymbol{I^0} + \tau\boldsymbol{\Delta I}$, and $\boldsymbol{I^0} = (I_0/2)\boldsymbol{P}(\boldsymbol{n})$. Substituting this into Eqn. (2.36) we find, in matrix notation

$$\Pi^2(\boldsymbol{I^0} + \tau\boldsymbol{\Delta I}) \ = \ \frac{2\tau\mathrm{Tr}\left[2\boldsymbol{I^0\Delta I} - I_0\boldsymbol{\Delta I}\right] + \tau^2\left(2\mathrm{Tr}[\boldsymbol{\Delta I}^2] - \mathrm{Tr}[\boldsymbol{\Delta I}]^2\right)}{(I_0 + \tau\mathrm{Tr}[\boldsymbol{\Delta I}])^2} \ . \tag{2.38}$$

Now the unpolarized part of the beam is just a projection matrix multiplied by a

scalar, so it has the property:

$$\mathrm{Tr}[\boldsymbol{I^0}\boldsymbol{\Delta I}] = \frac{I_0}{2}\,\mathrm{Tr}[\boldsymbol{\Delta I}]\ . \tag{2.39}$$

Therefore the first trace in the numerator in Eqn. (2.38) vanishes. The second term in the denominator can be ignored in the limit of a small perturbation intensity, and the squared polarization magnitude reduces to

$$\Pi^2(\boldsymbol{I^0} + \tau\boldsymbol{\Delta I}) \approx \tau^2 \left(\frac{\mathrm{Tr}[\boldsymbol{\Delta I}]}{\mathrm{Tr}[\boldsymbol{I^0}]}\right)^2 \left[\frac{2\mathrm{Tr}[\boldsymbol{\Delta I}^2]}{\mathrm{Tr}[\boldsymbol{\Delta I}]^2} - 1\right]\ . \tag{2.40}$$

In other words, the polarization magnitude of the total beam is just that of the polarized perturbation multiplied by the ratio of the intensity of the polarized part relative to the unpolarized part:

$$\Pi(\boldsymbol{I^0} + \tau\boldsymbol{\Delta I}) \approx \tau \left(\frac{\mathrm{Tr}[\boldsymbol{\Delta I}]}{\mathrm{Tr}[\boldsymbol{I^0}]}\right) \Pi(\boldsymbol{\Delta I})\ . \tag{2.41}$$

Finally in this section, we note that the polarization matrices of incoherent beams associated with the same direction and frequency may simply be summed, by an obvious extension of the proof for coherency matrices mentioned in §2.1.

2.3 Extension to a covariant polarization tensor

The discussion so far has been in terms of electric fields measured in a particular Lorentz frame. In treating problems involving scattering from a moving medium, it is necessary to Lorentz transform the fields between frames. This can be done explicitly by writing down the time dependent electric and magnetic fields of the waves, and using the transformation law of the fields. However it turns out to be much simpler to use an extension of the matrix approach we have described in which the beam is described by a second rank tensor on spacetime. In this approach the Lorentz transformations become simple tensor (or matrix) relations. Indeed a full development of the radiative transfer of polarized light on a curved spacetime is possible with this covariant formalism. In this section we work in a curved spacetime initially but eventually restrict to flat spacetime, which is adequate for our application to the SZE. We use the Minkowski metric with the convention $g_{\mu\nu} = \mathrm{diag}\{-1, 1, 1, 1\}$. The coordinates of a point in spacetime will be denoted either abstractly as x, or as an upper index quantity $x^\mu = (t, x, y, z)$. Latin indices will denote components in the orthonormal basis $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$.

A truly covariant description of the electromagnetic field requires introduction of the field strength tensor $F_{\alpha\beta}$, and indeed a covariant description of the polarization of light can be accomplished entirely in terms of the field strength tensor (Dialetis, 1969). But we wish to maintain an explicit connection with the Stokes parameters which are defined as time averaged quadratic combinations of electric field amplitudes, as measured by an observer at rest in some Lorentz frame. Thus we must express the

electric field amplitudes measured in the rest frame of a given observer in a Lorentz covariant manner. The rest frame of the observer along the light beam can be defined by specifying a differentiable time-like vector field $\vec{v}(x)$ giving the observer 4-velocity all along the light cone (with $\vec{v} \cdot \vec{v} = -1$).

To generalize the coherency matrix of the previous sections, we need to find a covariant way to describe the time averaged product of electric fields. This must be done by constructing the electromagnetic field strength tensor for a plane wave in the WKB (or shortwave) approximation of geometrical optics (see e.g. Born & Wolf (1980); Schneider et al. (1992); Misner et al. (1973)). In this approximation we treat the antisymmetric electromagnetic field strength tensor $F_{\mu\nu}$ as a test field (meaning that we may ignore the influence of $F_{\mu\nu}$ on the gravitational field) and assume that there are no charges or currents in the region we are considering. The field tensor thus obeys the source free Maxwell equations:

$$\nabla_\mu F^{\mu\nu} = 0 \ , \quad \nabla_\alpha F_{\mu\nu} + \nabla_\mu F_{\nu\alpha} + \nabla_\nu F_{\alpha\mu} = 0 \ . \tag{2.42}$$

The geometrical optics approximation consists in assuming that the field strength tensor can be written as the product of a slowly varying complex amplitude and a relatively rapidly varying phase factor:

$$F_{\mu\nu} = \Re\mathrm{e}\left\{\widetilde{F}_{\mu\nu}(x)\exp[i\varphi(x)/\epsilon]\right\} \tag{2.43}$$

where $\epsilon \sim \lambda/L$ is a perturbation parameter with λ being the wavelength and L the length-scale over which the amplitude $\widetilde{F}_{\mu\nu}$ changes (roughly the local radius of curvature of spacetime). In the geometrical optics limit we expand the Maxwell equations in an asymptotic series in ϵ, take the limit $\epsilon \to 0$, and read off the lowest order terms. Then ϵ is absorbed into $\varphi(x)$, by replacing $\varphi(x)/\epsilon$ with $\widetilde{\varphi}(x)$ and then dropping the tilde. The lowest order terms describe the evolution of electromagnetic waves which, on scales which are large compared to λ but small compared to L, are plane and monochromatic to an excellent approximation.

Substituting equation (2.43) into the Maxwell equations (2.42) and working to lowest order in ϵ, we obtain

$$\begin{aligned} k_\mu \widetilde{F}^{\mu\nu} &= 0 \ , \\ k_\alpha \widetilde{F}_{\mu\nu} + k_\mu \widetilde{F}_{\nu\alpha} + k_\nu \widetilde{F}_{\alpha\mu} &= 0 \ . \end{aligned} \tag{2.44}$$

where the wavevector k_μ is a one-form field normal to surfaces of constant phase, defined by:

$$k_\mu(x) \equiv \nabla_\mu \varphi \ . \tag{2.45}$$

It follows from this, and the fact that covariant derivatives commute when applied to a scalar field, that

$$\nabla_\mu k_\nu = \nabla_\nu k_\mu \ . \tag{2.46}$$

Contracting the second equation in (2.44) with k^α, and assuming that $\widetilde{F}_{\mu\nu}$ vanishes

only on hyper-surfaces, we find

$$k^\mu k_\mu = 0 \ . \tag{2.47}$$

Thus the wavevector $\vec{k}$ is null. If desired we may associate a photon 4-momentum $\vec{p} = \hbar\vec{k}$ with the wavevector, and go over to a particle description. The frequency of the wave as measured by a local observer with worldline $x^\mu(\tau)$ and 4-velocity $u^\mu = dx^\mu/d\tau$ is given by $\omega = -\vec{k}\cdot\vec{u} = d\varphi/d\tau$ (taking $\epsilon = 1$). Eqns. (2.46) and (2.47) imply that the wavevector is parallel transported:

$$\nabla_k k_\mu \equiv k^\alpha \nabla_\alpha k_\mu = 0 \ . \tag{2.48}$$

The curves $x^\mu(\lambda)$ with $dx^\mu/d\lambda = k^\mu$ are called light rays (λ is an affine parameter along the ray). Note $\nabla_k = k^\mu\nabla_\mu = d/d\lambda$ is the directional derivative along the ray. As a consequence of Eqn. (2.48), the system of rays is equivalent to a Hamiltonian flow for particles with Hamiltonian

$$H(x,k) = \frac{1}{2} g^{\mu\nu}(x) k_\mu k_\nu. \tag{2.49}$$

Hamilton's equation $dk_\mu/d\lambda = -\partial H/\partial x^\mu$ is equivalent to Eqn. (2.48), which is the geodesic equation for photons, while $dx^\mu/d\lambda = \partial H/\partial k_\mu$ gives the advance of the wavefront along the ray. The Hamilton-Jacobi equation $H(x, \nabla\varphi) = 0$ is also known as the eikonal equation for the phase factor $\varphi(x)$.

Now we would like to express the components of $F_{\mu\nu} = -F_{\nu\mu}$ in terms of the electric field. Writing a propagation equation for the electric field requires that we have a differentiable time-like vector field v^μ ($\vec{v}\cdot\vec{v} = -1$) giving the 4-velocity (hence rest frame) of observers all along the light cone. In other words, the electric field is defined with respect to a family of observers with 4-velocity $\vec{v}(x)$. In the local Lorentz frame at point x of the observer with 4-velocity $\vec{v}(x)$, the electric field components are $E_{\hat{i}} = F_{\hat{i}\hat{0}}$ and the magnetic field components are $B_{\hat{i}} = \frac{1}{2}\epsilon_{\hat{i}\hat{l}\hat{m}}F^{\hat{l}\hat{m}}$ where Latin indices range over the spatial components and the carets indicate an orthonormal basis, with $\vec{e}_{\hat{0}} = \vec{v}$. The transversality from equation (2.44) implies $B_{\hat{i}} = \epsilon_{\hat{i}\hat{l}\hat{m}}\hat{k}^{\hat{l}}E^{\hat{m}}$ where $\hat{k}$ is a spatial unit vector along the wavevector. In a general basis, we promote the electric field to a 4-vector

$$E_\mu \equiv v^\nu F_{\mu\nu} \ . \tag{2.50}$$

By antisymmetry of $F_{\mu\nu}$, E_μ is orthogonal to the 4-velocity of the observer v^μ:

$$v^\mu E_\mu = 0 \ . \tag{2.51}$$

In the geometrical optics limit we may define the complex amplitude of the 4-vector electric field using the complex amplitude of the field strength tensor:

$$\widetilde{E}_\mu \equiv v^\nu \widetilde{F}_{\mu\nu} \ , \quad v^\mu \widetilde{E}_\mu = 0 \ . \tag{2.52}$$

Thus

$$E_\mu = \Re e\{\widetilde{E}_\mu \exp[i\varphi(x)/\epsilon]\} \ . \tag{2.53}$$

Eqns. (2.44) imply $k^\mu E_\mu = k^\mu \widetilde{E}_\mu = 0$, which correspond to the transversality of the electric field to the the wavevector. The electric field 4-vector may be factored as

$$E_\mu = E\epsilon_\mu \ , \tag{2.54}$$

where $\vec{\epsilon}$ is a vector which satisfies $g^{\mu\nu}\epsilon^*_\mu\epsilon_\nu = 1$, called the *electric polarization vector*. In the rest frame of $\vec{v}$, this reduces to a 4-vector with spatial parts equal to the usual polarization 3-vector.

Contracting the second of Eqns. (2.44) with v^α and substituting Eqn. (2.52) yields an expression for the field strength amplitude in terms of the electric field 4-vector amplitude:

$$\widetilde{F}_{\mu\nu} = k^{-1}(k_\mu \widetilde{E}_\nu - k_\nu \widetilde{E}_\mu) \quad \text{where} \quad k \equiv -k_\mu v^\mu \ . \tag{2.55}$$

Next we would like to know how the amplitudes $\widetilde{F}_{\mu\nu}$ and $\widetilde{E}_\mu$ change along a ray. We proceed by computing the divergence of the second of the Maxwell equations (2.42).

$$g^{\alpha\beta}\nabla_\beta \left[\nabla_\alpha F_{\mu\nu} + \nabla_\mu F_{\nu\alpha} + \nabla_\nu F_{\alpha\mu}\right] = 0 \ . \tag{2.56}$$

Note that swapping the order of the covariant derivatives in the second two terms kills each term by Maxwell's equations. Thus using the following identity for the commutator of covariant derivatives in terms of the Riemann tensor,

$$(\nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu)S_{\alpha\beta} = -S_{\sigma\beta}R^\sigma{}_{\alpha\mu\nu} - S_{\alpha\sigma}R^\sigma{}_{\beta\mu\nu} \ , \tag{2.57}$$

we find a wave equation for $F_{\mu\nu}$ with curvature terms:

$$g^{\alpha\beta}\nabla_\alpha\nabla_\beta F_{\mu\nu} - R_{\mu\alpha}F^\alpha{}_\nu + R_{\nu\alpha}F^\alpha{}_\mu + R_{\mu\nu\alpha\beta}F^{\alpha\beta} = 0 \ . \tag{2.58}$$

Substituting equation (2.43) and working to the two lowest orders in ϵ, one finds the following equation for the evolution of the field strength amplitude (the Riemann tensor terms do not appear to this order):

$$\nabla_k \widetilde{F}_{\mu\nu} = -\frac{1}{2}\theta\widetilde{F}_{\mu\nu} \ , \quad \theta \equiv \nabla_\alpha k^\alpha \ . \tag{2.59}$$

The amplitude of the electromagnetic field changes along rays due to curvature of the wavefronts. For example, diverging rays ($\theta > 0$) lead to a decrease in the electromagnetic field strength as the wave propagates.

Substituting equation (2.55) into equation (2.59) now gives an equation for the electric field evolution along a ray,

$$\nabla_k \widetilde{E}_\mu = \left(\nabla_k \ln k - \frac{1}{2}\theta\right)\widetilde{E}_\mu + \frac{k_\mu}{k}\left(\nabla_k v^\alpha\right)\widetilde{E}_\alpha \ . \tag{2.60}$$

Factoring the electric field into its magnitude and direction (polarization) vector,

$\widetilde{E}_\mu = \widetilde{E}\epsilon_\mu$ where $g^{\mu\nu}\epsilon_\mu\epsilon_\nu = 1$ and $v^\mu\epsilon_\mu = 0$, we obtain

$$\nabla_k \widetilde{E} = \left(\nabla_k \ln k - \frac{1}{2}\theta\right)\widetilde{E} \ , \tag{2.61a}$$

$$\nabla_k \epsilon_\mu = \frac{k_\mu}{k}\left(\nabla_k v^\alpha\right)\epsilon_\alpha \ . \tag{2.61b}$$

The first of these equations yields for example the $1/r$ fall off of the electric field magnitude expected for a radiation field. The right-hand side of both equations vanishes for a plane wave in flat space, but not for a curved wavefront (e.g. a spherical wave), or for a wave propagating in a general curved space.

It is perhaps surprising that the electric polarization vector is not parallel transported in a curved spacetime. This fact leads to a rotation of the polarization vector when a beam passes through a strong gravitational field. This effect has been noted before by several authors (Skrotskii, 1957; Plebanski, 1960; Nouri-Zonoz, 1999; Kopeikin & Mashhoon, 2002), and is important in considering for example the propagation of polarized radiation in the vicinity of a black hole. However it is true that if one defines the polarization vector to be parallel to the vector potential rather than the electric field of the electromagnetic wave, then it is parallel transported in the geometrical optics limit (see e.g. Misner et al. (1973), Schneider et al. (1992)). This turns out to be consistent with the electric field evolution due to the enforcement of the gauge choice of the vector potential all along the photon worldline. Thus in considering the propagation of polarized photons on a curved spacetime it is more convenient to use a polarization tensor constructed from the vector potential to evolve the polarization state along the ray, and then make the transformation to electric fields. If the photon path does not pass through regions with an exceptionally strong gravitational field however, the rotation of the polarization vector resulting from this gravitational effect is small (but note that, strictly speaking, the rotation arises from the acceleration of the local observers, $\nabla_k v^\alpha = dv^\alpha/d\lambda$, which can be large even in flat spacetime if a peculiar vector field of observer 4-velocities is chosen). In considering the propagation of photons through a cluster of galaxies for example, the effect is entirely negligible, and so henceforth we will restrict the discussion to flat spacetime, and work with the more physical polarization tensors we defined in terms of the electric fields. In flat spacetime we may drop the right hand sides in Eqns. (2.61).

Having described the propagation of electromagnetic waves in the geometrical optics approximation, and defined the electric field in a covariant manner, we are equipped to construct the covariant version of the coherency matrix. We consider a plane electromagnetic wave propagating in flat spacetime, in the geometrical optics limit, with wavevector $\vec{k}$ and associated photon momentum $\vec{p}$. Henceforth we will write the complex amplitude of the 4-vector electric field of the wave as $\widetilde{E}^\mu$, dropping the tilde for brevity. The 4-vector $\vec{E}$ has the property that its spatial components E^i in the rest frame of the observer, in which $v^\mu = (1,0,0,0)$, are equal to the measured electric field, and also $E^0 = 0$ in this frame. Thus by analogy with the 3×3 polarization matrix Q_{ij} defined in Eqn. (2.22), we are lead to define a complex

valued rank $(0,2)$ tensor called the *polarization tensor*:

$$Q_{\mu\nu}(x,\vec{p},\vec{v}) \equiv \langle E_\mu E^*_\nu \rangle \ . \tag{2.62}$$

The spatial components of this tensor in the rest frame of the observer are entirely equivalent to the elements of the 3×3 coherency matrix considered in the previous section. It is related to the stress-energy tensor $T^{\mu\nu} = F^\mu{}_\alpha F^{\nu\alpha} - \frac{1}{4} g^{\mu\nu} F^{\alpha\beta} F_{\alpha\beta}$ (in Heaviside-Lorentz units). In particular, the time-average energy density in the geometrical optics limit is $v_\mu v_\nu \langle T^{\mu\nu}\rangle = Q \equiv Q^\mu{}_\mu$ where angle brackets denote averaging over a few periods. Note that $Q^{\mu\nu} v_\mu = Q^{\mu\nu} p_\mu = 0$ (where p^μ is the four momentum of the photon).

To define Stokes parameters, we need to specify a set of polarization basis vectors. The natural choice is the orthonormal tetrad basis vectors $\{\vec{\epsilon}_a\}$:

$$\vec{\epsilon}_0 = \vec{v} \ , \quad \vec{\epsilon}_3 = p^{-1}\vec{p} - \vec{v} \ , \quad \vec{\epsilon}_1 \ , \quad \vec{\epsilon}_2 \tag{2.63}$$

where $p \equiv -\vec{v}\cdot\vec{p}$. These vectors have the property $\vec{\epsilon}_a \cdot \vec{\epsilon}_b = \eta_{ab}$. Latin indices $\{a,b,\ldots\}$ are tetrad indices; Greek indices $\{\mu,\nu,\ldots\}$ are coordinate indices. The spatial direction of the photon momentum for an observer with 4-velocity $\vec{\epsilon}_0$ is $\vec{\epsilon}_3$. The remaining basis vectors, $\vec{\epsilon}_1$ and $\vec{\epsilon}_2$, give the physical polarization space. We call this tetrad the polarization tetrad. There are associated basis 1–forms $\{\tilde{\epsilon}^a\}$), which are dual to the basis vectors: $\langle \tilde{\epsilon}^a, \vec{\epsilon}_b\rangle = \delta^a{}_b$. (Note that the polarization tetrad depends on the photon momentum, i.e. $\vec{\epsilon}_a = \vec{\epsilon}_a(x,p)$. Thus, in general a different basis is needed for each photon momentum). The coherency matrix of Eqn. (2.11) is then given by

$$J_{ab} = \epsilon^\mu_a \epsilon^\nu_b \, Q_{\mu\nu}(x,\vec{p},\vec{v}) \ , \qquad (a,b) \in (1,2) \ . \tag{2.64}$$

In the case of beam which has a definite polarization vector $\vec{\epsilon}$ (lying in the polarization subspace spanned by $\{\vec{\epsilon}_1,\vec{\epsilon}_2\}$) which does not vary with time, i.e. a pure beam, the polarization tensor is given by

$$Q_{\mu\nu}(x,\vec{p},\vec{v}) = Q \ \epsilon_\mu \epsilon_\nu \ . \tag{2.65}$$

So far we have used the tensor $Q^{\mu\nu}$ to describe a polarized EM wave. However if we wish to consider energy transfer between photons and some scattering medium, free electrons for example, we must consider the trajectories of photons in phase space. To describe an ensemble of polarized photons we must define a distribution function on phase space. The matrix $Q^{\mu\nu}$ is not very useful because it describes a single classical system (the classical counterpart of a pure state) with specified wavevector $k_\mu(x)$. Developing a kinetic theory requires an ensemble of systems encompassing a continuous distribution of wave-vectors k_μ at each spacetime point. We accomplish this heuristically by analogy with the usual treatment of the unpolarized case.

In general, the stress-energy tensor of a system of photons may be written (in flat

space) as

$$T^{\mu\nu}(x) \;=\; \int \frac{d^3p}{p}\, p^\mu p^\nu f(x,\vec{p}) \;, \tag{2.66}$$

where $p = p^0$ and $f(x,\vec{p})$ is the (unpolarized) photon distribution function, which determines the total number of photons in the quantum state corresponding to phase space element $d^3x\, d^3p$ according to

$$dN = f(x,\vec{p})\; d^3x\, d^3p \;. \tag{2.67}$$

The occupation number is given by $n(x,\vec{p}) = h^3 f(x,\vec{p})$. It is not hard to prove that n and f are Lorentz scalars (see for example Lightman et al. (1975) for a proof).

To incorporate polarization we define the *distribution function polarization tensor* $f_{\mu\nu}(x,\vec{p},\vec{v})$ in a manner similar to the scalar distribution function $f(x,\vec{p})$. The polarization tensor of a general superposition of waves at a given spacetime point, according to the local observer with 4-velocity $\vec{v}$, may be defined as $Q_{\mu\nu}(x,\vec{v}) \equiv \langle E_\mu E^*_\nu \rangle$. Then by analogy with (2.66) we have

$$Q^{\mu\nu}(x,\vec{v}) = \int \frac{d^3p}{p}\, p^2\; f^{\mu\nu}(x,\vec{p},\vec{v}) \;. \tag{2.68}$$

This obviously does not uniquely define $f_{\mu\nu}(x,\vec{p},\vec{v})$. A rigorous definition requires a more sophisticated discussion (as in Bildhauer (1989a)). However we do not run into any difficulties if we simply regard $f_{\mu\nu}(x,\vec{p},\vec{v})$ as a tensor generalization of the scalar distribution function which satisfies Eqn. (2.68). $f^{\alpha\beta}$ has the property that $f^{\alpha\beta}\epsilon_\alpha\epsilon^*_\beta\; d^3p\; d^3x$ is proportional to the number of photons in the phase space element $d^3p\; d^3x$ passing per unit time through a polarizer oriented to select polarization ϵ^α (where this 4-vector must lie in the polarization subspace spanned by the vectors $\vec{\epsilon}_1$, $\vec{\epsilon}_2$ of Eqn. (2.63)). Contraction with the metric yields the scalar distribution function:

$$g_{\mu\nu} f^{\mu\nu}(\vec{p},\vec{v}) = f(\vec{p}) \;. \tag{2.69}$$

The distribution function tensor also has the properties

$$v_\alpha f^{\alpha\beta}(\vec{p},\vec{v}) = p_\alpha f^{\alpha\beta}(\vec{p},\vec{v}) = 0 \;, \quad f^{\beta\alpha} = \left(f^{\alpha\beta}\right)^* \;. \tag{2.70}$$

The generalization of Eqn. (2.36) for the polarization magnitude is

$$\Pi^2(\vec{p}) = \frac{2 f^{\mu\nu}(\vec{p},\vec{v}) f_{\mu\nu}(\vec{p},\vec{v})}{\left[g_{\alpha\beta} f^{\alpha\beta}(\vec{p},\vec{v})\right]^2} - 1 \;, \tag{2.71}$$

which is manifestly a Lorentz scalar.

Similarly to the case for coherency matrices and 3×3 polarization matrices, the polarization tensors $f_1^{\mu\nu}(\vec{p},\vec{v})$ and $f_2^{\mu\nu}(\vec{p},\vec{v})$ of two incoherent beams associated with the same photon momentum may be summed to yield the total polarization tensor. We shall always make the assumption that two separate beams are incoherent and

have polarization tensors that may be superimposed in this manner.

It is useful to define a covariant polarization tensor with dimensions of specific intensity (whose components are combinations of Stokes parameters). In the unpolarized case, the specific intensity $I(x,\vec{p})$ is introduced by defining

$$T^{00} = \int d^3p\, p\, f(x,\vec{p}) = \int d\nu\, d\Omega\; I(x,\vec{p}) \;, \tag{2.72}$$

where $d\Omega$ is the solid angle element about the photon direction. It follows from (2.66) that

$$I = hp^3\, f = h^4\nu^3\, f = h\nu^3 n \;. \tag{2.73}$$

We define a specific intensity (or brightness) tensor by analogy with the unpolarized case:

$$I^{\mu\nu}(x,\vec{p},\vec{v}) \equiv hp^3 f^{\mu\nu}(x,\vec{p},\vec{v}) \;. \tag{2.74}$$

The intensity polarization matrix is zero outside from the two-dimensional polarization space spanned by $\vec{\epsilon}_1$, $\vec{\epsilon}_2$, where it may be written in terms of the usual Stokes parameters I, Q, U, and V:

$$f_{ab} = \frac{1}{2hp^3}\begin{pmatrix} I+Q & U-iV \\ U+iV & I-Q \end{pmatrix} \tag{2.75}$$

The Stokes parameters are functions of frequency (photon energy); $I = I_\nu$ is the spectral intensity. In an arbitrary basis the intensity is $I = (h\nu^3/c^2)g^{\mu\nu}f_{\mu\nu}$. The normalization factor is chosen so that f_{11} is the photon occupation number (phase space density divided by h^3) for photons passed by a linear polarizer oriented along $\vec{\epsilon}_1$ (and similarly for other directions). In terms of the total spectral intensity, we may write the polarization magnitude as $\Pi^2 = 2(I^{-2})I^{\mu\nu}I_{\mu\nu} - 1$.

Now we wish to obtain an equation for the evolution of $f^{\mu\nu}$ in time. In the absence of scattering, photons follows geodesics (free stream) and the distribution function evolves according to the Liouville equation. The Liouville equation for the unpolarized distribution function is simply $Df/d\lambda = 0$, where λ is an affine parameter along the ray:

$$\frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\mu} + \frac{dp_\mu}{d\lambda}\frac{\partial f}{\partial p_\mu} \;=\; 0 \;. \tag{2.76}$$

This may also be written in the more familiar form (valid in curved spacetime)

$$\nabla_p f - \Gamma^i_{\mu\nu}p^\mu p^\nu \frac{\partial f}{\partial p^i} \;=\; 0 \quad \text{where } \nabla_p = p^\alpha \nabla_\alpha \;, \tag{2.77}$$

provided one regards f as a function of the 3-momentum in some frame, $f = f(\boldsymbol{p})$ (not $f = f(\vec{p})$), by enforcing the mass shell constraint $p^\mu p_\mu = 0$.

Now we consider the generalization to the polarized case. From Eqns. (2.61) it follows that the evolution equation for $Q_{\mu\nu}$ in the geometrical optics approximation

in flat spacetime is

$$\nabla_k Q_{\mu\nu} = 0 \ . \tag{2.78}$$

This is suggestive that the Liouville equation for $f_{\mu\nu}$ in flat spacetime is simply

$$\nabla_p f_{\mu\nu} = 0 \ , \tag{2.79}$$

where in flat spacetime $\nabla_p = p^\alpha \partial_\alpha$ of course. This obviously reduces to the correct evolution of the unpolarized distribution function on taking the trace. This is in fact the correct Liouville equation in the polarized case, as has been argued in the more sophisticated treatments of Dautcourt & Rose (1978); Bildhauer (1989a,b, 1990); Breuer & Ehlers (1980, 1981).

The full polarized Boltzmann (or kinetic) equation for the distribution function polarization tensor is

$$\nabla_p f_{\mu\nu} = C_{\mu\nu} \ , \tag{2.80}$$

where the effect of scattering is contained in the *scattering term* $C_{\mu\nu}$. In Chapter 3, the general form of the scattering term is derived for the case of Compton scattering.

2.4 Lorentz transformation properties of the polarization tensor

On performing a Lorentz transformation between inertial frames, it is well known that the propagation direction of an EM wave (or equivalently, photon) is aberrated and its frequency (or momentum) Doppler boosted. The transformation of the polarization state of the beam is less well known. Here we derive the transformation law of the polarization tensor between frames. This leads to the transformation law for the Stokes parameters, which turns out to be very simple (in fact, they are invariant) provided a certain choice of polarization basis is made.

First, we find the transformation of the 4-vector electric field $E^\mu(\vec{v})$ under a change of the local observer vector field from $\vec{v}(x)$ to $\vec{v}'(x)$. The spatial components of $E^i(\vec{v})$ are the electric field (3-vector) components measured by the observer with four-velocity $\vec{v}$ (in her rest frame). Let us find the relationship between the electric fields $E^\mu(\vec{v})$ and $E^\mu(\vec{v}')$. To determine this, recall from (2.55) that the definition of E^μ implies the following relation between E_μ and the field strength tensor for a plane wave:

$$F_{\mu\nu} = p^{-1}(p_\mu E_\nu - p_\nu E_\mu) \ , \quad p \equiv -\vec{p}\cdot\vec{v} \ . \tag{2.81}$$

Therefore, since $E^\mu(\vec{v}) = F^{\mu\nu} v_\nu$,

$$E^\mu(\vec{v}') \ = \ \frac{p'}{p}\left(g^\mu{}_\nu + \frac{p^\mu v'_\nu}{p'}\right) E^\nu(\vec{v}) \ , \tag{2.82}$$

where $p' \equiv -\vec{v}'\cdot\vec{p}$. These relations suggest introduction of a tensor $P_{\mu\nu}(\vec{p},\vec{v}\,)$ which projects onto the physical polarization plane $\vec{\epsilon}_1$-$\vec{\epsilon}_2$ by eliminating components in the

surface spanned by $\vec{\epsilon}_0$ and $\vec{\epsilon}_3$ (or $\vec{v}$ and $\vec{p}$):

$$\begin{aligned} P_{\mu\nu}(\vec{p},\vec{v}) &= g_{\mu\nu} + \epsilon_{0\mu}\epsilon_{0\nu} - \epsilon_{3\mu}\epsilon_{3\nu} \\ &= g_{\mu\nu} + \frac{1}{p}(p_\mu v_\nu + p_\nu v_\mu) - \frac{p_\mu p_\nu}{p^2} \qquad \text{where} \quad p \equiv -v^\mu p_\mu \ . \end{aligned} \tag{2.83}$$

This satisfies the idempotency relation $P^\mu{}_\alpha P^\alpha{}_\nu = P^\mu{}_\nu$, so that $P^\mu{}_\nu$ is a projection tensor, henceforth denoted the *screen projection tensor* which will prove to be important. The transformation law for measured electric fields, equation (2.82), may be written in terms of the screen projection tensor as follows:

$$E_\mu(\vec{v}') = \frac{v'_\alpha p^\alpha}{v_\beta p^\beta}\, P^\nu{}_\mu(\vec{p},\vec{v}') E_\nu(\vec{v}) \ , \tag{2.84}$$

since the second and the last term in $P^\nu{}_\mu$ vanish when contracted with E^ν. In the geometrical optics limit, taking components in the appropriate Lorentz frame, Eqn. (2.84) reproduces the usual relativistic transformation of electromagnetic fields. The dependence on the four-momentum appears because the boosted electric field depends on the magnetic field, which in the geometrical optics limit is $\hat{p} \times \vec{E}$.

The transformation law of $Q^{\mu\nu}(x,\vec{v})$ follows from Eqn. (2.62):

$$Q^{\mu'\nu'}(\vec{p},\vec{v}') = \left(\frac{\vec{v}'\cdot\vec{p}}{\vec{v}\cdot\vec{p}}\right)^2 P^{\mu'}{}_\mu(\vec{p},\vec{v}') P^{\nu'}{}_\nu(\vec{p},\vec{v}')\, Q^{\mu\nu}(\vec{p},\vec{v}) \ . \tag{2.85}$$

Since the integration measure in Eqn. (2.68) is Lorentz invariant, the transformation of $Q^{\mu\nu}$ implies that $f^{\mu\nu}$ transforms in the following way under change of the local observer 4-velocity:

$$f^{\mu'\nu'}(\vec{p},\vec{v}') = P^{\mu'}{}_\mu(\vec{p},\vec{v}') P^{\nu'}{}_\nu(\vec{p},\vec{v}') f^{\mu\nu}(\vec{p},\vec{v}) \ . \tag{2.86}$$

Note that the following transformation property of the specific intensity tensor follows immediately from the transformation properties of the distribution function and Eqn. (2.74):

$$I^{\mu'\nu'}(\vec{p},\vec{v}') = \left(\frac{\vec{v}'\cdot\vec{p}}{\vec{v}\cdot\vec{p}}\right)^3 P^{\mu'}{}_\mu(\vec{p},\vec{v}') P^{\nu'}{}_\nu(\vec{p},\vec{v}') I^{\mu\nu}(\vec{p},\vec{v}) \ . \tag{2.87}$$

In the unpolarized limit, the trace of this reduces to the familiar statement that I/ν^3 is Lorentz invariant. In the general polarized case, one sees that all of I/ν^3, Q/ν^3, U/ν^3, V/ν^3 are invariant under a boost along the photon direction.

The transformation properties of the Stokes parameters under a boost in a general direction are clearly dependent on the polarization basis chosen in each frame. To work out the general case, we consider the transformation from frame K' (the rest frame) with 4-velocity $\vec{v}_e$ into frame K (the lab frame) with 4-velocity $\vec{v}_l$. In lab frame coordinates, let $v_e^\mu = \gamma(1,\boldsymbol{v})$. In the rest frame, the brightness tensor $I^{\mu'\nu'}(\vec{p},\vec{v}_e)$ contains all polarization and intensity data of a photon with 4-momentum $\vec{p}$. We

denote the photon momentum in rest frame coordinates, as $p^{\mu'} = p'(1, \boldsymbol{n}')$, and in lab frame coordinates as $p^{\mu} = p(1, \boldsymbol{n})$.

The Stokes parameters measured in K' are defined by specifying a set of orthonormal polarization basis vectors $\{\vec{s}_1, \vec{s}_2\}$, where $\vec{s}_1 \cdot \vec{p} = \vec{s}_2 \cdot \vec{p} = 0$, and $\vec{s}_1 \cdot \vec{v}_e = \vec{s}_2 \cdot \vec{v}_e = 0$. Since the vectors $\vec{s}_a$ are purely spatial in the rest frame, we may write $s_a^{\mu'} = (0, \boldsymbol{\epsilon}'_a)$, $a \in \{1, 2\}$ with $\boldsymbol{\epsilon}'_a \cdot \boldsymbol{n}' = 0$, $\boldsymbol{\epsilon}'_a \cdot \boldsymbol{\epsilon}'_b = \delta_{ab}$. The Stokes parameters measured in K' are determined by the quantities:

$$J'_{ab} = I^{\mu'\nu'}(\vec{p}, \vec{v}_e) s_{a\mu'} s_{b\nu'} \ , \qquad (a, b) \in \{1, 2\} \ . \tag{2.88}$$

To determine the Stokes parameters measured in K, we must specify lab frame basis vectors $\{\vec{t}_1, \vec{t}_2\}$ which satisfy $\vec{t}_1 \cdot \vec{p} = \vec{t}_2 \cdot \vec{p} = 0$, and $\vec{t}_1 \cdot \vec{v}_l = \vec{t}_2 \cdot \vec{v}_l = 0$. We write $t_a^{\mu'} = (0, \boldsymbol{\epsilon}_a)$, $a \in \{1, 2\}$ with $\boldsymbol{\epsilon}_a \cdot \boldsymbol{n} = 0$. The analogous quantities to those in Eqn. (2.88) in K are

$$J_{ab} = I^{\mu\nu}(\vec{p}, \vec{v}_l) t_{a\mu} t_{b\nu} \ , \qquad (a, b) \in \{1, 2\} \ . \tag{2.89}$$

The vectors $\{\vec{t}_a\}$ are not uniquely determined, but there is a natural choice of basis which keeps the transformation of the Stokes parameters simple. Applying the transformation law (2.87) to Eqn. (2.89), and replacing

$$I^{\alpha\beta}(\vec{p}, \vec{v}_e) \to P^{\alpha}{}_{\gamma}(\vec{p}, \vec{v}_e) P^{\beta}{}_{\delta}(\vec{p}, \vec{v}_e) I^{\gamma\delta}(\vec{p}, \vec{v}_e) \ , \tag{2.90}$$

we obtain

$$J_{ab} = \left(\frac{p}{p'}\right)^3 P^{\mu\nu}{}_{\gamma\delta}(\vec{p}, \vec{v}_e, \vec{v}_l) I^{\gamma\delta}(\vec{p}, \vec{v}_e) t_{a\mu} t_{b\nu} \ . \tag{2.91}$$

where

$$P^{\mu\nu}{}_{\alpha\beta}(\vec{p}, \vec{v}_e, \vec{v}_l) \equiv P^{\mu}{}_{\alpha}(\vec{p}, \vec{v}_l) P^{\nu}{}_{\beta}(\vec{p}, \vec{v}_l) P^{\alpha}{}_{\gamma}(\vec{p}, \vec{v}_e) P^{\beta}{}_{\delta}(\vec{p}, \vec{v}_e) \ , \tag{2.92}$$

and

$$p' \equiv -\vec{v}_e \cdot \vec{p} \, , \qquad p \equiv -\vec{v}_l \cdot \vec{p} \, . \tag{2.93}$$

Comparing Eqn. (2.91) to Eqn. (2.88), it is apparent that if we demand that the vectors $\vec{t}_a$ satisfy:

$$P^{\alpha}{}_{\gamma}(\vec{p}, \vec{v}_e) P^{\mu}{}_{\alpha}(\vec{p}, \vec{v}_l) \ t_{a\mu} = s_{a\gamma} \ , \tag{2.94}$$

then the transformation law of the quantities J_{ab} reduces to

$$J_{ab} = \left(\frac{p}{p'}\right)^3 J'_{ab} \ , \tag{2.95}$$

and thus the Stokes parameter Q, for example, transforms simply as

$$Q = \left(\frac{p}{p'}\right)^3 Q' \ , \tag{2.96}$$

and similarly for the other Stokes parameters. Since $t_{a\mu}$ is assumed to be purely spatial in K ($\vec{t}_a \cdot \vec{v}_l = 0$), it follows that $P^\mu{}_\alpha(\vec{p}, \vec{v}_l)\, t_{a\mu} = t_{a\alpha}$, and the transformation (2.94) simplifies to

$$P^\alpha{}_\gamma(\vec{p}, \vec{v}_e) t_{a\alpha} = s_{a\gamma} \ . \tag{2.97}$$

In 4-vector notation, using $\vec{p} \cdot \vec{t}_a = 0$ we find

$$\vec{s}_a = \vec{t}_a + \frac{1}{p'}(\vec{v}_e \cdot \vec{t}_a)\vec{p} \tag{2.98}$$

This manifestly satisfies $\vec{s}_a \cdot \vec{v}_e = 0$. Since $\vec{t}_a$ must be purely spatial in K we have

$$\vec{t}_a \cdot \vec{v}_e = \left(\frac{p'}{p}\right) \vec{s}_a \cdot \vec{v}_l \ , \tag{2.99}$$

which yields

$$\vec{t}_a = \vec{s}_a - \frac{1}{p}(\vec{v}_l \cdot \vec{t}_a)\vec{p} \ . \tag{2.100}$$

The transformation law of the polarization basis 3-vectors $\boldsymbol{\epsilon}_a$, $\boldsymbol{\epsilon}'_a$ now follows. Denoting $\vec{s}_a$ in lab frame coordinates as $s^\mu_a = (s^0_a, \boldsymbol{s}_a)$, and Lorentz transforming $s^{\mu'}_a$ into K we obtain

$$\begin{aligned} s^0_a &= \gamma \boldsymbol{v} \cdot \boldsymbol{\epsilon}'_a \ , \\ \boldsymbol{s}_a &= \boldsymbol{\epsilon}'_a + (\gamma - 1)\frac{\boldsymbol{v}(\boldsymbol{v} \cdot \boldsymbol{\epsilon}'_a)}{v^2} \ . \end{aligned} \tag{2.101}$$

Thus the spatial part of Eqn. (2.100) yields

$$\boldsymbol{\epsilon}_a = \boldsymbol{\epsilon}'_a + (\gamma - 1)\frac{\boldsymbol{v}(\boldsymbol{v} \cdot \boldsymbol{\epsilon}'_a)}{v^2} - \gamma \boldsymbol{n}(\boldsymbol{v} \cdot \boldsymbol{\epsilon}'_a) \ . \tag{2.102}$$

This transformation law was previously obtained by Challinor & van Leeuwen (2002). One may check, using the transformation law of $\boldsymbol{n}$ (see Eqn. (3.59)), that the polarization basis 3-vectors $\boldsymbol{\epsilon}_a$ are indeed orthonormal and orthogonal to $\boldsymbol{n}$. The fact that such a complicated transformation of basis is needed to ensure that the Stokes parameters transform in a simple fashion demonstrates that the tensor approach is more convenient when dealing with relativistic transformations of polarized beams.

The *screen projection tensor* $P^{\mu\nu}(\vec{p}, \vec{v})$ defined in Eqn. (2.83) is an important tool in this polarization tensor formalism. It projects onto the "screen" subspace orthogonal to the photon momentum $\vec{p}$ and local observer 4-velocity $\vec{v}$, in the sense that it leaves $f^{\mu\nu}(\vec{p}, \vec{v})$ invariant:

$$P^\mu{}_\alpha(\vec{p}, \vec{v}) f^{\alpha\beta}(\vec{p}, \vec{v}) = f^{\mu\beta}(\vec{p}, \vec{v}) \ . \tag{2.103}$$

In a local Lorentz frame $P^{\mu\nu}$ is simply the 2×2 identity matrix in the subspace orthogonal to v^μ and p^μ. It is appropriate now to discuss some of its properties, which will be useful to refer to in later chapters. It may also be written in the form

(used in Challinor (2000) and Thorne (1981) for example)

$$P_{\mu\nu} = g_{\mu\nu} + v_\mu v_\nu - n_\mu n_\nu \ , \tag{2.104}$$

where n^μ is a spacelike unit vector giving the propagation direction of the photon with respect to the observer:

$$\vec{p} = p\,(\vec{v} + \vec{n}) \ , \quad \vec{n}\cdot\vec{v} = 0 \ , \quad \vec{n}\cdot\vec{n} = 1 \ , \quad \vec{n}\cdot\vec{p} = p \ . \tag{2.105}$$

In the rest frame of the observer with 4-velocity $\vec{v}$, the 00 and $0i$ components of $P_{\mu\nu}(\vec{p},\vec{v})$ vanish, and the spatial components are given by

$$P_{ij}(\vec{p},\vec{v}) = \delta_{ij} - n_i n_j \ . \tag{2.106}$$

where $\boldsymbol{n}$ is the photon direction 3-vector in the rest frame. By an obvious generalization of the argument leading to Eqn. (2.35), the distribution function tensor of an unpolarized beam is given by

$$f_{\mu\nu}(\vec{p},\vec{v}) = \frac{1}{2} f(\vec{p}) P_{\mu\nu}(\vec{p},\vec{v}) \ . \tag{2.107}$$

There are the simple properties:

$$\begin{aligned} v^\mu P_{\mu\nu}(\vec{p},\vec{v}) &= 0 \ , \quad p^\mu P_{\mu\nu}(\vec{p},\vec{v}) = 0 \ , \\ g_{\mu\nu}P^{\mu\nu}(\vec{p},\vec{v}) &= 2 \ , \quad P^{\mu\nu}(\vec{p},\vec{v})P_{\mu\nu}(\vec{p},\vec{v}) = 2 \ , \\ P^\mu{}_\alpha(\vec{p},\vec{v})P^\alpha{}_\nu(\vec{p},\vec{v}) &= P^\mu{}_\nu(\vec{p},\vec{v}) \ . \end{aligned} \tag{2.108}$$

Contracting two projection tensors with the same photon momentum but different observer velocities yields, with $p_1 \equiv -\vec{p}\cdot\vec{v}_1,\ p_2 \equiv -\vec{p}\cdot\vec{v}_2$:

$$P^\beta{}_\lambda(\vec{p},\vec{v}_1)P^{\gamma\lambda}(\vec{p},\vec{v}_2) = g^{\beta\gamma} + \frac{p^\gamma v_2^\beta}{p_2} + \frac{p^\beta v_1^\gamma}{p_1} + \frac{\vec{v}_1\cdot\vec{v}_2}{p_1 p_2} p^\beta p^\gamma \ . \tag{2.109}$$

Another contraction yields:

$$P^\alpha{}_\gamma(\vec{p},\vec{v}_1)P^\beta{}_\lambda(\vec{p},\vec{v}_1)P^{\gamma\lambda}(\vec{p},\vec{v}_2) = P^{\alpha\beta}(\vec{p},\vec{v}_1) \ , \tag{2.110}$$

which proves that a beam that is unpolarized according to some observer is also unpolarized according to any other observer.

We close this chapter with a demonstration of the Lorentz transformation of the polarization state of a beam using polarization matrix manipulations. This will serve as an introduction to the more complicated matrix manipulations used in the derivation of the SZE later.

We consider a photon with a general polarization state propagating in the z–direction with momentum p. We will compute the polarization matrix of the beam measured by an observer moving in the x–direction with velocity v. We work in the inertial frame K (unprimed) with basis 4-vectors $\vec{e}_i$, and observer 4-velocity $\vec{v}_0 = \vec{e}_t$

with unprimed components $v_0^\mu = (1,0,0,0)$. Let us consider a partially polarized photon beam propagating in the $\vec{e}_z$ direction, with 4-momentum $\vec{p}$ with unprimed components $p^\mu = p(1,0,0,1)$, and distribution function polarization tensor $f^{\mu\nu}(\vec{v}_0)$ as measured by observer $\vec{v}_0$. We suppress the photon 4-momentum argument of the polarization tensor since we deal here with a monochromatic beam.

We may perform tensor manipulations by defining 4×4 matrices with entries equal to tensor components, with no distinction between raised and lowered indices, provided there is a separate matrix for each combination of raised and lowered indices. Thus we define $f^{\mu\nu}(\vec{v}_0) = [\boldsymbol{f}(\vec{v}_0)]_{\mu\nu}$ where

$$\boldsymbol{f}(\vec{v}_0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & b^* & d & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \ , \tag{2.111}$$

where the row elements from left to right and the column elements from up to down refer to the (t,x,y,z) components. Choosing polarization basis vectors $\boldsymbol{\epsilon}_1 = \boldsymbol{x}$, $\boldsymbol{\epsilon}_2 = \boldsymbol{y}$, the coefficients (a,b,c,d) are related to the usual specific intensity Stokes parameters (here $h = c = 1$):

$$\begin{aligned} a &= \frac{I+Q}{2p^3}, \qquad b = \frac{U+iV}{2p^3} \ , \\ b^* &= \frac{U-iV}{2p^3}, \qquad d = \frac{I-Q}{2p^3} \ . \end{aligned} \tag{2.112}$$

Now we consider the polarization tensor measured by an observer moving perpendicular to the photon momentum in the unprimed frame, with 4-velocity $\vec{v}$. The rest frame of this observer is K'. We take 4-velocity $\vec{v}$ to have unprimed velocity components $v^\mu = \gamma(1,v,0,0)$, where $\gamma = 1/\sqrt{1-v^2}$. Then the polarization tensor measured by the observer with 4-velocity $\vec{v}$ has components in the unprimed frame as follows:

$$f^{\mu\nu}(\vec{v}) = P^\mu{}_\alpha(\vec{v},\vec{p})\, P^\nu{}_\beta(\vec{v},\vec{p}) f^{\alpha\beta}(\vec{v}_0) \ , \tag{2.113}$$

where the projection tensor $P^\mu{}_\nu$ is given by:

$$P^\mu{}_\nu(\vec{p},\vec{v}) = \eta^\mu{}_\nu + \frac{1}{p'}\left(p^\mu v_\nu + v^\mu p_\nu\right) - \frac{p^\mu p_\nu}{p'^2} \ , \tag{2.114}$$

and $p' \equiv -\vec{p}\cdot\vec{v} = \gamma p$. In matrix form $P^\mu{}_\nu = [\boldsymbol{P}_1]_{\mu\nu}$, where

$$\boldsymbol{P_1} = \begin{pmatrix} -v^2 & v & 0 & v^2 \\ -v & 1 & 0 & v \\ 0 & 0 & 1 & 0 \\ -v^2 & v & 0 & v^2 \end{pmatrix} \ . \tag{2.115}$$

The lowered index quantity $P_{\mu\nu}$ is represented by a matrix $\boldsymbol{P}_2$ with different entries:

$$\boldsymbol{P}_2 = \begin{pmatrix} v^2 & -v & 0 & -v^2 \\ -v & 1 & 0 & v \\ 0 & 0 & 1 & 0 \\ -v^2 & v & 0 & v^2 \end{pmatrix} . \tag{2.116}$$

The idempotency relation satisfied by the projection tensor, $P^{\mu}{}_{\alpha}P_{\mu\beta} = P_{\alpha\beta}$, implies the matrix relation

$$\boldsymbol{P}_1^T \boldsymbol{P}_2 \;\; = \;\; \boldsymbol{P}_2 \; , \tag{2.117}$$

which is satisfied by the matrices above. Using the projection matrices we find $f^{\mu\nu}(\vec{v}) = [\boldsymbol{f}(\vec{v})]^{\mu\nu}$ where

$$\boldsymbol{f}(\vec{v}) = \boldsymbol{P}_1 \boldsymbol{f}(\vec{v}_0) \boldsymbol{P}_1^T = \begin{pmatrix} av^2 & av & bv & av^2 \\ av & a & b & av \\ b^*v & b^* & d & b^*v \\ av^2 & av & bv & av^2 \end{pmatrix} . \tag{2.118}$$

Now we would like to obtain the Stokes parameters measured by the observer with 4-velocity $\vec{v}$. This is given by the Lorentz transformed matrix $\boldsymbol{f'} = \boldsymbol{\Lambda I \Lambda}^T$, where in this case the matrix $\boldsymbol{\Lambda}$ is a boost matrix in the x-direction:

$$\boldsymbol{\Lambda} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.119}$$

Thus the polarization matrix of the beam measured by the observer at rest in K', in primed coordinates, is

$$\boldsymbol{f'}(\vec{v}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a/\gamma^2 & b/\gamma & av/\gamma \\ 0 & b^*/\gamma & d & b^*v \\ 0 & av/\gamma & bv & av^2 \end{pmatrix} . \tag{2.120}$$

Adding the diagonal elements of this matrix yields $a + d$, since the total photon occupation number $f'^{\mu\nu}\eta_{\mu\nu} = I/p^3 = a+d$ is a Lorentz invariant quantity. The ti and it elements are zero since the electric field 4-vectors in this frame are purely spatial by definition of the tensor $f^{\mu\nu}(\vec{v})$. A matrix of this form is termed a *physical polarization matrix*, since its elements correspond to quantities measured by a polarimeter in this frame.

We now examine the Stokes parameters measured by an observer with 4-velocity $\vec{v}$ according to the polarization tensor derived, in order to check that our formalism agrees with the known transformation properties of the Stokes parameters. To obtain the Stokes parameters in the boosted frame, we need to define a photon polarization

basis. The basis given by Eqn. (2.102) should guarantee that the Stokes (divided by the cube of the momentum) are invariant under the boost. In the unprimed frame, the basis vectors were $\boldsymbol{\epsilon}_1 = \boldsymbol{x},\ \boldsymbol{\epsilon}_2 = \boldsymbol{y}$. In the lab frame (here primed), the photon momentum is $p^{\mu'} = p(\gamma, -\gamma v, 0, 1)$, the photon direction is $\boldsymbol{n'} = (-v, 0, 1/\gamma)$, and the lab frame velocity is $\boldsymbol{v'} = -v\boldsymbol{x}$. Substituting these into (2.102) yields

$$\begin{aligned} \boldsymbol{\epsilon}'_1 &= (0, 1/\gamma, 0, v)\ , \\ \boldsymbol{\epsilon}'_2 &= (0, 0, 1, 0)\ , \end{aligned} \tag{2.121}$$

which are clearly orthonormal and orthogonal to the primed photon momentum, and reduce in the limit $v \to 0$ to the unprimed basis. A more general polarization basis is obtained by rotating these vectors through an angle χ about the photon momentum, as follows:

$$\begin{aligned} \boldsymbol{\epsilon'_1} &\to \cos\chi\ \boldsymbol{\epsilon}'_1 - \sin\chi\ \boldsymbol{\epsilon'_2} = (0, \cos\chi/\gamma, -\sin\chi, v\cos\chi)\ , \\ \boldsymbol{\epsilon'_2} &\to \cos\chi\ \boldsymbol{\epsilon'_2} + \sin\chi\ \boldsymbol{\epsilon}'_1 = (0, \sin\chi/\gamma, \cos\chi, v\sin\chi)\ . \end{aligned} \tag{2.122}$$

In this rotated basis, the Stokes parameters in the boosted frame are given by the quantities

$$\begin{aligned} a' &= \boldsymbol{\epsilon}'^T_1 \cdot \boldsymbol{f'}\boldsymbol{\epsilon}'_1 = a\cos^2\chi - (b+c)\cos\chi\sin\chi + d\sin^2\chi\ , \\ b' &= \boldsymbol{\epsilon}'^T_1 \cdot \boldsymbol{f'}\boldsymbol{\epsilon'_2} = b\cos^2\chi + (a-d)\cos\chi\sin\chi - b^*\sin^2\chi\ , \\ d' &= \boldsymbol{\epsilon}'^T_2 \cdot \boldsymbol{f'}\boldsymbol{\epsilon'_2} = d\cos^2\chi + (b+c)\cos\chi\sin\chi + a\sin^2\chi\ . \end{aligned} \tag{2.123}$$

Thus the measured Stokes parameters in the primed frame are given by:

$$\begin{aligned} \frac{I'}{(p')^3} &= a' + d' = a + d = \frac{I}{p^3}\ , \\ \frac{Q'}{(p')^3} &= a' - d' = (a-d)\cos 2\chi - (b+b^*)\sin 2\chi = \frac{Q}{p^3}\cos 2\chi - \frac{U}{p^3}\sin 2\chi\ , \\ \frac{U'}{(p')^3} &= b' + b'^* = (b+b^*)\cos 2\chi + (a-d)\sin 2\chi = \frac{U}{p^3}\cos 2\chi + \frac{Q}{p^3}\sin 2\chi\ , \\ \frac{V'}{(p')^3} &= i(b' - b'^*) = i(b - b^*) = \frac{V}{p^3}\ . \end{aligned} \tag{2.124}$$

We find that with the choice $\chi = 0$, the Stokes parameters transform as claimed in Eqn. (2.96), and with a general χ the Stokes Q, U, V transform in the expected fashion under rotation of the polarization basis vectors.

We now describe the general procedure for transforming polarization tensors between frames in the case of an arbitrary boost direction, since this will be used in section §4.4 in a Monte Carlo simulation of the polarized SZ effect. Let the photon 4-momentum in frame S be $p^\mu = (\sqrt{p_i p_i}, p_x, p_y, p_z)$ and the components of the polarization tensor be given by some 4×4 matrix $\boldsymbol{f}(\vec{v}_0)$ in which none of the entries are necessarily zero. To obtain the polarization tensor measured by an observer with 4-velocity $v^\mu = \gamma(1, v_x, v_y, v_z)$ the following steps are taken:

1. Compute $p \equiv \gamma(\sqrt{p_i p_i} - p_i v_i)$.

2. Define the following row vectors corresponding to the upper and lower index 4-momenta:

$$\begin{aligned} \boldsymbol{p_1} &= (\sqrt{p_i p_i}, p_x, p_y, p_z) \ , \\ \boldsymbol{p_2} &= (-\sqrt{p_i p_i}, p_x, p_y, p_z) \ . \end{aligned} \tag{2.125}$$

3. Define the following row vectors corresponding to the upper and lower index 4-velocities:

$$\begin{aligned} \boldsymbol{v_1} &= \gamma(1, v_x, v_y, v_z) \ , \\ \boldsymbol{v_2} &= \gamma(-1, v_x, v_y, v_z) \ . \end{aligned} \tag{2.126}$$

4. Construct the 4×4 matrix $\boldsymbol{P}$ whose entries are given by:

$$[\boldsymbol{P}]_{\mu\nu} = \delta_{\mu\nu} + \frac{1}{p}\left([\boldsymbol{p_1}]_\mu [\boldsymbol{v_2}]_\nu + [\boldsymbol{p_2}]_\mu [\boldsymbol{v_1}]_\nu\right) - \frac{1}{p^2}[\boldsymbol{p_1}]_\mu [\boldsymbol{p_2}]_\nu \ . \tag{2.127}$$

5. Construct the boost matrix $\boldsymbol{\Lambda}(\boldsymbol{v})$.

6. Perform the following matrix multiplications:

$$\begin{aligned} \boldsymbol{f}(\vec{v}) &= \boldsymbol{P}\boldsymbol{f}(\vec{v}_0)\boldsymbol{P}^T \ , \\ \boldsymbol{f'}(\vec{v}) &= \boldsymbol{\Lambda}(\boldsymbol{v})\boldsymbol{f}(\vec{v})\boldsymbol{\Lambda}(\boldsymbol{v})^T \ . \end{aligned} \tag{2.128}$$

These steps are easy to implement in a computer program or computer algebra script.

Chapter 3

Kinetic theory of Compton scattering of polarized photons

The Sunyaev-Zeldovich effects (SZE) arise due to inverse Compton scattering of the CMB. The evolution of the polarization matrix of the CMB is determined by the Boltzmann (or kinetic) equation

$$\nabla_p f_{\mu\nu} = C_{\mu\nu} \ , \tag{3.1}$$

where $C_{\mu\nu}$ is the Compton scattering term. In this chapter we derive this scattering term for arbitrarily relativistic electrons and polarized photons. In fact, since the CMB photons have negligible momentum in comparison to the electron rest mass, the SZE can be calculated accurately with a simpler scattering term derived in the Thomson limit, in which the electron recoil is ignored. However, we go through the complete relativistic calculation in any case since there are other applications (not considered here) in which the recoil effect cannot be ignored.

We do however ignore the effect of induced (or stimulated) scattering, which is required for example to obtain the Kompaneets equation often used to derive the thermal SZ distortion. But the terms due to induced scattering in the Kompaneets equation are negligible in the case of cluster SZE, and in general in the unpolarized case it is known that induced scattering is a negligible effect unless electron energies are comparable to the electron rest mass. In any case a rigorous derivation of the induced effects require a quantum treatment (Nagirner & Poutanen, 2001), which we have not developed here.

The structure of this chapter is as follows. In §3.1, the classical non-relativistic physics of the generation of polarization by Thomson scattering in the electron rest frame is discussed using the polarization tensor approach. An equation for the time evolution of the distribution function polarization tensor in the electron rest frame due to Thomson scattering is derived. Then in §3.2 we derive the Boltzmann collision integral using a phenomenological approach based on the master equation of kinetic theory, still in the Thomson limit. As a check, we construct the matrix analogue of the radiative transfer equation in the case of a scattering medium composed of stationary electrons, which agrees with the results of Chandrasekhar (1960). In §3.3

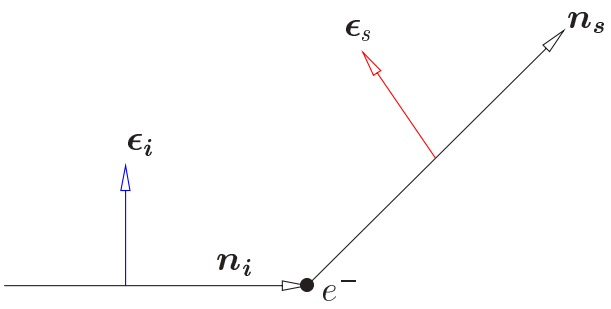


Figure 3-1: Thomson scattering of a pure incident beam from an electron at rest into a specified final polarization state.

the full relativistic kinetic equation is obtained, working in the rest frame of the initial electron – following the procedure used in the Thomson limit, but using the Klein-Nishina cross section and taking into account recoil. The transformation to a common lab frame is then taken, to obtain the kinetic equation for scattering from electrons with a general distribution of velocities. We check that this can be expressed in a manifestly covariant form. Note that throughout this chapter we work in flat spacetime.

3.1 Thomson scattering

In this section we present a derivation of the equation for the time evolution of the distribution function polarization tensor due to Thomson scattering from a distribution of stationary electrons, starting from the classical results for Thomson scattering, ignoring the effects of electron recoil and induced scattering.

Recall that for a completely linearly polarized beam, $Q^{\mu\nu}\epsilon^*_\mu\epsilon_\nu$ is the time-average energy density for electromagnetic radiation of polarization ϵ_μ, where ϵ_μ is spacelike and normalized, $\vec{\epsilon}^{\,*}\cdot\vec{\epsilon} \equiv g^{\mu\nu}\epsilon^*_\mu\epsilon_\nu = 1$. Consider a completely polarized beam with polarization vector $\vec{\epsilon}_i$ and momentum $\vec{p}_i$ incident upon an electron at rest (Fig. 3-1). The polarization matrix of the incident beam is $Q_i\vec{\epsilon}_i\otimes\vec{\epsilon}^{\,*}_i$ where cQ_i is the incident flux (we choose units such that $c = 1$). Normalization of the polarization vector implies $Q_i = Q_i^{\mu\nu}\epsilon^*_{i\,\mu}\epsilon_{i\,\nu}$. In the Thomson limit, in which the electron recoil is negligible, the differential cross section for Thomson scattering of a beam into final momentum $\vec{p}_s$ and polarization $\vec{\epsilon}_s$ is (Jackson, 1998)

$$\frac{d\sigma}{d\Omega_s} = \frac{3\sigma_{\rm T}}{8\pi}\left|\vec{\epsilon}_i^{\,*}\cdot\vec{\epsilon}_s\right|^2 \quad . \tag{3.2}$$

Thus the power per unit solid angle in the scattered beam is

$$\frac{dP_s}{d\Omega_s} = \frac{3\sigma_{\rm T}}{8\pi} Q_i \, |\vec{\epsilon}_i{}^* \cdot \vec{\epsilon}_s|^2 \tag{3.3}$$

where $d\Omega_s$ is the element of solid angle associated with the direction of $\vec{p}_s$. We may also write $Q_i|\vec{\epsilon}_i{}^* \cdot \vec{\epsilon}_s|^2 = Q_i^{\mu\nu}\epsilon^*_{s\,\mu}\epsilon_{s\,\nu}$.

Next consider a gas of electrons all at rest with number density n_e: we work in the rest frame of the electrons throughout this section. Assuming incoherent scattering, multiplying Eqn. (3.3) by $n_e d\Omega_s$ converts scattered power per electron to the rate of change of energy density in final polarization state $\vec{\epsilon}_s$:

$$\frac{dQ_s^{\mu\nu}}{dt}\,\epsilon^*_{s\,\mu}\epsilon_{s\,\nu} = \frac{3\sigma_{\rm T}}{8\pi} n_e Q_i^{\mu\nu}\epsilon^*_{s\,\mu}\epsilon_{s\,\nu}\, d\Omega_s \ . \tag{3.4}$$

Note that the time derivative d/dt here should actually be interpreted as a total derivative taken along the ray, $d/dt = \partial/\partial t + \boldsymbol{n}\cdot\boldsymbol{\nabla}$, since eventually the left hand side of the evolution equation will be replaced with the left hand side of Eqn. (3.1). Using equation (2.68), and setting $p_i = p_s$ since we are working in the Thomson limit, we may convert this to the change in the phase space density matrix, giving

$$\frac{df_s^{\mu\nu}}{dt}\,\epsilon^*_{s\,\mu}\epsilon_{s\,\nu} d\Omega_s = \frac{3\sigma_{\rm T}}{8\pi} n_e f_i^{\mu\nu}\epsilon^*_{s\,\mu}\epsilon_{s\,\nu}\, d\Omega_i d\Omega_s \ . \tag{3.5}$$

We would like an equation for the change in $f_{\mu\nu}$ due to scattering, but Eqn. (3.5) gives the change only for a particular (but arbitrary) polarization of the outgoing wave, $\vec{\epsilon}_s$. We cannot remove the polarization factors and conclude $df_s^{\mu\nu} \propto f_i^{\mu\nu}$ because the polarization of the incoming wave does not lie in the same plane as the polarization of the scattered wave. For a given outgoing momentum $\vec{p}_s$, the outgoing polarization is a linear combination of the two basis vectors $\vec{\epsilon}_1$ and $\vec{\epsilon}_2$ (associated with the photon of momentum $\vec{p}_s$) of §2.3. Thus, $f_i^{\mu\nu}\epsilon^*_{s\,\mu}\epsilon_{s\,\nu}$ projects out of the incoming density matrix $f_i^{\mu\nu}$ only those components lying in the $\vec{\epsilon}_1$-$\vec{\epsilon}_2$ plane. This projection is equivalent to first projecting $f_i^{\mu\nu}$ with $\vec{\epsilon}_1 \otimes \vec{\epsilon}_1 + \vec{\epsilon}_2 \otimes \vec{\epsilon}_2$. But this is exactly the projection tensor of Eqn. (2.83), with $\vec{p} = \vec{p}_s$ being the outgoing photon momentum and $\vec{v}$ being the electron 4-velocity. Projecting the final polarization vector with $P_{\mu\nu}$ does not change it: $P^\mu{}_\alpha \epsilon_{s\,\mu} = \epsilon_{s\,\alpha}$. It follows that $f_i^{\mu\nu}\epsilon^*_{s\,\mu}\epsilon_{s\,\nu} = f_i^{\alpha\beta}P^\mu{}_\alpha P^\nu{}_\beta \epsilon^*_{s\,\mu}\epsilon_{s\,\nu}$. Now it is safe to remove the outgoing polarization vectors from Eqn. (3.5).

We conclude that, for any initial and final polarizations,

$$\frac{df_s^{\mu\nu}(\vec{p}_s,\vec{v}_e)}{dt} = \frac{3\sigma_{\rm T}}{8\pi} n_e \, P^\mu{}_\alpha(\vec{p}_s,\vec{v}_e\,)P^\nu{}_\beta(\vec{p}_s,\vec{v}_e\,) \int d\Omega_i \, f_i^{\alpha\beta}(\vec{p}_i,\vec{v}_e) \ . \tag{3.6}$$

Eqn. (3.6) is the key result for Thomson scattering in the polarization tensor formalism. It gives the photon scattering rate per unit volume for given momenta and polarizations. The projection tensors are easy to understand: the scattered density matrix is simply proportional to the incident density matrix after the unphysical polarization components (those proportional to $\vec{v}_e$ and $\vec{n}_s$) are eliminated.

If the integration time is sufficiently short, we may replace $n_e \sigma_{\rm T}\, dt$ with the optical depth to Thomson scattering, τ. Then we have

$$f_s^{\mu\nu}(\vec{p}_s, \vec{v}_e) = \frac{3\tau}{8\pi}\, P^{\mu}{}_{\alpha}(\vec{p}_s, \vec{v}_e\,) P^{\nu}{}_{\beta}(\vec{p}_s, \vec{v}_e\,) \int d\Omega_i\, f_i^{\alpha\beta}(\vec{p}_i, \vec{v}_e) \ . \tag{3.7}$$

It follows that scattering of a photon with a given incident momentum $\vec{p}_i$ leads to a scattered beam with normalized polarization tensor

$$\phi_s^{\mu\nu}(\vec{p}_s, \vec{v}_e) = \frac{P^{\mu}{}_{\alpha}(\vec{p}_s, \vec{v}_e\,) P^{\nu}{}_{\beta}(\vec{p}_s, \vec{v}_e\,)\, \phi_i^{\alpha\beta}(\vec{p}_i, \vec{v}_e)}{P_{\gamma\delta}(\vec{p}_s, \vec{v}_e\,) \phi_i^{\gamma\delta}(\vec{p}_i, \vec{v}_e)} \ . \tag{3.8}$$

Taking the trace of the scattering rate Eqn. (3.6) yields the evolution equation for the scalar distribution functions $f_i(\vec{p}_i) = g_{\alpha\beta} f_i^{\alpha\beta}(\vec{p}_i)$, $f_s(\vec{p}_i) = g_{\alpha\beta} f_s^{\alpha\beta}(\vec{p}_s)$, which may be written in the following form (by definition of the differential scattering cross section):

$$\frac{d}{dt}(f_s\, d^3x\, d^3p_s) = (n_e\, d^3x)\, \frac{d\sigma}{d\Omega_s}\, d\Omega_s\, (f_i\, d^3p_i) \ , \tag{3.9}$$

which yields the differential scattering cross section in the rest frame (and the Thomson limit):

$$\frac{d\sigma}{d\Omega_s} = \frac{3\sigma_{\rm T}}{8\pi}\, P_{\alpha\beta}(\vec{p}_s, \vec{v}_e) \phi_i^{\alpha\beta}(\vec{p}_i, \vec{v}_e) \ . \tag{3.10}$$

In rest frame coordinate, we may deal with 3×3 matrices rather than tensors. Then we may write the normalized polarization matrix of the scattered beam in terms of that of the incident beam as:

$$\boldsymbol{\phi}_s(\boldsymbol{n}_s) = \frac{\boldsymbol{P}(\boldsymbol{n}_s) \boldsymbol{\phi}_i(\boldsymbol{n}_i) \boldsymbol{P}(\boldsymbol{n}_s)}{{\rm Tr}[\boldsymbol{P}(\boldsymbol{n}_s) \boldsymbol{\phi}_i(\boldsymbol{n}_i)]} \ . \tag{3.11}$$

The polarization magnitude of the beam reduces to a familiar form in the case of an unpolarized incident beam. For example, consider the case of an unpolarized beam incident in the z-direction in the rest frame, and let the rest frame direction vector of the scattered beam have components

$$\boldsymbol{n}_s = (\cos\varphi_s \sqrt{1-\mu_s^2}, \sin\varphi_s \sqrt{1-\mu_s^2}, \mu_s) \ . \tag{3.12}$$

The incident normalized polarization matrix is $\boldsymbol{\phi}(\boldsymbol{n}_i) = \boldsymbol{P}(\boldsymbol{n}_i)/2$. The polarization magnitude of the scattered beam is given by

$$\Pi^2(\mu_s, \varphi_s) = 2{\rm Tr}[\boldsymbol{\phi}_s^2]/{\rm Tr}[\boldsymbol{\phi}_s]^2 - 1 = \left(\frac{1-\mu_s^2}{1+\mu_s^2}\right)^2 \ , \tag{3.13}$$

which is independent of φ_s since the incident unpolarized beam picks out no preferred azimuth. In the case of an incident beam with a general polarization state, we may

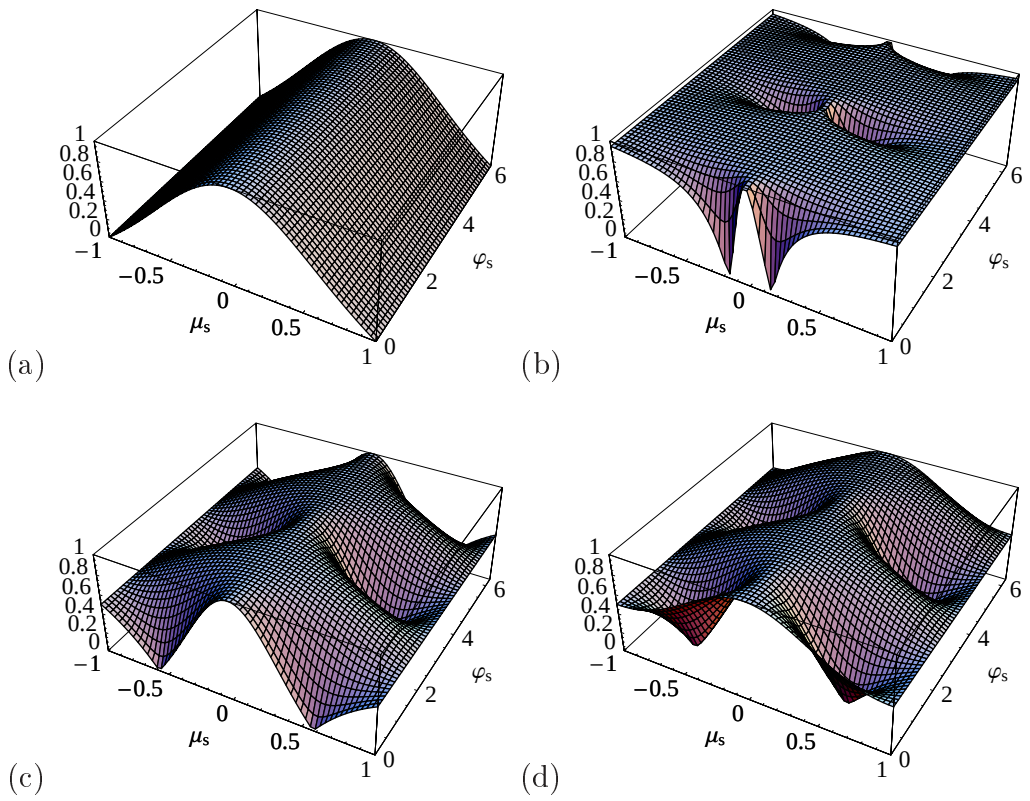


Figure 3-2: Polarization magnitude versus rest frame scattering angles (in polar coordinates). The plots for incident beams with Stokes parameters: (a) $Q = U = V = 0$, (b) $Q/I = 0.9$, $U = V = 0$, (c) $Q/I = 0.5$, $U = V = 0$, (d) $U/I = 0.5$, $Q = V = 0$.

choose polarization basis 3-vectors [1] for the incident beam $\boldsymbol{\epsilon}_1 = \boldsymbol{e}_x$, $\boldsymbol{\epsilon}_2 = \boldsymbol{e}_y$ and write, in the 2×2 polarization subspace,

$$\boldsymbol{\phi}_i = \frac{1}{2I} \begin{pmatrix} I+Q & U+iV \\ U-iV & I-Q \end{pmatrix} . \tag{3.14}$$

Then we find the scattered polarization magnitude

$$\Pi^2(\mu_s, \varphi_s) = 1 + \frac{4\,(-I^2+Q^2+U^2+V^2)\,{\mu_s}^2}{I^2[1+{\mu_s}^2+(-1+{\mu_s}^2)\,(Q\,\cos(2\,\varphi_s)+U\,\sin(2\,\varphi_s))\,/I]^2} . \tag{3.15}$$

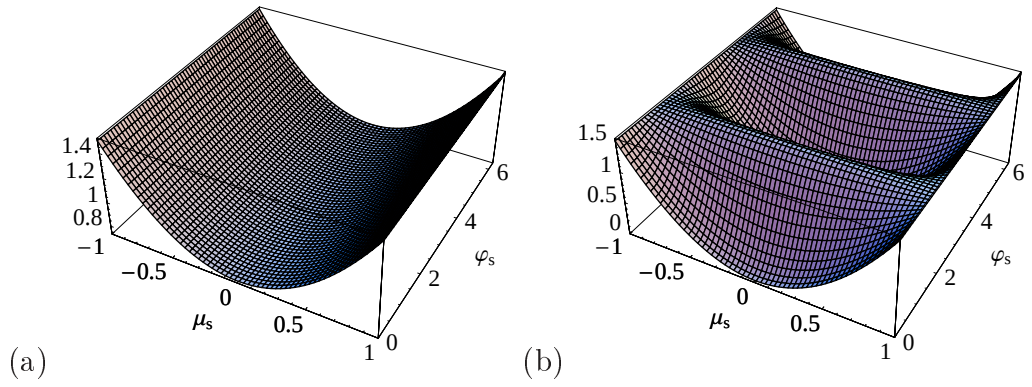
This may also be written as

$$\Pi^2(\mu_s, \varphi_s) = 1 + \frac{(-I^2+Q^2+U^2+V^2)\,{\mu_s}^2}{I^2\left[1-n_{si}n_{sj}\phi_i^{ij}\right]^2} . \tag{3.16}$$

This function is plotted for incident beams with various polarization states in Fig. 3-2.

From Eqn. (3.11) we can determine the probability for a photon to Thomson

[1] The unit 3-vectors pointing along the Cartesian coordinate axes x, y, z are denoted $\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$.



(a) (b)

Figure 3-3: Phase function Φ versus rest frame scattering angles (in polar coordinates, for a beam incident along the z-axis with Stokes parameters: (a) $Q = U = V = 0$, (b) $Q/I = 1$, $U = V = 0$.

scatter into a particular solid angle element $d\Omega_s$, which is conventionally termed the *phase function*. This is simply proportional to the differential cross section, which in matrix notation is

$$\frac{d\sigma}{d\Omega_s} = \frac{3\sigma_{\rm T}}{8\pi}\,{\rm Tr}[\boldsymbol{\phi}_i(\boldsymbol{n}_i)\boldsymbol{P}(\boldsymbol{n}_s)]\ . \tag{3.17}$$

Thus the phase function for Thomson scattering is a function of the scattered direction vector $\boldsymbol{n}_s$ and the elements of the incident polarization matrix $\boldsymbol{\phi}_i(\boldsymbol{n}_i)$ (the dependence on $\boldsymbol{n}_i$ is implicit in $\boldsymbol{\phi}_i(\boldsymbol{n}_i)$). Denoting the phase function as $\Phi[\boldsymbol{n}_s, \boldsymbol{\phi}_i(\boldsymbol{n}_i)]$, we use the normalization

$$\int \Phi[\boldsymbol{n}_s, \boldsymbol{\phi}_i(\boldsymbol{n}_i)]\,\frac{d\Omega_s}{4\pi} = 1\ . \tag{3.18}$$

Since $\int P_{ij}(\boldsymbol{n}_s)\,d\Omega_s/4\pi = 2\delta_{ij}/3$, we have

$$\Phi[\boldsymbol{n}_s, \boldsymbol{\phi}_i(\boldsymbol{n}_i)] = \frac{3}{2}\,{\rm Tr}[\boldsymbol{\phi}(\boldsymbol{n}_i)\boldsymbol{P}(\boldsymbol{n}_s)]\ . \tag{3.19}$$

For example, consider the case of an incident beam with $\boldsymbol{n}_i = (0, 0, 1)$, and intensity polarization matrix with Stokes parameters defined with respect to polarization basis vectors (associated with the incident beam) $\boldsymbol{\epsilon}_1 = \boldsymbol{e}_x$, $\boldsymbol{\epsilon}_2 = \boldsymbol{e}_y$:

$$\boldsymbol{I}(\boldsymbol{n}) = \frac{1}{2}\begin{pmatrix} I+Q & U & 0 \\ U & I-Q & 0 \\ 0 & 0 & 0 \end{pmatrix}\ . \tag{3.20}$$

and let the scattered direction have the components (3.12). Then we obtain the phase function (Code & Whitney, 1995):

$$\Phi[\boldsymbol{n}_s, \boldsymbol{\phi}_i(\boldsymbol{n}_i)] \;=\; \frac{3}{4}\left[1 + \mu_s^2 - \,(1-\mu_s^2)\,\left(Q\cos 2\phi_s + U\sin 2\phi_s\right)/I\right]\ . \tag{3.21}$$

In Fig. 3-3 this function is compared for unpolarized and completely polarized incident beams. The polarization of the incident beam destroys the azimuthal symmetry of the differential cross section and phase function.

Before proceeding further with the derivation of the Boltzmann equation, it is instructive to see how Eqn. (3.6) leads to the classical Rayleigh phase matrix (as defined by Chandrasekhar (1960)). Consider an incoming ray along $\boldsymbol{e}_z$, scattering into an outgoing ray along $\boldsymbol{e}_r(\mu_s,\varphi_s) = \sqrt{1-\mu_s^2}(\boldsymbol{e}_x\cos\varphi_s + \boldsymbol{e}_y\sin\varphi_s) + \boldsymbol{e}_z\mu_s$. For the incoming wave we choose polarization basis $\boldsymbol{\epsilon}_{i\,1} = \boldsymbol{e}_x\cos\varphi_s + \boldsymbol{e}_y\sin\varphi_s$ and $\boldsymbol{\epsilon}_{i\,2} = \boldsymbol{e}_{\hat{\varphi}_s} = -\boldsymbol{e}_x\sin\varphi_s + \boldsymbol{e}_y\cos\varphi_s$. For the scattered wave, $\boldsymbol{\epsilon}_{i\,2}$ is a good polarization vector (it is orthogonal to the momentum direction $\boldsymbol{e}_r$) but the other basis vector is changed to $\boldsymbol{\epsilon}_{s\,1} = \boldsymbol{e}_{\hat{\theta}} = \boldsymbol{e}_x\mu_s\cos\varphi_s + \boldsymbol{e}_y\mu_s\sin\varphi_s - \boldsymbol{e}_z\sqrt{1-\mu_s^2} = \boldsymbol{\epsilon}_{i\,1}\mu_s - \boldsymbol{e}_z\sqrt{1-\mu_s^2}$. In the basis $\{\boldsymbol{\epsilon}_{s\,1}, \boldsymbol{\epsilon}_{i\,2}\}$, we find

$$P^{\mu}{}_{\alpha}(\vec{p}_s,\vec{v}_e\,) = \begin{pmatrix} \mu_s^2 & 0 & -\mu_s\sqrt{1-\mu_s^2} \\ 0 & 1 & 0 \\ -\mu_s\sqrt{1-\mu_s^2} & 0 & 1-\mu_s^2 \end{pmatrix} .$$

For $f_i^{\alpha\beta}P^{\mu}{}_{\alpha}(\vec{p}_s,\vec{v}_e\,)P^{\nu}{}_{\beta}(\vec{p}_s,\vec{v}_e\,)$ this gives

$$P^{\dagger}f_iP = \frac{1}{2}\begin{pmatrix} A & B & C \\ B^* & I-Q & D \\ C & D^* & E \end{pmatrix} , \tag{3.22}$$

where $A = (I+Q)\mu_s^4$, $B = (U-iV)\mu_s^2$, $C = -(I+Q)\mu_s^3\sqrt{1-\mu_s^2}$, $D = -(U+iV)\mu_s\sqrt{1-\mu_s^2}$, $E = (I+Q)\mu_s^2(1-\mu_s^2)$. In the same basis, the scattered matrix is

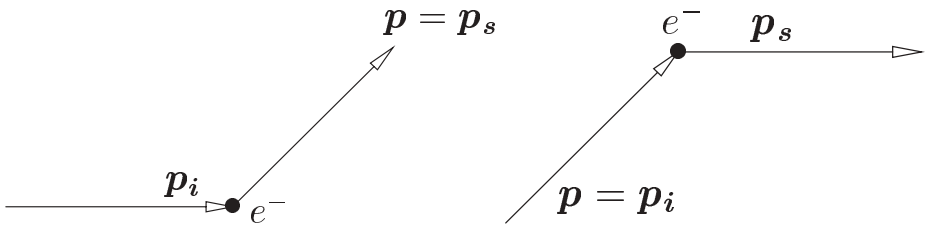
$$f_s = \frac{1}{2}\begin{pmatrix} A_s & B_s & C_s \\ B_s^* & I_s-Q_s & D_s \\ C_s & D_s^* & E_s \end{pmatrix} , \tag{3.23}$$

where $A_s = (I_s+Q_s)\mu_s^2$, $B_s = (U_s-iV_s)\mu_s$, $C_s = -(I_s+Q_s)\mu_s\sqrt{1-\mu_s^2}$, $D_s = -(U_s+iV_s)\sqrt{1-\mu_s^2}$, $E_s = (I_s+Q_s)(1-\mu_s^2)$. We find

$$\frac{d}{dt}\begin{pmatrix} I_s^{\parallel} \\ I_s^{\perp} \\ U_s \\ V_s \end{pmatrix} = \frac{3\sigma_{\rm T}}{8\pi}n_e d\Omega_i \begin{pmatrix} \mu_s^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \mu_s & 0 \\ 0 & 0 & 0 & \mu_s \end{pmatrix}\begin{pmatrix} I_i^{\parallel} \\ I_i^{\perp} \\ U_i \\ V_i \end{pmatrix} , \tag{3.24}$$

where $I^{\parallel} = I+Q$, $I^{\perp} = I-Q$. Eqn. (3.24) gives the classical Thomson (or Rayleigh) phase matrix (Chandrasekhar, 1960).

This completes our general discussion of the generation of polarization by classical Thomson scattering in the electron rest frame.



(a) Gain (b) Loss

Figure 3-4: Gain and loss processes in kinetic/master equation for Thomson scattering.

3.2 Kinetic equation in the Thomson limit

We now use the preceding results to derive the Boltzmann collision term in the electron rest frame. This is derived by the following heuristic line of reasoning. If we ignore polarization and assign a scalar distribution function $f(\boldsymbol{p})$ to each photon, the scattering rate is given by Eqn. (3.9) with the cross-section for the transition from $\boldsymbol{p}_i$ to $\boldsymbol{p}_s$ replaced by its unpolarized form, which in the Thomson limit is

$$\frac{d\sigma}{d\Omega_s} = \frac{3\sigma_{\rm T}}{16\pi}\left[1 + (\boldsymbol{n}_i \cdot \boldsymbol{n}_s)^2\right] \ . \tag{3.25}$$

We could then write the rate of change of phase space density by subtracting from Eqn. (3.9) the rate of scattering out of $\boldsymbol{p}_s$. That result is known as the *master equation* or Boltzmann equation for f (Binney & Tremaine, 1987; Groot et al., 1980):

$$\frac{d}{dt}f(\boldsymbol{p}) = n_e \int d\Omega_s \int d^3p_i \, \frac{d\sigma}{d\Omega_s}(\boldsymbol{p}_s;\boldsymbol{p}_i)\, f(\boldsymbol{p}_i)\left[\delta^3(\boldsymbol{p}_s - \boldsymbol{p}) - \delta^3(\boldsymbol{p}_i - \boldsymbol{p})\right] \ . \tag{3.26}$$

The meaning of the master equation is that the rate of change of the photon number in a given phase space element is given by summing over all scatterings into and out of this element. In this expression, $\boldsymbol{p}_s$ is not a free variable, it is a function of the incident photon momentum and scattering angles, $\boldsymbol{p}_s = \boldsymbol{p}_s(\boldsymbol{p}_i, \boldsymbol{n}_s)$, determined by the scattering kinematics. In the Thomson limit, $|\boldsymbol{p}_s| = |\boldsymbol{p}_i|$, so the scattered photon momentum is simply given by $\boldsymbol{p}_s(\boldsymbol{p}_i, \boldsymbol{n}_s) = |\boldsymbol{p}_i|\boldsymbol{n}_s$. This allows completion of the integral over the first delta function.

The first and second terms inside the square brackets correspond to scatterings into and out of the beam (with momentum $\boldsymbol{p}$) respectively, and are termed the *gain* and *loss* terms. The delta functions select the appropriate states, as indicated in Fig. 3-4. Eqn. (3.26) is simply a statement of photon number conservation combined with the rate of scattering into the final momentum state (3.9).

Now we wish to generalize this to the polarized case. The polarization tensor allows us to extend equation (3.26) to a general polarization state. and write down the kinetic equation for polarization corresponding to Eqn. (3.6). Because the transition

rate is linear in $\boldsymbol{f}_i$ and $\boldsymbol{f}_s$ it is possible to write the scattering rate for a linear superposition of initial states to a linear superposition of final states. Assuming linear superposition for incoherent light, we can write the most general incident state as $f^{\alpha\beta}(\vec{p}_i, \vec{v}_e)$ and ask for the transition rate of each element of this matrix. The transition rate is a linear transformation from $f^{\alpha\beta}(\vec{p}_i, \vec{v}_e)$ to $f^{\mu\nu}(\vec{p}_s, \vec{v}_e)$ and must therefore take the following form,

$$\frac{d}{dt}[f^{\mu\nu}(\vec{p}_s, \vec{v}_e)\, d^3p_s] = n_e\, \Phi^{\mu\nu}_{\alpha\beta}(\vec{p}_s, \vec{v}_e; \vec{p}_i, \vec{v}_e)\, f^{\alpha\beta}(\vec{p}_i, \vec{v}_e)\, d^3p_i\, d\Omega_s \ , \tag{3.27}$$

with some matrix $\Phi^{\mu\nu}_{\alpha\beta}$ that we call the *polarization scattering tensor*. It is convenient to write this as $\Phi^{\mu\nu}_{\alpha\beta}(s;i)$, where the arguments (i) and (s) are abbreviations for the pairs of 4-vectors $(\vec{p}_i, \vec{v}_e)$ and $(\vec{p}_s, \vec{v}_e)$,

The polarization scattering tensor is effectively a 4×4 matrix giving the transition rate between all possible initial and final polarization states. It follows from Eqn. (3.6) that in the Thomson limit the polarization scattering tensor is given by

$$\Phi^{\mu\nu}_{\alpha\beta}(s;i) = \frac{3\sigma_{\rm T}}{8\pi} P^{\mu}{}_{\alpha}(\vec{p}_s, \vec{v}_e) P^{\nu}{}_{\beta}(\vec{p}_s, \vec{v}_e) \ . \tag{3.28}$$

Now we make the following ansatz for polarized analogue of the master equation corresponding to Eqn. (3.26):

$$\begin{aligned}\frac{d}{dt} f^{\mu\nu}(\vec{p}, \vec{v}_e) = n_e \int d\Omega_s \int d^3p_i\, \Phi^{\alpha\beta}_{\gamma\delta}(\vec{p}_s, \vec{v}_e; \vec{p}_i, \vec{v}_e)\, f^{\gamma\delta}(\vec{p}_i, \vec{v}_e) \\ \times \left[\delta^{\mu}{}_{\alpha}\delta^{\nu}{}_{\beta}\, \delta^3(\boldsymbol{p}_s - \boldsymbol{p}) - g_{\alpha\beta}\, \phi^{\mu\nu}(\vec{p}, \vec{v}_e)\, \delta^3(\boldsymbol{p}_i - \boldsymbol{p})\right] \ .\end{aligned} \tag{3.29}$$

This is the rest frame form of the scattering term $C^{\mu\nu}$ in Eqn. (3.1). With this ansatz for the master equation, it may be checked that for any two initial and scattered *pure* states $f^{\alpha\beta}(\vec{p}_i, \vec{v}_e) = f_i \epsilon_i^{\alpha} (\epsilon_i^{\beta})^*$ and $f^{\mu\nu}(\vec{p}_s, \vec{v}_e) = f_s \epsilon_s^{\mu} (\epsilon_s^{\nu})^*$, Eqn. (3.29) reduces to Eqn. (3.26) with the Thomson cross section Eqn. (3.10). Then since any polarized beam can be written as some superposition of pure states, it follows that Eqn. (3.29) is true for all polarization states. This verifies that the ansatz (3.29) is correct in the Thomson limit.

The first term in the square brackets is exactly the gain term of Eqn. (3.6). The second term represents losses to any final polarization state; the sum over polarizations is given by $g_{\alpha\beta}$. Each incident photon beam with polarization tensor $\phi^{\mu\nu}(\vec{p}, \vec{v}_e)$ is lost by scattering implying that the loss term must be proportional to $\phi^{\mu\nu}(\vec{p}, \vec{v}_e)$. Just as the loss term in Eqn. (3.26) is proportional to the same quantity occurring on the left-hand side of the equation, the same is true here. In fact this loss term is simply the phase function multiplied by the incident beam, and is thus proportional to the probability of a photon scattering from momentum $\vec{p}_i$ to $\vec{p}_s$. To see this, note that the loss term contains the scalar obtained by contracting the projection tensors which are orthogonal to the incident and scattered photons:

$$P^{\alpha\beta}(\vec{p}_i, \vec{v}_e) P_{\alpha\beta}(\vec{p}_s, \vec{v}_e) = 2 + \frac{2}{p_i p_s}\, \vec{p}_i \cdot \vec{p}_s + \frac{(\vec{p}_i \cdot \vec{p}_s)^2}{(p_i p_s)^2} \ . \tag{3.30}$$

where $p_i \equiv -\vec{p}_i \cdot \vec{v}_e$, $p_s \equiv -\vec{p}_s \cdot \vec{v}_e$. In the rest frame of $\vec{v}_e$, $p_i^\mu = p_i(1, \boldsymbol{n}_i)$, $p_s^\mu = p_s(1, \boldsymbol{n}_s)$, giving $\vec{p}_i \cdot \vec{p}_s = -p_i p_s(1 - \boldsymbol{n}_i \cdot \boldsymbol{n}_s)$, and the loss term scalar has the form

$$P^{\alpha\beta}(\vec{p}_i, \vec{v}_e) P_{\alpha\beta}(\vec{p}_s, \vec{v}_e) = 1 + (\boldsymbol{n}_i \cdot \boldsymbol{n}_s)^2 \ , \tag{3.31}$$

which is the familiar angular dependence of the differential cross section for Thomson scattering of unpolarized radiation. The total loss term is thus simply proportional to the incident beam multiplied by the total cross section (in this case, since we have restricted to the Thomson limit, the Thomson cross section).

Note that the form of Eqn. (3.29) guarantees photon number conservation (Compton scattering cannot change the overall photon number):

$$\int d^3p \ g_{\mu\nu} \frac{df^{\mu\nu}(\vec{p}, \vec{v}_e)}{dt} = 0 \ . \tag{3.32}$$

Integration over the delta functions yields the much simplified form of the rest frame kinetic equation [2] :

$$\frac{d}{dt} f^{\mu\nu}(\vec{p}, \vec{v}_e) = n_e \sigma_{\rm T} \left[\frac{3}{2} \int \frac{d\Omega_i}{4\pi} P^\mu{}_\alpha(\vec{p}, \vec{v}_e) P^\nu{}_\beta(\vec{p}, \vec{v}_e) f^{\alpha\beta}(\vec{p}_i, \vec{v}_e) - f^{\mu\nu}(\vec{p}, \vec{v}_e) \right] \ . \tag{3.33}$$

This equation, in conjunction with the transformation to lab frame and integration over electron velocities, discussed in the next sections, is sufficient to compute all the Sunyaev-Zeldovich effects.

At this point we consider the form of Eqn. (3.33) in the case where the polarization tensors on the right hand side are taken to be unpolarized. This will be used later in the computation of the polarization generated by a single scattering of an unpolarized radiation field. In this case the integrand of the gain term in Eqn. (3.33) is proportional to the following combination of projection tensors:

$$G^{\mu\nu}(\vec{p}_i, \vec{p}_s, \vec{v}_e) \equiv P^\mu{}_\alpha(\vec{p}_s, \vec{v}_e) P^\nu{}_\beta(\vec{p}_s, \vec{v}_e) P^{\alpha\beta}(\vec{p}_i, \vec{v}_e) \ . \tag{3.34}$$

This combination can be simplified considerably. Defining $p_i \equiv -\vec{p}_i \cdot \vec{v}_e$, $p_s \equiv -\vec{p}_s \cdot \vec{v}_e$ as before, an intermediate result is:

$$\begin{aligned} P^\mu{}_\alpha(\vec{p}_s, \vec{v}_e) P^{\alpha\beta}(\vec{p}_i, \vec{v}_e) \ &= \ g^{\mu\beta} + \frac{v_e^\mu p_s^\beta}{p_s} + \frac{p_i^\mu v_e^\beta}{p_i} + \frac{p_i^\beta p_s^\mu}{p_i p_s} - \frac{p_s^\mu p_s^\beta}{p_s^2} - \frac{p_i^\mu p_i^\beta}{p_i^2} \\ &\quad + \frac{\vec{p}_i \cdot \vec{p}_s}{p_i p_s} \left[v_e^\mu v_e^\beta + \frac{p_s^\mu p_i^\beta}{p_i p_s} - \frac{p_s^\mu v_e^\beta}{p_s} - \frac{p_i^\mu v_e^\beta}{p_i} \right] \ . \end{aligned} \tag{3.35}$$

Contracting with $P^\nu{}_\beta(\vec{p}_s, \vec{v}_e)$ we find:

$$G^{\mu\nu}(\vec{p}_i, \vec{p}_s, \vec{v}_e) = P^{\mu\nu}(\vec{p}_s, \vec{v}_e) - N_\perp^\mu(\vec{p}_i, \vec{p}_s, \vec{v}_e) N_\perp^\nu(\vec{p}_i, \vec{p}_s, \vec{v}_e) \ , \tag{3.36}$$

[2] It is important to remember that on the left hand side, the derivative d/dt stands for the operator $\partial/\partial t + \boldsymbol{n} \cdot \boldsymbol{\nabla}$, where $p^\mu = p(1, \boldsymbol{n})$ in a local Lorentz frame.

where

$$\vec{N}_\perp(\vec{p}_i,\vec{p}_s,\vec{v}_e) \equiv \frac{\vec{p}_i}{p_i} - \frac{\vec{p}_s}{p_s} + \frac{\vec{p}_i\cdot\vec{p}_s}{p_i p_s}\left[\vec{v}_e - \frac{\vec{p}_s}{p_s}\right] \ . \tag{3.37}$$

Clearly

$$\vec{N}_\perp(\vec{p}_i,\vec{p}_s,\vec{v}_e)\cdot\vec{p}_s = 0 \ , \tag{3.38}$$

as required by transversality of the Thomson scattered polarization tensor to the scattered photon momentum. Dotting $\vec{N}_\perp$ into itself yields:

$$\vec{N}_\perp\cdot\vec{N}_\perp = -\frac{\vec{p}_i\cdot\vec{p}_s}{p_i p_s}\left[2 + \frac{\vec{p}_i\cdot\vec{p}_s}{p_i p_s}\right] \ . \tag{3.39}$$

We can simplify things a little by defining spacelike vectors $\vec{n}_i$ and $\vec{n}_s$ whose spatial parts are the direction 3-vectors of the photons according to the observer $\vec{v}_e$:

$$\begin{aligned}
\vec{p}_i &= p_i\,(\vec{v}_e + \vec{n}_i) \ , \quad \vec{n}_i\cdot\vec{v}_e = 0 \ , \quad \vec{n}_i\cdot\vec{n}_i = 1 \ , \quad \vec{n}_i\cdot\vec{p}_i = p_i \ , \\
\vec{p}_s &= p_s\,(\vec{v}_e + \vec{n}_s) \ , \quad \vec{n}_s\cdot\vec{v}_e = 0 \ , \quad \vec{n}_s\cdot\vec{n}_s = 1 \ , \quad \vec{n}_s\cdot\vec{p}_s = p_s \ .
\end{aligned} \tag{3.40}$$

We find

$$\vec{N}_\perp = \vec{n}_i - (\vec{n}_i\cdot\vec{n}_s)\vec{n}_s \ . \tag{3.41}$$

This vector is obviously orthogonal to $\vec{n}_s$:

$$\vec{N}_\perp\cdot\vec{n}_s = 0 \ . \tag{3.42}$$

In the rest frame of $\vec{v}_e$, where $n_i^\mu = (0,\boldsymbol{n}_i)$, $n_s^\mu = (0,\boldsymbol{n}_s)$, $\cos\theta \equiv \boldsymbol{n}_i\cdot\boldsymbol{n}_s$, we obtain:

$$\begin{aligned}
N_\perp^\mu &= (0,\boldsymbol{N}_\perp) \quad \text{where} \quad \boldsymbol{N}_\perp = \boldsymbol{n}_s\times(\boldsymbol{n}_i\times\boldsymbol{n}_s) \ , \\
\vec{N}_\perp\cdot\vec{N}_\perp &= -(\boldsymbol{n}_i\cdot\boldsymbol{n}_s)\,[2 + \boldsymbol{n}_i\cdot\boldsymbol{n}_s] = \sin^2\theta \ .
\end{aligned} \tag{3.43}$$

We find from Eqns. (3.36) and (3.41) that in the rest frame, the spatial components of the integrand of the gain term are given by

$$\begin{aligned}
G^{ij}(\vec{p}_i,\vec{p}_s,\vec{v}_e) &= \left(\delta^i_k - n^i_s n_{s,k}\right)\left(\delta^j_l - n^j_s n_{s,l}\right)\left(\delta^{kl} - n^k_i n^l_i\right) \\
&= (\delta^{ij} - n^i_i n^j_i) - n^i_s n^j_s[1 + (\boldsymbol{n}_s\cdot\boldsymbol{n}_i)^2] + (n^i_s n^j_s + n^j_s n^i_s)(\boldsymbol{n}_s\cdot\boldsymbol{n}_i) \ .
\end{aligned} \tag{3.44}$$

This expression is very convenient for the calculation of the polarization generated by scattering of the CMB quadrupole, in §5.3. In that section we use Eqn. (3.44) to prove that only the quadrupole component (in a spherical harmonic expansion) of the intensity of the unpolarized incident beam leads to generation of polarization, which is not obvious from the general form of the kinetic equation (3.33).

We close this section with a demonstration that the rest frame form of the kinetic equation, Eqn. (3.33), yields the well known results of Chandrasekhar (1960) for the polarized radiative transfer equations in the case of Thomson scattering from cold (i.e. stationary) electrons in a slab geometry. Since Chandrasekhar used a different formalism based on transformations of the Stokes parameters to derive his expressions, this is an important check of the formalism we have developed.

Consider a plane parallel atmosphere of cold electrons with uniform density filling the half-space $z < 0$, illuminated by an unpolarized beam of monochromatic radiation incident along the normal to the plane (we do not specify the boundary conditions here, since we are only interested in deriving the form of the transfer equations and not their solution). By symmetry the radiation field in this case is clearly axisymmetric. We work entirely in the electron rest frame, where $I^{00} = I^{0i} = 0$, and following Chandrasekhar (1960) define τ to be the optical depth from the surface $z = 0$ along the downward normal. The Cartesian coordinates in the plane are x and y, and The polar and azimuthal angles about the z-axis are denoted θ and φ (and $\mu = \cos\theta$).

Then the kinetic equation for the lab frame intensity polarization matrix $\boldsymbol{I}(\mu, \varphi, \tau)$ is given by Eqn. (3.33) (replacing distribution function tensors by intensity tensors, which is trivially allowed here since the beam is monochromatic):

$$\mu\, \partial_\tau \boldsymbol{I}(\mu, \varphi, \tau) \;=\; -\frac{3}{2}\boldsymbol{P}(\boldsymbol{n}) \int_{-1}^{1} d\mu_s \int_0^{2\pi} \frac{d\varphi_s}{4\pi} \boldsymbol{I}(\mu_s, \varphi_s, \tau)\boldsymbol{P}^{\boldsymbol{T}}(\boldsymbol{n}) + \boldsymbol{I}(\mu, \varphi, \tau)\ . \tag{3.45}$$

where $\boldsymbol{n}(\theta, \varphi) = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta)$ is the direction 3-vector of the beam. Note that in this expression, following Chandrasekhar (1960), the gain and loss terms have picked up a minus sign since the direction of increasing optical depth (along the normal to the boundary $z = 0$ directed into the half-space $z < 0$) is defined to be opposite to the polar axis.

In contrast to the Stokes vector approach of Chandrasekhar (1960), the polarization matrix even in this azimuthally symmetric problem has azimuthal dependence. But we are free to exploit the azimuthal symmetry here by choosing a convenient azimuth to evaluate the projection factors, and then the result obtained for this azimuth may be transformed into the other directions trivially. Choosing $\varphi = 0$, we have $(n^x, n^y, n^z) = (\sqrt{1-\mu^2}, 0, \mu)$. In the $\varphi = 0$ direction, I^{ij} has the form (since the yx and yz cross-terms must vanish in order that the place of polarization is parallel or perpendicular to the $y - z$ plane as required by axisymmetry).

$$\boldsymbol{I}(\mu, 0, \tau) = \begin{pmatrix} I_l\mu^2 & 0 & I_l\mu\sqrt{1-\mu^2} \\ 0 & I_r & 0 \\ I_l\mu\sqrt{1-\mu^2} & 0 & I_l(1-\mu^2) \end{pmatrix}\ . \tag{3.46}$$

Here $I_l(\mu, \tau)$ and $I_r(\mu, \tau)$ are the azimuth independent Stokes parameters, parallel and perpendicular respectively to the meridian plane, as defined in Chandrasekhar (1960). The matrices inside the $d\varphi_s$ integrals range over all values of φ_s though, so an expression for $I^{kl}(\mu_s, \varphi_s, \tau)$ is required. By azimuthal symmetry, this is simply given by rotating $I^{kl}(\mu_s, 0, \tau)$ through an angle φ_s about the $\hat{z}$ axis (since under rotation polarization matrices transform according to the vector rotation of the electric field strength vectors):

$$\boldsymbol{I}(\mu_s, \varphi_s, \tau) = \boldsymbol{R}(-\varphi_s\hat{z})\boldsymbol{I}(\mu_s, 0, \tau)\boldsymbol{R}^{\boldsymbol{T}}(\varphi_s\hat{z})\ . \tag{3.47}$$

where $\boldsymbol{R}(\varphi_s\hat{z})$ is the matrix which rotates through angle φ_s about the z-axis. Performing the matrix multiplication, and integrating over the azimuth yields

$$\mu\ \partial_\tau I_l(\mu,\tau) = I_l(\mu,\tau) \\ -\frac{3}{8}\int_{-1}^{1} d\mu_s\ \left\{ I_l(\mu_s,\tau)\left[2(1-\mu_s^2)(1-\mu^2)+\mu^2\mu_s^2\right] + I_r(\mu_s,\tau)\mu^2 \right\} . \quad (3.48)$$

The equation for I_r is obtained similarly:

$$\mu\ \partial_\tau I_r(\mu,\tau) \ = \ I_r(\mu,\tau) - \frac{3}{8}\int_{-1}^{1} d\mu_s \left\{ I_l(\mu_s,\tau)\mu_s^2 + I_r(\mu_s,\tau) \right\} . \quad (3.49)$$

These are the coupled radiative transfer equations for the Stokes parameters for Thomson scattering in slab geometry, first obtained by Chandrasekhar (1960).

3.3 Klein-Nishina scattering

Up to now we have worked in the Thomson limit. In this section we extend to the general case of Compton scattering with the full Klein-Nishina form of the scattering cross section, and taking into account the electron recoil.

For now we work still in the rest frame of the electron before scattering (the "initial" electron). To derive the polarization tensor kinetic equation our starting point is the Klein-Nishina differential cross-section (Greiner & Reinhardt, 1994), in the initial electron rest frame, for photons with 3-momentum $\boldsymbol{p}_i$ and polarization $\boldsymbol{\epsilon}_i$ to scatter into 3-momentum $\boldsymbol{p}_s$ and polarization $\boldsymbol{\epsilon}_s$, generalized to allow for arbitrary elliptical polarization [3] (Stedman & Pooke, 1982):

$$\frac{d\sigma}{d\Omega_s} = \frac{3\sigma_{\rm T}}{8\pi}\left(\frac{p_s}{p_i}\right)^2 \left[|\boldsymbol{\epsilon}_i\cdot\boldsymbol{\epsilon}_s^*|^2 + \frac{1}{4}\left(\frac{p_s}{p_i}+\frac{p_i}{p_s}-2\right)\left(1+|\boldsymbol{\epsilon}_i\cdot\boldsymbol{\epsilon}_s^*|^2-|\boldsymbol{\epsilon}_i\cdot\boldsymbol{\epsilon}_s|^2\right)\right] . \quad (3.50)$$

where $p_i \equiv E(\boldsymbol{p}_i) = |\boldsymbol{p}_i|$ and $p_s \equiv E(\boldsymbol{p}_s) = |\boldsymbol{p}_i|$. We allow the polarization vector to be complex in order to treat elliptical polarization; the polarization vectors are normalized by $\boldsymbol{\epsilon}_i\cdot\boldsymbol{\epsilon}_i^* = \boldsymbol{\epsilon}_s\cdot\boldsymbol{\epsilon}_s^* = 1$. The factor $(1+|\boldsymbol{\epsilon}_i\cdot\boldsymbol{\epsilon}_s^*|^2-|\boldsymbol{\epsilon}_i\cdot\boldsymbol{\epsilon}_s|^2)$ is usually not given as it reduces to unity for linearly polarized light, but we allow for light of arbitrary polarization. Eqn. (3.50) assumes the transverse gauge condition $\boldsymbol{\epsilon}_i\cdot\boldsymbol{p}_i = \boldsymbol{\epsilon}_s\cdot\boldsymbol{p}_s = 0$ and that the time component of both polarization 4-vectors vanishes in the initial electron rest frame. The factor $(p_s/p_i)^2$ in the cross section is a phase space factor.

Conservation of energy-momentum relates the initial and final momenta and scattering angle:

$$\frac{p_s}{p_i} = \left[1+\frac{p_i}{m_e}(1-\boldsymbol{n}_i\cdot\boldsymbol{n}_s)\right]^{-1} = 1-\frac{p_s}{m_e}(1-\boldsymbol{n}_i\cdot\boldsymbol{n}_s) , \quad (3.51)$$

[3] This has recently been verified by A. Guth, by performing the explicit QED computation with complex polarization vectors

where $\boldsymbol{n}_i$ and $\boldsymbol{n}_s$ are unit three-vectors along the spatial parts of the photon momenta $\boldsymbol{p}_i = p_i \boldsymbol{n}_i$ and $\boldsymbol{p}_s = p_s \boldsymbol{n}_s$. Note that for fixed directions, $dp_s/dp_i = (p_s/p_i)^2$.

Equations (3.50) and (3.51) both assume that all quantities are given in the rest frame of the incident electron. We note, incidentally, that the Klein-Nishina formula should be symmetric under the interchange of the initial and final states, but the gauge condition which was imposed to derive this form of the cross section required that the incident and scattered photon polarization basis 4-vectors be orthogonal to the *incident* electron, and it thus appears rather asymmetric. However, it is true that there is nothing special about this choice of gauge. In Appendix A, it is demonstrated that the cross section is manifestly symmetric under interchange of the initial and final states.

The 4-velocity of the incident electron will from now on be denoted $\vec{v}_i$, rather than $\vec{v}_e$ (since now we have to distinguish between the incident and scattered electron momenta). Now following the procedure in the previous section, we write the transition rate of the scattered polarization matrix $f^{\mu\nu}(\vec{p}_s, \vec{v}_i)$ as a linear transformation of the incident matrix $f^{\alpha\beta}(\vec{p}_i, \vec{v}_i)$, as in Eqns. (3.27) and (3.28), except we pull the phase space factor out for convenience:

$$\frac{d}{dt}[f^{\mu\nu}(\vec{p}_s, \vec{v}_i)\, d^3p_s] = n_e\, \Phi^{\mu\nu}_{\alpha\beta}(\vec{p}_s, \vec{v}_i; \vec{p}_i, \vec{v}_i)\, f^{\alpha\beta}(\vec{p}_i, \vec{v}_i) \left(\frac{p_s}{p_i}\right)^2 d^3p_i\, d\Omega_s \ , \tag{3.52}$$

The scattering polarization tensor $\Phi^{\mu\nu}_{\alpha\beta}$ for the Klein-Nishina cross-section is then given by the modified form:

$$\begin{aligned} \Phi^{\mu\nu}_{\alpha\beta}(s;i) \ &= \ \frac{3\sigma_{\rm T}}{8\pi}\Bigg\{ P^{\mu\nu}_{\alpha\beta}(s) + \frac{1}{4}\left(\frac{p_s}{p_i} + \frac{p_i}{p_s} - 2\right) \\ &\qquad \times \left[P^{\mu\nu}(s) P_{\alpha\beta}(i) + P^{\mu\nu}_{\alpha\beta}(s) - P^{\mu\nu}_{\beta\alpha}(s)\right]\Bigg\} \ , \end{aligned} \tag{3.53}$$

where

$$P^{\mu\nu}_{\alpha\beta}(s) \equiv P^{\mu}{}_{\alpha}(s) P^{\nu}{}_{\beta}(s) \ , \tag{3.54}$$

and the arguments (i) and (s) are abbreviations for the pairs of 4-vectors $(\vec{p}_i, \vec{v}_i)$ and $(\vec{p}_s, \vec{v}_i)$, and $p_s = -\vec{v}_i \cdot \vec{p}_s$ and $p_i = -\vec{v}_i \cdot \vec{p}_i$. As a check, if we consider an unpolarized incident beam with $\phi^{\alpha\beta}(i) = \frac{1}{2} P^{\alpha\beta}(i)$ and sum over final polarizations, we get the usual spin-summed result for the Klein-Nishina cross section:

$$\frac{1}{2} P_{\mu\nu}(\vec{p}_s, \vec{v}_i) \Phi^{\mu\nu}_{\alpha\beta}(\vec{p}_s, \vec{v}_i; \vec{p}_i, \vec{v}_i) P^{\alpha\beta}(\vec{p}_i, \vec{v}_i) = \frac{3\sigma_{\rm T}}{16\pi}\left(\frac{p_s}{p_i} + \frac{p_i}{p_s} - \sin^2\theta\right) \ , \tag{3.55}$$

where θ is the scattering angle, $\cos\theta = \boldsymbol{n}_i \cdot \boldsymbol{n}_s$.

Now, as in the derivation in the Thomson limit case, we make an ansatz for the

rest frame kinetic equation:

$$\frac{d}{dt} f^{\mu\nu}(\vec{p},\vec{v}_i) = n_e \int d\Omega_s \int d^3p_i \left(\frac{p_s}{p_i}\right)^2 \Phi^{\alpha\beta}_{\gamma\delta}(\vec{p}_s,\vec{v}_i;\vec{p}_i,\vec{v}_i)\, f^{\gamma\delta}(\vec{p}_i,\vec{v}_i) \\ \times \left[\delta^{\mu}{}_{\alpha}\delta^{\nu}{}_{\beta}\,\delta^3(\boldsymbol{p}_s-\boldsymbol{p}) - g_{\alpha\beta}\,\phi^{\mu\nu}(\vec{p},\vec{v}_i)\,\delta^3(\boldsymbol{p}_i-\boldsymbol{p})\right] \ . \tag{3.56}$$

That this expression is correct can be verified as in the Thomson limit case by substituting the matrices of pure states and checking that Eqn. (3.26) is regained with the the Klein-Nishina cross section (3.50). It is important to understand that, as in Eqn. (3.26), the scattered electron 3-momentum $\boldsymbol{p}_s$ in the expression above is *not* a free variable — it is determined by the scattering kinematics as $\boldsymbol{p}_s = p_s\boldsymbol{n}_s$, where p_s is given as a function of $\boldsymbol{p}_i$ and $\boldsymbol{n}_s$ by Eqn. (3.51).

Having obtained the equation for polarized radiation transfer in the rest frame of the scattering electron, we now consider the general case of scattering from a distribution of electrons with varying velocity. To obtain this, first it is necessary to transform the kinetic equation to a common lab frame. Since we are now dealing with two coordinate frames, for clarity the coordinates used to describe the scattering are shown schematically on the Feynman diagrams representing the Compton scattering process in Figs. 3-5, 3-6, and 3-7. We take $c = 1$ throughout. Henceforth, components of 4-vectors in the rest frame of the initial electron are denoted with primes, and those in the lab frame without primes. The incident particles are denoted with a subscript i, and the scattered particle with a subscript s. Photon momenta are denoted by p, and electron momenta by q. As usual, quantities with arrows are 4-vectors, and boldface quantities are 3-vectors. The 4-momenta satisfy the mass shell conditions $\vec{p}_i\cdot\vec{p}_i = \vec{p}_s\cdot\vec{p}_s = 0$, $\vec{q}_i\cdot\vec{q}_i = \vec{q}_s\cdot\vec{q}_s = -m_e^2$. The electron energies are given by $E(\boldsymbol{q}) = \sqrt{m_e^2+|\boldsymbol{q}|^2}$.

In lab frame, the initial and scattering electron 4-momenta are written in terms of the lab frame electron 3-velocities as follows:

$$\begin{aligned} \vec{q}_i &= m_e\vec{v}_i = \gamma_i m_e(1,\boldsymbol{v}_i) \ , \quad \gamma_i \equiv \frac{1}{\sqrt{1-v_i^2}} \\ \vec{q}_s &= m_e\vec{v}_s = \gamma_s m_e(1,\boldsymbol{v}_s) \ , \quad \gamma_s \equiv \frac{1}{\sqrt{1-v_s^2}} \ . \end{aligned} \tag{3.57}$$

The Lorentz transformation into the rest frame of the initial electron is given by the matrix:

$$\Lambda^0_0 = \gamma_i, \quad \Lambda^0_i = -\gamma_i\,[\boldsymbol{v}_i]_i \quad \Lambda^i_j = \delta^i_j + (\gamma_i-1)\frac{[\boldsymbol{v}_i]_i\,[\boldsymbol{v}_i]_j}{v_i^2} \ . \tag{3.58}$$

This yields the transformation of the photon direction vector between frames:

$$\boldsymbol{n}_i' = [\gamma_i(1-\boldsymbol{n}_i\cdot\boldsymbol{v}_i)]^{-1}\left[\boldsymbol{n}_i + \frac{\gamma_i^2}{\gamma_i^2+1}\boldsymbol{v}_i(\boldsymbol{n}_i\cdot\boldsymbol{v}_i) - \gamma_i\boldsymbol{v}_i\right] \ . \tag{3.59}$$

The relationship between the rest and lab frame momenta of the incident and scat-

Figure 3-5: Scattering process in four-vector notation

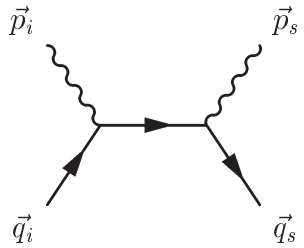


Figure 3-6: Scattering process in rest frame coordinates

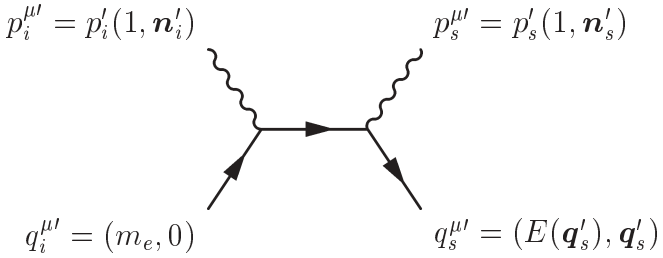
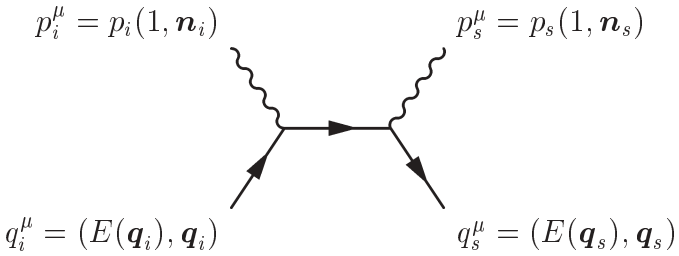


Figure 3-7: Scattering process in lab frame coordinates



tered photons is

$$\begin{aligned} p'_s &= -\vec{p}_s \cdot \vec{q}_i/m_e = \frac{1}{m_e}\left[p_s q_i - \boldsymbol{p}_s \cdot \boldsymbol{q}_i\right] \ , \\ p'_i &= -\vec{p}_i \cdot \vec{q}_i/m_e = \frac{1}{m_e}\left[p_i q_i - \boldsymbol{p}_i \cdot \boldsymbol{q}_i\right] \ . \end{aligned} \tag{3.60}$$

Or in terms of the γ_i factor,

$$\begin{aligned} \frac{p'_i}{p_i} &= \gamma_i \left(1 - \boldsymbol{n}_i \cdot \boldsymbol{v}_i\right) \ , \\ \frac{p'_s}{p_s} &= \gamma_i \left(1 - \boldsymbol{n}_s \cdot \boldsymbol{v}_i\right) \ . \end{aligned} \tag{3.61}$$

The scattered and incident photon energies in the rest frame are related by the familiar Compton scattering formula:

$$\frac{p'_s}{p'_i} = \frac{1}{1 + (p'_i/m_e)\left[1 - \boldsymbol{n}'_i \cdot \boldsymbol{n}'_s\right]} \ . \tag{3.62}$$

The lab frame version of this is

$$p_s = \frac{p_i(1 - \boldsymbol{n}_i \cdot \boldsymbol{v}_i)}{1 - \boldsymbol{n}_s \cdot \boldsymbol{v}_i + (p_i/\gamma_i m_e)\left(1 - \boldsymbol{n}_i \cdot \boldsymbol{n}_s\right)} \ . \tag{3.63}$$

We are now in a position to Lorentz transform Eqn. (3.56) to a lab frame in which the electrons have 3-velocity $\boldsymbol{v}_i$. Four quantities need to be transformed: n'_e, dt', the 3-vector $\boldsymbol{p}'$ of the beam on the left hand side of the master equation, and $f^{\mu\nu}(\vec{p}, \vec{v}_i)$ itself because of its dependence on $\vec{v}_i$. Let the four-vector $\vec{p}$ have spatial components $p\boldsymbol{n}$ in the lab frame and $p'\boldsymbol{n}'$ in the electron rest frame. The transformation laws of p and $\boldsymbol{n}$ have already been derived:

$$\begin{aligned} p' &= \gamma_i p(1 - \boldsymbol{n} \cdot \boldsymbol{v}_i) \ , \\ \boldsymbol{n}' &= \left[\gamma_i(1 - \boldsymbol{n} \cdot \boldsymbol{v}_i)\right]^{-1} \left[\boldsymbol{n} + \frac{\gamma_i^2}{\gamma_i^2 + 1}\boldsymbol{v}_i(\boldsymbol{n} \cdot \boldsymbol{v}_i) - \gamma_i \boldsymbol{v}_i\right] \ . \end{aligned} \tag{3.64}$$

The Lorentz transformation of the electron density is

$$n'_e = \gamma_i^{-1} n_e \ . \tag{3.65}$$

This is simply due to the Lorentz contraction of the volume element.

The Lorentz transformation of the time element is more subtle; dt transforms like p, not like n_e:

$$dt' = \gamma_i(1 - \boldsymbol{n} \cdot \boldsymbol{v}_i)\, dt \ . \tag{3.66}$$

From a mathematical point of view, this is because the transport operator on the left-hand side of the Boltzmann equation is actually the directional derivative $d/d\lambda \equiv$

$(dx^\mu/d\lambda)\partial/\partial x^\mu + (dp^\mu/d\lambda)\partial/\partial p^\mu = p^\mu \partial/\partial x^\mu$ (in flat space). In a local Lorentz frame, $d/d\lambda = pd/dt$ where $d/dt \equiv \partial/\partial t + \boldsymbol{n}\cdot\boldsymbol{\nabla}$. Physically, the transformation of dt arises due to the enhancement of the rate of scattering of photons from electrons which are approaching compared to that from electrons which are receding, which is due to the dependence of the flux of photons incident on the electrons on their relative velocity. Thus we refer to the factor $\gamma_i(1-\boldsymbol{n}\cdot\boldsymbol{v}_i)$ which appears in Eqn. (3.66) as the *flux factor*. This factor is crucial in the derivation of the SZ effects!

It follows from the transformation of the time element and the invariance of the trace of the distribution function tensor, that the left-hand sides of the Boltzmann equation in rest and lab frame are related by [4]

$$\frac{df'(\vec{p})}{dt'} = \frac{1}{\gamma_i(1-\boldsymbol{n}\cdot\boldsymbol{v}_i)}\frac{df(\vec{p})}{dt} \ , \tag{3.67}$$

or alternatively

$$\frac{df(\vec{p})}{dt} = \frac{1}{\gamma_i(1+\boldsymbol{n}'\cdot\boldsymbol{v}_i)}\frac{df'(\vec{p})}{dt'} \ . \tag{3.68}$$

(The photon occupation numbers in rest and lab frames also satisfy these equations, of course).

We now have all the ingredients needed to transform equation (3.56) to the rest frame of an observer with 4-velocity $\vec{v}_l$ (the subscript standing for "lab"). The transformation of the polarization tensors from lab to rest frame was derived in §2.4:

$$f^{\mu'\nu'}(\vec{p},\vec{v}_i) = P^{\mu'}{}_{\mu}(\vec{p},\vec{v}_i)P^{\nu'}{}_{\nu}(\vec{p},\vec{v}_i)f^{\mu\nu}(\vec{p},\vec{v}_l) \ . \tag{3.69}$$

It is convenient to insert the tensors which project into the electron rest frame into the scattering tensor, by redefining the tensor $P^{\mu\nu}_{\alpha\beta}(s)$ which appears in Eqn. (3.53) as

$$P^{\mu\nu}_{\alpha\beta}(s;i) \equiv P^{\mu}{}_{\gamma}(s)P^{\nu}{}_{\delta}(s)P^{\gamma}{}_{\alpha}(i)P^{\delta}{}_{\beta}(i) \ . \tag{3.70}$$

where the arguments (i) and (s) are abbreviations for the pairs of 4-vectors $(\vec{p}_i,\vec{v}_i)$ and $(\vec{p}_s,\vec{v}_i)$, and $p_s = -\vec{v}_i\cdot\vec{p}_s$ and $p_i = -\vec{v}_i\cdot\vec{p}_i$.

Finally, we can generalize the electron density to a distribution of electrons, $n_e = \int d^3q_i\, g_e(\boldsymbol{q}_i)$ where $\boldsymbol{q}_i$ is the electron 3-momentum, and g_e is the scalar phase space distribution function for the electrons. Putting everything together, we obtain

$$\begin{aligned} p\frac{d}{dt}f^{\kappa\lambda}(\vec{p},\vec{v}_l) = P^{\kappa}{}_{\mu}(\vec{p},\vec{v}_l)P^{\lambda}{}_{\nu}(\vec{p},\vec{v}_l)\int\frac{d^3q_i}{E(\boldsymbol{q}_i)}\,g_e(\boldsymbol{q}_i)\,m_e p' \\ \times\int d\Omega'_s\int d^3p'_i\left(\frac{p'_s}{p'_i}\right)^2\Phi^{\alpha\beta}_{\gamma\delta}(\vec{p}_s,\vec{v}_i;\vec{p}_i,\vec{v}_i)\,f^{\gamma\delta}(\vec{p}_i,\vec{v}_l) \\ \times\left[\delta^{\mu}{}_{\alpha}\delta^{\nu}{}_{\beta}\,\delta^3(\boldsymbol{p}'_s-\boldsymbol{p}') - g_{\alpha\beta}\,\phi^{\mu\nu}(\vec{p}_i,\vec{v}_l)\,\delta^3(\boldsymbol{p}'_i-\boldsymbol{p}')\right] \ , \end{aligned} \tag{3.71}$$

where $E(\boldsymbol{q}_i) = \gamma_i m_e$. Primes denote components in the rest frame of $\vec{v}_i = \vec{q}_i/m_e$, e.g. $p' \equiv -\vec{v}_i\cdot\vec{p}$. The flux factor is contained in the p' factor in the first line on the

[4] Recall that $\vec{p}$ is a Lorentz covariant 4-vector in these expressions.

right hand side (the Lorentz invariant measure $d^3q_i/E(\boldsymbol{q}_i)$ is pulled out to facilitate the derivation of the covariant form to follow). Note that the projection operators which project $f^{\gamma\delta}$ into the rest frame of $\vec{v}_i$ are already present in $\Phi^{\alpha\beta}_{\gamma\delta}$. Similarly, it does not matter whether $\phi^{\mu\nu}$ is evaluated in the rest frame of $\vec{v}_i$ or $\vec{v}_l$, because of the projection operators in front of the integral. It follows that we may, without loss of generality, drop the 4-velocity argument from $f^{\mu\nu}$ and $\phi^{\mu\nu}$ provided it is understood that the final results must always be projected into the physical polarization space of the observer.

Eqn. (3.71) looks complicated and is not the most convenient form for computation. The delta functions can be integrated over resulting in a simpler expression. This requires the Jacobian of $(\boldsymbol{p}'_i, \boldsymbol{n}'_s)$ and $(\boldsymbol{p}'_s, \boldsymbol{n}'_i)$, which follows from Eqns. (3.62) and (3.64), yielding [5] :

$$d\Omega'_s\, d^3p'_i \left(\frac{p'_s}{p'_i}\right)^2 = d\Omega'_i\, d^3p'_s \left(\frac{p'_i}{p'_s}\right)^2 \quad . \tag{3.72}$$

The Boltzmann equation now becomes

$$\begin{aligned}\frac{d}{dt}f^{\mu\nu}(\vec{p}) = &\int d^3q_i\, g_e(\boldsymbol{q}_i)\,(1-\boldsymbol{n}\cdot\boldsymbol{v}_i) \\ &\times\left[\int d\Omega'_i \left(\frac{p'_i}{p'}\right)^2 \Phi^{\mu\nu}_{\alpha\beta}(\vec{p},\vec{v}_i;\vec{p}_i,\vec{v}_i)\, f^{\alpha\beta}(\vec{p}_i)\right. \\ &\left. - \int d\Omega'_s \left(\frac{p'_s}{p'}\right)^2 \phi^{\mu\nu}(\vec{p}) g_{\alpha\beta}\Phi^{\alpha\beta}_{\gamma\delta}(\vec{p}_s,\vec{v}_i;\vec{p},\vec{v}_i) f^{\gamma\delta}(\vec{p})\right] .\end{aligned} \tag{3.73}$$

This form is convenient for both analytic and Monte Carlo calculations. The flux factor is explicit in the integration over the electron momenta. (Recall that $f^{\mu\nu}$ must be projected into the observer frame at the end of the calculation.)

The covariant kinetic equation for Compton scattering was derived for unpolarized photons by Dodelson & Jubas (1995). Their expression for the time evolution of the photon phase space distribution function has the form (Eqn. (2.3) of Dodelson & Jubas (1995)):

$$\begin{aligned}p_1\frac{d}{dt}f(\vec{p}_1) = &\int \frac{d^3q_1}{E(\boldsymbol{q}_1)}\int \frac{d^3q_2}{E(\boldsymbol{q}_2)}\int \frac{d^3p_2}{E(\boldsymbol{p}_2)}\,|M|^2\,\delta^4(\vec{p}_1+\vec{q}_1-\vec{p}_2-\vec{q}_2) \\ &\left[f(\vec{p}_2)(1+f(\vec{p}_1))g_e(\boldsymbol{q}_2) - f(\vec{p}_1)(1+f(\vec{p}_2))g_e(\boldsymbol{q}_1)\right].\end{aligned} \tag{3.74}$$

where $|M|^2$ is the squared matrix element for unpolarized Compton scattering. Note that this contains stimulated emission factors, which we ignore in our treatment of the polarized case. The integration measures are Lorentz invariant: $\int d^3q_1/E(\boldsymbol{q}_1) = \int d^4q_1\delta[\frac{1}{2}(\vec{q}_1\cdot\vec{q}_1+m_e^2)]$. $E(\boldsymbol{q}) \equiv q^0$ is set by the mass shell condition. We complete our

[5] This also follows from the well known transformation law of the solid angle element.

discussion of the Boltzmann equation by checking that the lab frame kinetic equation can be recast in such a manifestly covariant form. We will need the following identities:

$$\int \frac{d^3q_1}{E(\boldsymbol{q}_1)} \int \frac{d^3q_2}{E(\boldsymbol{q}_2)} \int \frac{d^3p_2}{E(\boldsymbol{p}_2)}\, \delta^4(\vec{p}_1 + \vec{q}_1 - \vec{p}_2 - \vec{q}_2)$$
$$= \int \frac{d^3q_1}{E(\boldsymbol{q}_1)} \int \frac{d\Omega_2\, p_2^2}{(-\vec{p}_2 \cdot \vec{q}_2)} = \int \frac{d^3q_2}{E(\boldsymbol{q}_2)} \int \frac{d\Omega_2\, p_2^2}{(-\vec{p}_2 \cdot \vec{q}_1)}\ . \tag{3.75}$$

These follow from writing the time part of the delta function as $\delta[E(\boldsymbol{p}_1) + E(\boldsymbol{q}_1) - E(\boldsymbol{p}_2) - E(\boldsymbol{p}_1 + \boldsymbol{q}_1 - p_2\boldsymbol{n}_2)]$ or $\delta[E(\boldsymbol{p}_1) + E(p_2\boldsymbol{n}_2 + \boldsymbol{q}_2 - \boldsymbol{p}_1) - E(\boldsymbol{p}_2) - E(\boldsymbol{q}_2)]$. To get the forms that we finally need, we replace the denominators of the angular integrals using the identities $\vec{p}_1 \cdot \vec{q}_1 = \vec{p}_2 \cdot \vec{q}_2$ and $\vec{p}_1 \cdot \vec{q}_2 = \vec{p}_2 \cdot \vec{q}_1$, which follow from conservation of total 4-momentum.

With these identities, the Boltzmann equation finally takes a manifestly covariant form,

$$p_1 \frac{d}{dt} f^{\mu\nu}(\vec{p}_1\,) = \int \frac{d^3q_1}{E(\boldsymbol{q}_1)} \int \frac{d^3q_2}{E(\boldsymbol{q}_2)} \int \frac{d^3p_2}{E(\boldsymbol{p}_2)}\, \delta^4(\vec{p}_1 + \vec{q}_1 - \vec{p}_2 - \vec{q}_2)$$
$$\times \left[\Phi^{\mu\nu}_{\gamma\delta}(\vec{p}_1, \vec{v}_2; \vec{p}_2, \vec{v}_2) f^{\gamma\delta}(\vec{p}_2) g_e(\boldsymbol{q}_2) - \phi^{\mu\nu}(\vec{p}_1) g_{\alpha\beta} \Phi^{\alpha\beta}_{\gamma\delta}(\vec{p}_2, \vec{v}_1; \vec{p}_1, \vec{v}_1) f^{\gamma\delta}(\vec{p}_1) g_e(\boldsymbol{q}_1) \right]. \tag{3.76}$$

The integration measures are Lorentz invariant: $\int d^3q_1/E(\boldsymbol{q}_1) = \int d^4q_1 \delta[\frac{1}{2}(\vec{q}_1 \cdot \vec{q}_1 + m_e^2)]$. It may be checked that working backwards from this equation, integration over the 4-dimensional delta function yields the rest frame form of the master equation Eqn. (3.71).

Chapter 4

Theory of the polarized Sunyaev-Zeldovich effects

In this chapter we apply the polarization tensor methods described in the previous chapters to the calculation of the polarized Sunyaev-Zeldovich (SZ) effect in clusters of galaxies.

In §4.1, we review briefly the unpolarized thermal and kinematic effects. In §4.2, the generation of polarization by scattering of the unpolarized isotropic part of the incident CMB (the monopole in the spherical harmonic expansion) by the electron distribution is considered. In this case, motion of electrons in a cluster with a bulk velocity of magnitude V_b with respect to the CMB rest frame produces an apparent CMB quadrupole in the rest frame of the scattering electrons, of order V_b^2. Thus a single scattering generates polarization of order $V_b^2\tau_{\rm T}$, where $\tau_{\rm T}$ is the optical depth to Thomson scattering along the line of sight through the cluster. We present a detailed calculation of the polarization matrix of the scattered radiation, which yields in addition to the polarization magnitude, the unpolarized thermal and kinematic effects also. We obtain the first relativistic corrections to the polarization and intensity distortions.

In §4.3 we discuss briefly the generation of polarization by double scattering of the CMB photons. We obtain expressions for the Stokes parameters of the effect, and present numerical results in the case of a homogeneous spherical cluster.

In §4.4 we describe a Monte Carlo procedure for simulation of the Sunyaev-Zeldovich effects, based on the polarization tensor master equation.

We defer a detailed discussion of the component of the polarization generated by scattering of the intrinsic CMB quadrupole until §5.3 in Chapter 5, where its use as a cosmological probe is considered.

4.1 Unpolarized thermal and kinematic SZ effects

The CMB photons propagating from the surface of last scattering have a blackbody spectrum with temperature $T_{\rm CMB} \approx 2.726K$. Recall that in a blackbody radiation

field at temperature $T_{\rm CMB}$, the specific intensity at frequency ν is given by

$$I_\nu = I_0 = i_0 \frac{x^3}{e^x - 1} \ , \tag{4.1}$$

where $x \equiv h\nu/k_B T_{\rm CMB}$ and $i_0 \equiv 2(k_B T_{\rm CMB})^3/(hc)^2$. The corresponding photon occupation number (averaged over the two polarization states) is

$$n_0 = \frac{1}{e^x - 1} \ . \tag{4.2}$$

As the photons pass through the electron-proton plasma in a galaxy cluster, with typical electron temperatures $T_e \approx 10^8$K, some are inverse Compton scattered by the electrons, on average to higher energies (scattering off the protons is negligible because the cross section is smaller by a factor of m_e/m_p). This may happen multiple times before they leave the cluster, but cluster optical depths are roughly 0.01, so the probability that a photon scatters more than once is small. Consequently the CMB spectrum observed on the sky in the direction of a cluster is slightly distorted relative to that seen away from a cluster. Roughly speaking, the distortion is an increase in the specific intensity in the Wien tail and a decrement in the Rayleigh-Jeans tail.

The canonical way to derive the frequency dependence of the distortion produced by a homogeneous cloud of electrons at temperature T_e is via the *Kompaneets equation*, which is an equation for the rate of change of the photon occupation number due to the process of inverse-Compton scattering in a homogeneous plasma:

$$\frac{dn}{dt} = (n_e \sigma_{\rm T} c) \left(\frac{k_B T_e}{m_e c^2} \right) \frac{1}{x^2} \frac{\partial}{\partial x} \left[x^4 \left(\frac{T_e}{T_{\rm CMB}} \frac{\partial n}{\partial x} + n + n^2 \right) \right] \ , \tag{4.3}$$

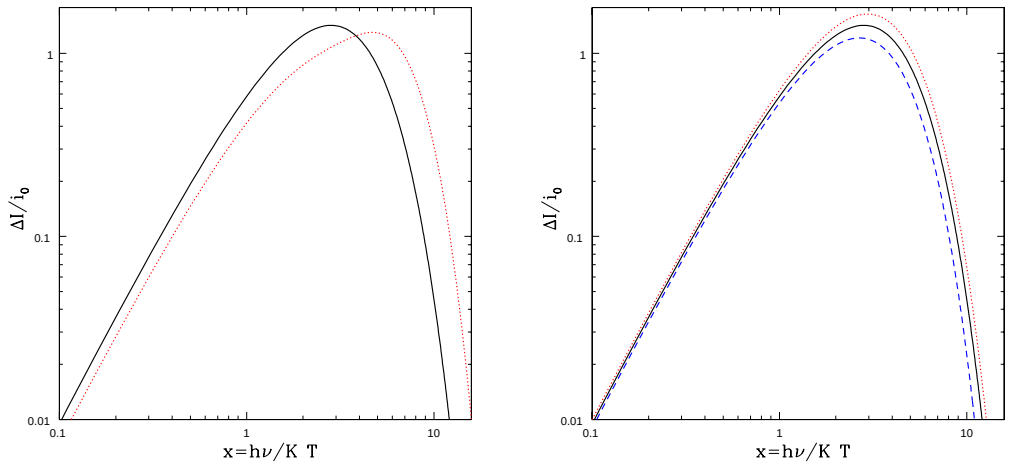
This equation is obtained via a Fokker-Planck type approximation where the typical photon frequency shift is assumed to be small compared to the photon frequency (see e.g. Rybicki & Lightman (1979) for a derivation). Since $T_e \gg T_{\rm CMB}$ in galaxy clusters, the second two terms in round brackets (the second of which, incidentally, arises from the quantum effect of stimulated emission) may be neglected in comparison to the first. Then substituting Eqn. (4.2) into the right hand side, and replacing $n_e \sigma_{\rm T} c$ with the Thomson optical depth $\tau_{\rm T}$ (a valid procedure in the limit of low optical depth), yields the well known *thermal SZ* distortion (Zeldovich & Sunyaev, 1969; Sunyaev & Zeldovich, 1980a):

$$\Delta I_\nu = i_0 y \frac{x^4 e^x}{(e^x - 1)^2} F(x) \ , \quad F(x) \equiv -4 + x \coth \left(\frac{x}{2} \right) \ , \tag{4.4}$$

where the *Compton y-parameter* is:

$$y = \int \frac{k_B T_e}{m_e c^2} \, n_e \sigma_T dl \ , \tag{4.5}$$

the integral ranging over the line of sight through the cluster.



(a) Thermal effect (b) Kinematic effect

Figure 4-1: The frequency dependence of the intensity distortion in the CMB due to the (a) thermal, and (b) kinematic SZ effects, for a cluster with highly exaggerated temperature, radial velocity and optical depth. Note that the thermal SZ distortion has a zero at $\nu \approx 217\,\mathrm{Ghz}$ ($x \approx 3.83$).

Note that it is common practice to express the intensity distortion in terms of the Rayleigh-Jeans (RJ) brightness temperature — the temperature of a blackbody with the same intensity at the given frequency, in the RJ limit. This is given by

$$\Delta T_B = \frac{c^2}{2k_B\nu^2}\Delta I_\nu = \frac{T_{\mathrm{CMB}}}{i_0 x^2}\Delta I_\nu \ . \tag{4.6}$$

The thermal SZ effect has attracted considerable attention in the astronomical community because, in conjunction with measurements of the X-ray luminosity of clusters, it offers the prospect of a measurement of the Hubble constant (Birkinshaw, 1999; Rephaeli, 1995a; Holzapfel et al., 1997). So far the measurements have had fairly limited success in this regard due to the many sources of systematic error in the measurement. Several authors have derived the relativistic corrections to the thermal effect, which turn out to be quite substantial (Rephaeli, 1995b; Challinor & Lasenby, 1998; Itoh et al., 1998; Nozawa et al., 1998; Sazonov & Sunyaev, 1998a; Molnar & Birkinshaw, 1999). Measurement of the relativistic corrections would allow measurement of the Hubble constant even without the X-ray observations (Diego et al., 2003).

In the *kinematic SZ* effect (KSZ) (Sunyaev & Zeldovich, 1980b; Phillips, 1995), the bulk motion of the cluster with respect to the CMB rest frame (the frame in which the CMB dipole vanishes) leads to a Doppler boosting of the incident CMB radiation. This leads to an enhancement or decrement in the CMB intensity depending on whether the cluster is approaching or receding along the line of sight. The frequency

dependence is simply (Rephaeli, 1995a),

$$\Delta I_\nu = -i_0 \frac{x^4 e^x}{(e^x-1)^2} \frac{V_r}{c} \tau_{\rm T} \ , \tag{4.7}$$

where V_r is the velocity component along the line of sight (directed toward the observer if $V_r < 0$). Thus via the kinematic SZ effect, the radial component of the peculiar velocity of the cluster could be measured (Kashlinsky & Atrio-Barandela, 2000), This has yet to be achieved in practice. Relativistic corrections to the kinematic effect have also been derived (Sazonov & Sunyaev, 1998a), but are less important than in the thermal effect. The thermal and kinematic effects are plotted for highly exaggerated cluster temperature and bulk velocity in Fig. 4-1. The cluster optical depth was set to $\tau_{\rm T} = 1$, and the distortions were computed with, in figure (a), electron temperature $k_B T_e = 50$ keV, in figure (b), radial bulk flow velocity $V_b = \pm 15000$ km s^{-1}. The upper and lower curves in (b) correspond to approaching and receding bulk velocity respectively. The solid line in both plots shows the undistorted Planck spectrum. For more realistic cluster parameters and optical depths, the thermal SZ brightness temperature distortion is typically ~ 1 mK, and the kinetic SZ distortion typically ~ 0.1 mK. Note that the SZ fractional intensity (or brightness) distortion is redshift independent, essentially since the photon momentum dependence appears only in the ratio $h\nu/k_B T_{\rm CMB}$. This fact makes SZ observations an especially useful probe of high redshift objects.

In the following section, we perform calculations of the polarization generated in the SZ effect using the kinetic equation for Thomson scattering derived in Chapter 3, which yields as a by-product derivations of both the thermal and kinematic SZ distortions, and their relativistic corrections. Note that the polarization magnitude can also be expressed as a brightness temperature distortion, using the formula Eqn. (4.6) with ΔI_ν replaced by $\sqrt{\Delta Q_\nu^2 + \Delta U_\nu^2}$.

4.2 Scattering of CMB monopole

In this section we compute the CMB intensity and polarization distortion, in the approximation of a single scattering, due to scattering of the unpolarized isotropic part of the incident CMB (the monopole) from moving electrons in the cluster gas. We deal only with an idealized galaxy cluster composed of a concentrated clump of electrons of density n_e and corresponding optical depth $\tau_{\rm T}$ in lab frame. The electrons are assumed to have a phase space density g_e given by a relativistic Maxwellian distribution with electron temperature T_e (which we henceforth write in the non-dimensionalized form $\theta_e \equiv k_B T_e / m_e c^2$), and a bulk 3-velocity $\boldsymbol{V}_b = \boldsymbol{\beta}_b c$ with respect to the CMB rest frame. The CMB rest frame will henceforth be called the "lab frame" in this section. We choose to align $\boldsymbol{V}_b$ with the z-axis of a Cartesian coordinate system.

We wish to calculate the polarization matrix $\boldsymbol{I}_s(\boldsymbol{n})$ resulting from Thomson scattering of the incident unpolarized CMB blackbody radiation in the lab frame into the viewing direction $\boldsymbol{n}$. Previous calculations have determined the intensity (Sunyaev &

Zeldovich, 1980b; Nozawa et al., 1998; Sazonov & Sunyaev, 1998a), and polarization magnitude (Sunyaev & Zeldovich, 1980b; Sazonov & Sunyaev, 1999; Audit & Simmons, 1999; Challinor et al., 2000; Itoh et al., 2000) of the distortion of the scattered radiation field as expansions in θ_e and β_b with various formalisms, but not in such a systematic and explicit fashion as we describe here.

In the single scattering limit, since the resulting scattered radiation field must be symmetric under rotations about the electron bulk velocity, the angular dependence of the intensity and polarization magnitude of the scattered radiation is a function only of the angle cosine $\mu = \boldsymbol{n} \cdot \boldsymbol{\beta}_b / |\boldsymbol{\beta}_b|$ between the electron bulk velocity and the line of sight. The resulting intensity distortion and polarization magnitude is an expansion in powers of β_b, μ, and the dimensionless electron temperature $\theta_e \equiv k_B T_e / m_e c^2$. By symmetry, the actual polarization vectors on the sky produced by this effect are simply all orthogonal to the direction of the cluster bulk velocity (they could also be parallel to it, depending on the sign of the Stokes parameter Q, but turn out to be orthogonal (Sazonov & Sunyaev, 1999)).

In a calculation with a real cluster with spatially extended structure we may replace $\tau_{\rm T}$ with the optical depth integrated along the line of sight to obtain the intensity distortion for each viewing angle.

We break the calculation into two stages. In §4.2.1 we perform the calculation in the case of a clump of electrons with zero temperature moving with a collective bulk velocity β_b along the z-direction in the lab frame, working entirely in the rest frame of the electrons. The polarization matrix of the scattered radiation is obtained, which on transformation to the lab frame yields the kinematic effects to any desired order in β_b. In §4.2.2, we extend this calculation to allow for thermal motion of the electrons. This is done by first generalizing the calculation of the rest frame scattered matrix in §4.2.1 to the case of a lab frame electron velocity in an arbitrary direction. Since the algebraic manipulations are lengthy and tedious, a computer algebra system is used. (One of advantages of the polarization matrix formalism is that it is quite simple to implement on a computer algebra system capable of handling matrix manipulations).

After transformation of the resulting scattered beam into lab, the integration over electron velocities is performed, yielding both the thermal and kinematic effects, and relativistic corrections.

4.2.1 Cold electrons

We begin by considering a simplified calculation in which the clump of electrons is taken to move collectively along the z-axis, i.e. the electron distribution has zero temperature. We work in the rest frame of the electrons, in which the Thomson limit form of the rest frame scattering term Eqn. (3.33) can be applied. We align the velocity 3-vector of the electrons in lab frame with the z-axis, and write the electron velocity in lab frame coordinates as

$$\vec{v}_e = \gamma_b (1, \beta_b \hat{\boldsymbol{z}}) \ , \quad \gamma_b \equiv \frac{1}{\sqrt{1-\beta_b^2}} \ . \tag{4.8}$$

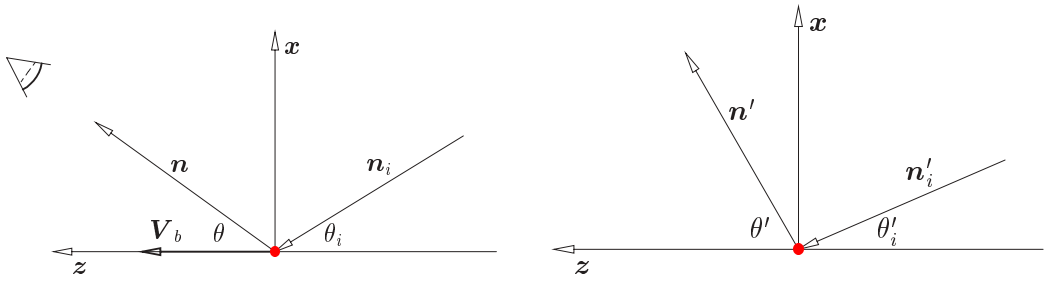


Figure 4-2: The coordinate system in the (a) lab frame, and (b) rest frame, used to evaluate the polarization matrix. In lab, the clump of electrons indicated at the origin travels along the z-axis with velocity V_b. Note that in the rest frame, we choose to consider the photons scattered in a direction $\boldsymbol{n}'$ in the x-z plane, but the incident photon direction $\boldsymbol{n}'_i$ is in a general direction.

The lab frame 4-velocity is denoted $\vec{v}_l$. The rest frame momentum of the incident photon is $p_i^{\mu'} = p'_i(1, \boldsymbol{n}'_i)$, where the rest frame direction vector is expressed in polar coordinates with respect to the z-axis:

$$\boldsymbol{n}'_i = \left(\cos\psi'_i\sqrt{1-\mu_i'^2},\ \sin\psi'_i\sqrt{1-\mu_i'^2},\ \mu'_i\right) \ . \tag{4.9}$$

The coordinate system is illustrated in Fig. 4-2. The corresponding lab frame momentum is $p_i^{\mu} = p_i(1, \boldsymbol{n}_i)$, where the lab frame direction vector is:

$$\boldsymbol{n}_i = \left(\cos\psi_i\sqrt{1-\mu_i^2},\ \sin\psi_i\sqrt{1-\mu_i^2},\ \mu_i\right) \ . \tag{4.10}$$

Assuming unpolarized isotropic incident CMB radiation in lab frame, the intensity polarization matrix of a photon incident in the lab frame with 4-momentum $\vec{p}_i$ is given by

$$I^{\mu\nu}(\vec{p}_i, \vec{v}_l) = \frac{1}{2} I_0(p_i) P^{\mu\nu}(\vec{p}_i, \vec{v}_l) \ , \tag{4.11}$$

where I_0 is the Planck function at the mean temperature of the CMB, $T_{\rm CMB}$:

$$I_0(p) = \frac{2c}{h^2}\frac{p^3}{e^{pc/k_B T_{\rm CMB}} - 1} \ . \tag{4.12}$$

The incident photon momentum in the lab frame is Doppler shifted on going to the rest frame:

$$p'_i = \gamma_b p_i(1 - \beta_b \mu_i) \ , \tag{4.13}$$

This may be written in terms of the incident polar angle in the rest frame. Using the formula for relativistic aberration,

$$\mu_i = \frac{\mu'_i + \beta_b}{1 + \beta_b \mu'_i} \ , \tag{4.14}$$

we obtain

$$p_i' = \frac{p_i}{\gamma_b(1+\beta_b\mu_i')} \ . \tag{4.15}$$

The specific intensity tensor in the rest frame can be obtained from that in the lab frame using the transformation law of the intensity between frames,

$$\begin{aligned} I'(p_i',\mu_i') &= \left(\frac{p_i'}{p_i}\right)^3 I_0(p_i) = \frac{I_0[\gamma_b p_i'(1+\beta_b\mu_i')]}{\gamma_b^3(1+\beta_b\mu_i')^3} \\ &= \frac{2h}{c^2}\frac{{p_i'}^3}{e^{\gamma_b p_i'(1+\beta_b\mu_i')/k_B T_{\rm CMB}}-1} \ . \end{aligned} \tag{4.16}$$

The isotropic specific intensity in the lab frame, $I_0(p_i)$, transforms into an anisotropic intensity in the rest frame which is still of blackbody form but with a temperature with angular dependence:

$$T(\mu_i') = \frac{T_{\rm CMB}}{\gamma_b(1+\beta_b\mu_i')} \ . \tag{4.17}$$

The incident radiation field in the rest frame is of course also unpolarized and has intensity tensor:

$$\boldsymbol{I'}(\vec{p}_i,\vec{v}_e) = I^{\mu'\nu'}(\vec{p}_i,\vec{v}_e) = \frac{1}{2}I'(p_i',\mu_i')P^{\mu'\nu'}(\vec{p}_i,\vec{v}_e) \ . \tag{4.18}$$

Now using the kinetic equation in the form Eqn. (3.33) we may write down the scattering term in the rest frame. The Thomson limit form is appropriate here since even with the boost from the lab to rest frame, the CMB photon momenta are a tiny fraction of the electron mass and therefore recoil is negligible. We evaluate the kinetic equation at photon 4-momentum $\vec{p}$, with the following components in rest frame coordinates:

$$p^{\mu'} = p'(1,\boldsymbol{n}'), \quad \boldsymbol{n}' = \left(\cos\psi'\sqrt{1-\mu'^2}, \sin\psi'\sqrt{1-\mu'^2}, \mu'\right) \ , \tag{4.19}$$

where $p' = p_i'$ since we are working in the Thomson limit.

In the single scattering limit, we may insert the polarization tensor of the incident unpolarized radiation field in the right hand side of the kinetic equation to obtain the scattered beam:

$$\frac{d}{dt'}I^{\mu'\nu'}(\vec{p},\vec{v}_e) = \frac{3}{4}n_e'\sigma_T\int\frac{d\Omega_i'}{4\pi}I'(p_i',\mu_i')G^{\mu'\nu'}(\mu',\psi';\mu_i',\psi_i') - \frac{1}{2}I'(p',\mu')P^{\mu'\nu'}(\vec{p},\vec{v}_e) \ . \tag{4.20}$$

where

$$G^{\mu'\nu'}(\mu',\psi';\mu_i',\psi_i') = P^{\mu'}{}_{\alpha'}(\vec{p},\vec{v}_e)P^{\nu'}{}_{\beta'}(\vec{p},\vec{v}_e)P^{\alpha'\beta'}(\vec{p}_i,\vec{v}_e) \ , \tag{4.21}$$

The 00 and $0i$ components of this tensor equation obviously vanish when evaluated in electron rest frame coordinates. We evaluate the gain term by first performing the

integral over azimuthal angles $d\psi_i'$, using the explicit form of the spatial part of the projection tensor $P^{i'j'}(\vec{p}_i, \vec{v}_e)$:

$$\begin{pmatrix} 1+(\mu_i'^2-1)\cos^2\psi_i' & (\mu_i'^2-1)\cos\psi_i'\sin\psi_i' & -\mu_i'\sqrt{1-\mu_i'^2}\cos\psi_i' \\ (\mu_i'^2-1)\cos\psi_i'\sin\psi_i' & 1+(\mu_i'^2-1)\sin^2\psi_i' & -\mu_i'\sqrt{1-\mu_i'^2}\sin\psi_i' \\ -\mu_i'\sqrt{1-\mu_i'^2}\cos\psi_i' & -\mu_i'\sqrt{1-\mu_i'^2}\sin\psi_i' & 1-\mu_i'^2 \end{pmatrix} \quad , \quad (4.22)$$

Performing the integral over azimuth yields

$$\int_0^{2\pi} \frac{d\psi_i'}{2\pi} P^{i'j'}(\vec{p}_i, \vec{v}_e) = \begin{pmatrix} \frac{1}{2}(1+\mu_i'^2) & 0 & 0 \\ 0 & \frac{1}{2}(1+\mu_i'^2) & 0 \\ 0 & 0 & \frac{1}{2}(1-\mu_i'^2) \end{pmatrix} \quad . \quad (4.23)$$

To further simplify, we may evaluate the rest of the scattering term at $\psi' = 0$, since by azimuthal symmetry of the radiation field in the rest frame the intensity tensor for a general (μ', ψ') is related to that at $\psi' = 0$ by a rotation about the z–axis through angle ψ', and the polarization magnitude is independent of azimuth. Putting $\boldsymbol{n}' = \left(\sqrt{1-\mu'^2}, 0, \mu'\right)$, the explicit form of the spatial part of the projection tensor is $P^{i'j'}(\vec{p}, \vec{v}_e)$ is:

$$P^{i'j'}(\vec{p}, \vec{v}_e) = \begin{pmatrix} \mu'^2 & 0 & -\mu'\sqrt{1-\mu'^2} \\ 0 & 1 & 0 \\ -\mu'\sqrt{1-\mu'^2} & 0 & 1-\mu'^2 \end{pmatrix} \quad . \quad (4.24)$$

Now performing the multiplication of two of the matrices in Eqn. (4.24) with the matrix in Eqn. (4.23), yields the azimuthal integral of Eqn. (4.21) required in the gain term of the kinetic equation (4.20):

$$\bar{G}^{i'j'} \equiv \int_0^{2\pi} \frac{d\psi_i'}{2\pi} G^{i'j'}(\mu', 0; \mu_i', \psi_i') \; = \\ \frac{1}{2} \begin{pmatrix} G_\parallel(\mu', \mu_i')\mu'^2 & 0 & -\mu'\sqrt{1-\mu'^2}G_\parallel(\mu', \mu_i') \\ 0 & G_\perp(\mu', \mu_i') & 0 \\ -\mu'\sqrt{1-\mu'^2}G_\parallel(\mu', \mu_i') & 0 & G_\parallel(\mu', \mu_i')\left(1-\mu'^2\right) \end{pmatrix} \quad , \quad (4.25)$$

where

$$\begin{aligned} G_\parallel(\mu', \mu_i') &= 2 - \mu'^2 + \mu_i'^2\left(3\mu'^2 - 2\right) \quad , \\ G_\perp(\mu', \mu_i') &= 1 + \mu_i'^2 \quad . \end{aligned} \quad (4.26)$$

One can check that

$$\frac{1}{2}\int_{-1}^{1} d\mu_i' \, \bar{G}^{\mu'\nu'} = \frac{2}{3} P^{\mu'\nu'}(\vec{p}, \vec{v}_e) \; , \quad (4.27)$$

(evaluated at $\psi' = 0$). Thus as $\beta_b \to 0$, $\frac{d}{dt} I^{\mu'\nu'}(\vec{p}, \vec{v}_e) \to 0$, since by symmetry

scattering of an isotropic radiation field from a stationary electron cannot alter the radiation field.

Now putting together the gain and loss terms, and integrating μ_i', we find for a finite rest frame time interval $\Delta t'$ (all evaluated at $\psi' = 0$ in rest frame coordinates):

$$\frac{\boldsymbol{\Delta I'}(\vec{p},\vec{v}_e)}{\Delta\tau'_{\rm T}} = \frac{1}{2}\begin{pmatrix} I_\parallel(p',\mu')\mu'^2 & 0 & -\mu'\sqrt{1-\mu'^2}I_\parallel(p',\mu') \\ 0 & I_\perp(p',\mu') & 0 \\ -\mu'\sqrt{1-\mu'^2}I_\parallel(p',\mu') & 0 & I_\parallel(p',\mu')\left(1-\mu'^2\right) \end{pmatrix} , \tag{4.28}$$

where we set the optical depth in the rest frame to $\Delta\tau'_{\rm T} \equiv n'_e\sigma_T\Delta t'$, and defined

$$\begin{aligned} I_\parallel(p',\mu') &= \frac{3}{8}\left[(2-\mu'^2)J(p') + (3\mu'^2-2)K(p') - \frac{8}{3}I'(p',\mu')\right] , \\ I_\perp(p',\mu') &= \frac{3}{8}\left[J(p') + K(p') - \frac{8}{3}I'(p',\mu')\right] , \end{aligned} \tag{4.29}$$

and the functions (note that these are functions of p'_i, but in the Thomson limit this equals the scattered photon momentum p'):

$$\begin{aligned} J(p') &\equiv \int_{-1}^{1} d\mu'_i\; I'(p',\mu'_i) , \\ K(p') &\equiv \int_{-1}^{1} d\mu'_i\; I'(p',\mu'_i)\mu'^2_i . \end{aligned} \tag{4.30}$$

Using the dimensionless frequency $x' = cp'/k_BT_{\rm CMB}$, and the constant i_0 (recall Eqn. (4.1)), these functions are given by the integrals

$$\begin{aligned} J(x',\beta_b) &= i_0 x'^3 \int_{-1}^{1} d\mu'_i \left[e^{\gamma_b x'(1+\beta_b\mu'_i)} - 1\right]^{-1} , \\ K(x',\beta_b) &= i_0 x'^3 \int_{-1}^{1} \mu'^2_i d\mu'_i \left[e^{\gamma_b x'(1+\beta_b\mu'_i)} - 1\right]^{-1} . \end{aligned} \tag{4.31}$$

The intensity of the scattered radiation in the lab frame is thus given by the trace

$$\begin{aligned} \Delta I'(x',\mu') &= {\rm Tr}[\boldsymbol{\Delta I'}(\vec{p},\vec{v}_e)] \\ &= \frac{3}{16}\Delta\tau'_{\rm T}\left[(3-\mu'^2)J(x',\beta_b) + (3\mu'^2-1)K(x',\beta_b) - \frac{16}{3}I'(x',\mu')\right] \end{aligned} \tag{4.32}$$

where

$$I'(x',\mu') = i_0 x'^3 \left[e^{\gamma_b x'(1+\beta_b\mu')} - 1\right]^{-1} . \tag{4.33}$$

The polarization magnitude, in the limit of small $\Delta\tau'_{\rm T}$, is given by the formula (2.41):

$$\Pi(x',\mu') = \Pi(\boldsymbol{I'}(\vec{p},\vec{v}_e) + \tau_{\rm T}\boldsymbol{\Delta I}(\vec{p},\vec{v}_e)) = \Delta\tau'_{\rm T}\left(\frac{{\rm Tr}[\boldsymbol{\Delta I'}]}{{\rm Tr}[\boldsymbol{I'}]}\right)\Pi(\boldsymbol{\Delta I'}) . \tag{4.34}$$

Evaluating this, we find

$$\Pi(x',\mu') = \frac{3}{16}\Delta\tau'_{\rm T}\ \left(1-\mu'^2\right)\ \frac{3K(x',\beta_b)-J(x',\beta_b)}{I'(x',\mu')}\ . \tag{4.35}$$

This formula, and the μ'_i angular dependence inside the integrands of K and J, shows that only the quadrupole in the incident intensity in the rest frame generates polarization. The polarization magnitude has the familiar $\sin^2\theta'$ dependence. As $\beta_b \to 0$, the incident radiation field in the rest frame becomes isotropic, $I'(x',\mu') \to I_0$ (of Eqn. (4.1)), yielding $J \to 2I_0, K \to 2I_0/3$, and thus $\Pi(x',\mu') \to 0$ as expected.

We now expand the denominator of the integrands of $J(x',\beta_b)$ and $K(x',\beta_b)$ in powers of β_b:

$$\left[e^{\gamma_b x'(1+\beta_b\mu')}-1\right]^{-1} = n_0(x) - \frac{e^{x'}}{(e^{x'}-1)^2}\sum_{n=0}^{\infty}\left(\frac{e^{x'}}{e^{x'}-1}\right)^n(-1)^n\delta^{n+1}\ , \tag{4.36}$$

where

$$\delta \equiv e^{(\gamma_b-1)x'}e^{\gamma_b\beta_b\mu' x'}-1\ . \tag{4.37}$$

Expanding δ up to second order in β_b, we find:

$$\begin{aligned}
\int_{-1}^{1} d\mu'\ \delta &= \beta_b^2 x'\left(1+\frac{1}{3}x'\right)+O(\beta_b^4)\ ,\\
\int_{-1}^{1} d\mu'\ \delta^2 &= \frac{2}{3}\beta_b^2 x'^2+O(\beta_b^4)\ ,\\
\int_{-1}^{1} d\mu'\ \mu'^2\ \delta &= \frac{1}{3}\beta_b^2 x'\left(1+\frac{3}{5}x'\right)+O(\beta_b^4)\ ,\\
\int_{-1}^{1} d\mu'\ \mu'^2\ \delta^2 &= \frac{2}{5}\beta_b^2 x'^2+O(\beta_b^4)\ .
\end{aligned} \tag{4.38}$$

To obtain $\Pi(x',\mu')$ to second order in β_b, since the numerator is already second order, we may replace $I'(x',\mu')$ in Eqn. (4.35) with I_0.

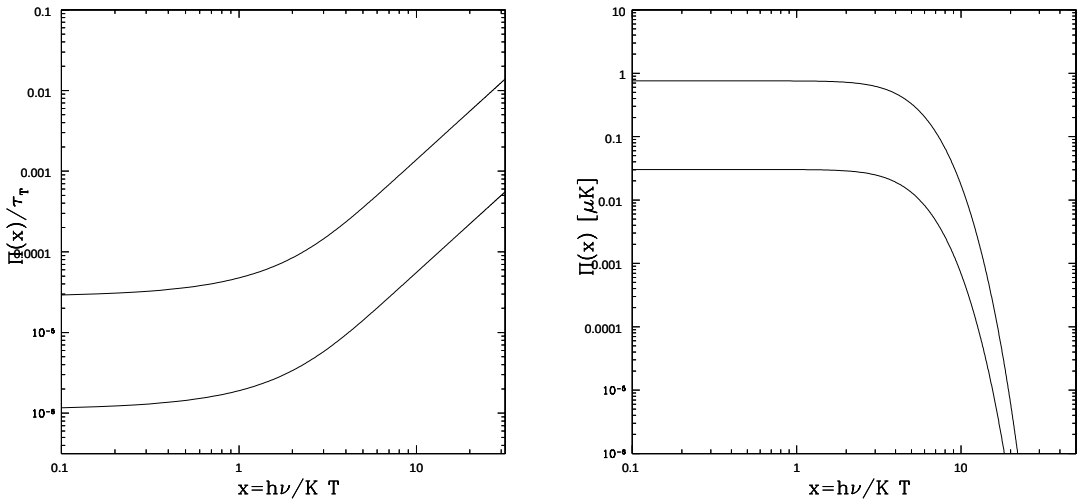
Using the results in the previous equation we obtain finally the polarization magnitude in the rest frame to second order in β_b:

$$\Pi(x',\mu') = \frac{1}{20}\ \Delta\tau'_{\rm T}x'^2\frac{e^{x'}(e^{x'}+1)}{(e^{x'}-1)^2}\beta_b^2\ (1-\mu'^2)\ . \tag{4.39}$$

On Lorentz transforming into the lab frame, the polarization magnitude of the scattered photons does not change, but the photon angle is aberrated, with

$$\mu' = \frac{\mu-\beta_b}{1-\beta_b\mu}\ . \tag{4.40}$$

Since $\mu'^2 = \mu^2 + O(\beta_b)$, and $x' = x + O(\beta_b)$, and the optical depth $\Delta\tau'_{\rm T}$ transforms



(a) Dimensionless $\Pi/\tau_{\rm T}$ (b) Brightness temperature Π

Figure 4-3: Polarization magnitude of the CMB scattered by a concentrated cloud of cold electrons with a bulk flow velocity V_b transverse to the line of sight. In (a), we plot the dimensionless polarization magnitude $\Pi(x)$ divided through by the optical depth $\tau_{\rm T}$. In (b), we plot the polarization magnitude $\Pi(x)$ as a (Rayleigh-JeansJ) brightness temperature distortion, taking $\tau_{\rm T} = 0.01$. The upper and lower curves in both (a) and (b) correspond to $V_b = 5000$ km s^{-1} and $V_b = 1000$ km s^{-1} respectively.

like dt', we have the same result in the lab frame quantities to this order:

$$\Pi(x,\mu) = \frac{1}{20}\,\Delta\tau_{\rm T} x^2 \frac{e^x(e^x+1)}{(e^x-1)^2}\beta_b^2\,(1-\mu^2)\;. \tag{4.41}$$

This result was obtained before by Sazonov & Sunyaev (1999); Audit & Simmons (1999); Challinor et al. (2000), all using different methods. Note that the $(1-\mu^2)$ dependence implies that this component of the CMB polarization is a direct measure of the peculiar velocity of the cluster gas *perpendicular* to the line of sight, which in conjunction with the intensity measurement allows, in principle, measurement of all the components of the cluster peculiar velocity. This will be an important cosmological probe, if the polarization measurements can be made. This will be a considerable experimental challenge, since the polarization magnitude is rather small, typically $0.1\,\mu$K at most, as illustrated in Fig. 4-3 (note that cluster bulk velocities rarely exceed $1000\,{\rm km\,s^{-1}}$).

One might worry that the dimensionless polarization magnitude Π goes quadratically in x as $x \to \infty$, which would seem to be a problem since the polarization magnitude must be bounded by unity. However, the analysis we have given is only the lowest order result - at high photon frequencies, relativistic corrections will mod-

ify Eqn. (4.41). Since we essentially expanded in powers of $\beta_b x$, our analysis cannot be trusted for frequencies greater than $x \approx 1/\beta_b$.

We finish this section by expanding the total intensity of the scattered radiation, given in Eqn. (4.32), to second order in β_b, and performing the transformation to lab frame to obtain the kinematic SZ distortion and its first relativistic correction. To do this calculation it is convenient to work with the phase space density rather than the intensity. The intensity distortion $\Delta I'(x', \mu')$ is related to the phase space density distortion by

$$\frac{\Delta I'(x', \mu')}{I'(x', \mu')} = \frac{\Delta f'(x', \mu')}{f'(x', \mu')} \ . \tag{4.42}$$

From the transformation law of the left hand side of the kinetic equation, given by Eqn. (3.67), we find an equation for the rate of change of the phase space density in lab:

$$\frac{1}{f(x,\mu)} \frac{df(x,\mu)}{dt} = \gamma_b (1 - \beta_b \mu)\, n'_e \sigma_{\rm T} \left(\frac{\Delta I'(x', \mu')}{\Delta \tau'_{\rm T} I'(x', \mu')} \right) \ . \tag{4.43}$$

Expanding the functions $K(x', \beta_b)$ and $J(x', \beta_b)$ in Eqn. (4.32) up to second order in β_b as before, and expressing the right hand side in lab frame quantities by making the replacements $x' = \gamma_b x(1 - \beta_b \mu)$, $\mu' = (\mu - \beta_b)/(1 - \beta_b \mu)$ and $n'_e = n_e/\gamma_b$, we find to $O(\beta_b^2)$ the fractional intensity distortion in lab:

$$\frac{\Delta I(x,\mu)}{I_0(x)} = \Delta\tau_{\rm T} \frac{x e^x}{e^x - 1} \left[\mu\, \beta_b + \left(-1 - \mu^2 + \frac{x\,(3 + 11\,\mu^2)\, \coth(\frac{x}{2})}{20} \right) {\beta_b}^2 \right] \ , \tag{4.44}$$

where the lab frame optical depth is defined by $\Delta\tau_{\rm T} = n_e \sigma_{\rm T} \Delta t$. This is the first relativistic correction to the kinematic SZ effect, obtained previously by Sazonov & Sunyaev (1998a,b). Note that without the correct "flux factor" in Eqn. (4.43), this would differ at second order in β_b^2. The first term is simply the lowest order kinematic SZ distortion (4.7), where $V_r = -\mu\beta_b$ is the bulk velocity projected along the line of sight (which is opposite to the direction of the scattered photon momentum, hence the minus sign).

4.2.2 Hot electrons

We now extend to the more general case of a Maxwellian distribution of electrons with dimensionless temperature $\theta_e \equiv k_B T_e/m_e$ moving with a bulk velocity $\boldsymbol{V}_b$ with respect to the CMB rest frame (lab frame). In the single scattering limit, the Thomson scattering of isotropic blackbody radiation from a Maxwellian distribution of electrons moving with a bulk velocity $\boldsymbol{V}_b = \boldsymbol{\beta}_b c$ produces a scattered radiation field whose intensity and polarization magnitude are azimuthally symmetric about $\boldsymbol{V}_b$. Our goal is to compute the polarization matrix of the scattered radiation field, as an expansion in powers of V_b and θ. This computation will yield, to lowest order, the usual thermal and kinematic SZ distortion of the intensity, and the polarization magnitude to lowest order in V_b. Going to higher order yields the "interference" terms between the thermal and kinematic effects, in both the intensity and polarization, and the relativistic corrections.

The first task is to determine the lab frame polarization matrix of the scattered beam due to scattering of an incident unpolarized isotropic blackbody radiation field in lab by an electron with a general lab frame velocity $\boldsymbol{\beta}$. This is not the bulk velocity but rather the velocity of some of the electrons in the thermal distribution, which will eventually be integrated over. This part is just a generalization of the calculation performed in section §4.2.1. The resulting polarization matrix of the lab frame scattered radiation field as a function of electron velocity may then be averaged over a distribution of lab frame electron velocities to yield the observed lab frame result. The steps required to compute this are described below. The actual calculation, even at lowest order, is quite lengthy, so a computer algebra system (Mathematica) was used to perform the calculation. We do not give all the algebra but just outline the procedure. Henceforth in this section primed indices refer to components of 4-vectors in the electron rest frame, and unprimed indices to components in the lab frame.

It is convenient to integrate over angles in the electron rest frame, but to express the electron velocity and final state photon momentum in lab coordinates throughout (to avoid a cumbersome transformation of rest frame angles to lab frame). In lab frame coordinates, the electron 4-velocity is

$$v_e^\mu = \gamma(1, \boldsymbol{\beta}) \ , \quad \gamma \equiv \frac{1}{\sqrt{1-\beta^2}} \ , \quad \beta \equiv |\boldsymbol{\beta}| \ . \tag{4.45}$$

The Cartesian coefficients of $\boldsymbol{\beta}$ are denoted β_i. In rest frame coordinates, the velocity of the lab frame is of course

$$v_l^{\mu'} = \gamma(1, -\boldsymbol{\beta}) \ . \tag{4.46}$$

The scattered photon momentum in lab frame coordinates is written

$$p_s^\mu = p_s(1, \boldsymbol{n}_s) \ . \tag{4.47}$$

To simplify the computation, we set up a polar coordinate system with polar axis along the z-direction and evaluate the scattered polarization matrix at azimuth $\psi_s = 0$:

$$\boldsymbol{n}_s \ = \ \left(\sqrt{1-\mu_s^2},\ 0,\ \mu_s\right) \ . \tag{4.48}$$

This is no loss of generality provided we choose the bulk velocity $\boldsymbol{V}_b$ to lie along the z–direction, in which case the polarization matrix for a general $\boldsymbol{n}_s$ is related to the one calculated here by a simple rotation about the z–axis.

The scattered photon momentum in the electron rest frame is found by applying Lorentz transformation matrices to obtain $p_s^{\mu'} = \Lambda^{\mu'}{}_\mu(\boldsymbol{\beta})\, p_s^\mu$, where

$$\begin{aligned} \Lambda^0{}_0 \ &= \ \gamma = 1/\sqrt{1-\beta^2}, \qquad \Lambda^0{}_i = \Lambda^i{}_0 = -\gamma\beta^i \ , \\ \Lambda^i{}_j \ &= \ (\gamma-1)\frac{\beta^i\beta_j}{\beta^2} + \delta^i{}_j \ . \end{aligned} \tag{4.49}$$

Using the notation of §3.3, we denote the momenta of the incident photons in the lab

and rest frames as follows:

$$\begin{aligned} p_i^\mu &= p_i(1, \boldsymbol{n}_i) \ , \quad p_i^{\mu'} = p_i'(1, \boldsymbol{n}_i') \ , \\ \boldsymbol{n}_i' &= \left(\cos\psi_i' \sqrt{1-\mu_i'^2},\ \sin\psi_i' \sqrt{1-\mu_i'^2},\ \mu_i'\right) \ . \end{aligned} \tag{4.50}$$

with $p_i = -\vec{v}_l \cdot \vec{p}_i$, $p_i' = -\vec{v}_e \cdot \vec{p}_i$. In the lab frame, the scalar occupation number of the incident photons is isotropic with a Planck spectrum:

$$n_i(\vec{p}_i) = \frac{1}{e^{p_i/k_B T_{\rm CMB}} - 1} \ . \tag{4.51}$$

(Note, do not confuse $\boldsymbol{n}_i$, a direction vector, with n_i, the occupation number!). In the rest frame, $n_i'(\vec{p}_i) = n_i(\vec{p}_i)$, but the occupation number of the incident photons is no longer isotropic since photons with different momenta are aberrated through different angles. Thus n_i becomes a function of p_i' and $\boldsymbol{n}_i'$ through p_i:

$$n_i'(\vec{p}_i) = \frac{1}{e^{p_i(p_i', \boldsymbol{n}_i')/k_B T_{\rm CMB}} - 1} \ , \tag{4.52}$$

where in terms of $\boldsymbol{\beta}$,

$$p_i(p_i', \boldsymbol{n}_i') = \gamma p_i' \left(1 + \boldsymbol{\beta} \cdot \boldsymbol{n}_i'\right) \ . \tag{4.53}$$

The angular dependence of the incident radiation field in the rest frame is obtained by expanding (4.52) in powers of the velocity components β_i. For the lowest order polarization computation, the expansion must be taken up to at least second order in the velocity components.

Then as in the previous section, the right hand side of the rest frame master equation (3.33) is constructed, and the integration over the rest frame angles of the incident beam performed. The resulting matrix is then transformed into lab frame by the application of two projection tensors, and the lab frame fractional intensity distortion obtained, making sure, as in Eqn. (4.43), to multiply by the correct flux factor, which now has the form $\gamma(1 - \boldsymbol{\beta} \cdot \boldsymbol{n})$.

We thus obtain the lab frame polarization matrix as a function of the lab frame photon direction and the velocity components $\boldsymbol{\beta}$. In the lab frame, the integration over electron velocities is performed. To do this we need first to construct the distribution function of electron velocities in lab frame. In the "comoving frame", denoted with primes, in which the average electron velocity vanishes, the electron phase space distribution function as a function of the electron 3-momentum $\boldsymbol{q}'$ is assumed to be a relativistic Maxwellian at dimensionless temperature θ_e:

$$g_e(\boldsymbol{q}') = g_0 \, e^{-E(\boldsymbol{q}')/(m_e \theta_e)} \ . \tag{4.54}$$

where $E(\boldsymbol{q}') = \sqrt{\boldsymbol{q}'^2 + m_e^2}$, and g_0 is a normalization constant which depends on the total number density of electrons. We use a relativistic Maxwellian in order to retain the corrections to the SZ effect in a mildly relativistic plasma.

With $\boldsymbol{q}' = m_e \gamma' \boldsymbol{\beta}'$, where $\boldsymbol{\beta}'$ is the electron 3-velocity in the comoving frame, and

$\gamma' \equiv 1/\sqrt{1-\beta'^2}$, we have (as a function of β since the distribution is isotropic in the comoving frame)

$$g_e(\beta') = g_0\, e^{-\gamma'/\theta_e} \ . \tag{4.55}$$

The number density of electrons in each comoving frame momentum element d^3p' is thus

$$dn'_e = 4\pi g_0\, p'^2\ dp' e^{-\gamma'/\theta_e} \ . \tag{4.56}$$

Integrating this distribution over the element d^3p' yields the electron number density in the comoving frame n'_e. Using $p' = m_e\gamma'\beta'$, we find $p'^2dp' = m_e^3\gamma'^5\beta'^2d\beta'$. Thus

$$\begin{aligned} n'_e &= g_0 \int 4\pi\gamma'^5\beta'^2 d\beta' e^{-\gamma'/\theta_e} \\ &= 4\pi g_0\, m_e^3\ \theta_e K_2(1/\theta_e) \ , \end{aligned} \tag{4.57}$$

where K_2 is a modified Bessel function (for a derivation of this result see for example Lightman et al. (1975) or Synge (1957)).

Thus in the comoving frame, the number density of electrons in each comoving frame velocity element d^3v', is

$$dn'_e = n'_e \frac{d^3\beta'}{4\pi} \frac{\gamma'^5 e^{-\gamma'/\theta_e}}{\theta_e K_2(1/\theta_e)} \ . \tag{4.58}$$

For small θ_e, the denominator can be expanded:

$$4\pi\theta_e K_2(1/\theta_e) e^{1/\theta_e} = (2\pi\theta_e)^{3/2}\left[1 + 15\theta_e/8 + \cdots\right] \ , \tag{4.59}$$

yielding the familiar prefactor of the non-relativistic Maxwellian to lowest order.

Now we wish to compute the analogous lab frame quantity by a Lorentz transformation from the comoving frame. Since the distribution function is a Lorentz scalar, the number density element transforms like the momentum space volume element in Eqn. (4.56). Using the Lorentz invariance of the quantity $d^3p/p = d^3p'/p'$, it follows that

$$p'^2dp' = m_e^3\gamma'^5\beta'^2d\beta' = m_e^3\gamma^5\beta^2 d\beta(\gamma'/\gamma) \ . \tag{4.60}$$

Choosing the bulk velocity of the comoving frame with respect to the lab frame to be $V_b = \beta_b c$ in the z-direction, we have

$$\gamma'/\gamma = \gamma_b\left(1 - \beta_z\beta_b\right) \ , \qquad \gamma_b \equiv (1-\beta_b^2)^{-1/2} \ . \tag{4.61}$$

For calculations it is convenient to write the distribution function in lab frame in a form in which the non-relativistic part of the Maxwellian, which has Gaussian form, is pulled out and the rest expanded in a series in powers of the velocity relative to the dimensionless bulk velocity $\boldsymbol{\beta}_b = (0, 0, \beta_b)$:

$$\frac{dn_e}{n_e} = \frac{e^{-(\gamma-1)/\theta_e}}{\theta_e e^{1/\theta_e} K_2(1/\theta_e)} \left[\gamma^5\gamma_b\left(1 - \beta_z\beta_b\right) e^{-\gamma[\gamma_b(1-\beta_z\beta_b)-1]/\theta_e}\right] \frac{d^3\boldsymbol{\beta}}{4\pi} \ . \tag{4.62}$$

Making the substitution $\beta_z \to \tilde{\beta}_z + V_b$, the part in square brackets may be expanded straightforwardly about unity in powers of $\beta_x, \beta_y, \tilde{\beta}_z$, β_b and θ_e. Defining

$$\tilde{\beta}^2 \equiv \beta_x^2 + \beta_y^2 + \tilde{\beta}_z^2 \ , \tag{4.63}$$

the exponential factor in front can be written as

$$e^{-\tilde{\beta}^2/2\theta_e} \times \text{prefactor} \ , \tag{4.64}$$

where the prefactor is an expansion about unity in powers of $\tilde{\beta_z}, \beta_x, \beta_y$. The result is a Gaussian multiplied by a prefactor which is polynomial in the components β_i with coefficients which are functions of β_b and θ_e.

A further transformation is required before the lab frame integral can be done. The integral $d^3\boldsymbol{\beta}$ ranges over the velocity sphere $|\boldsymbol{\beta}| \leq 1$. To simplify the Gaussian integrals, it is easier to make the transformation $\beta_i \to u_i/\gamma$, and integrate $d^3\boldsymbol{u}$ over all space. With this transformation, we find

$$\gamma = \sqrt{1+u^2} \ , \qquad \gamma^5 \beta^2 d\beta = u^2 du \ . \tag{4.65}$$

Steps similar to those described above yield an expansion about the transformed bulk velocity $\boldsymbol{\beta}_b \to \boldsymbol{U}_b = (0, 0, \beta_b/\sqrt{1+\beta_b^2})$ in powers of $\tilde{u}_z = u_z - \beta_b/\sqrt{1-\beta_b^2}$, u_x, u_y. This form is then convenient for integration by a symbolic algebra package.

We expand the prefactor to terms up to sixth order in the coefficients, and up to second order in both β_b and θ_e. Integration over the electron distribution function then yields the lab frame polarization matrix as a function of the bulk velocity β_b and electron temperature θ_e. Taking the trace of this matrix gave the following result for the intensity distortion:

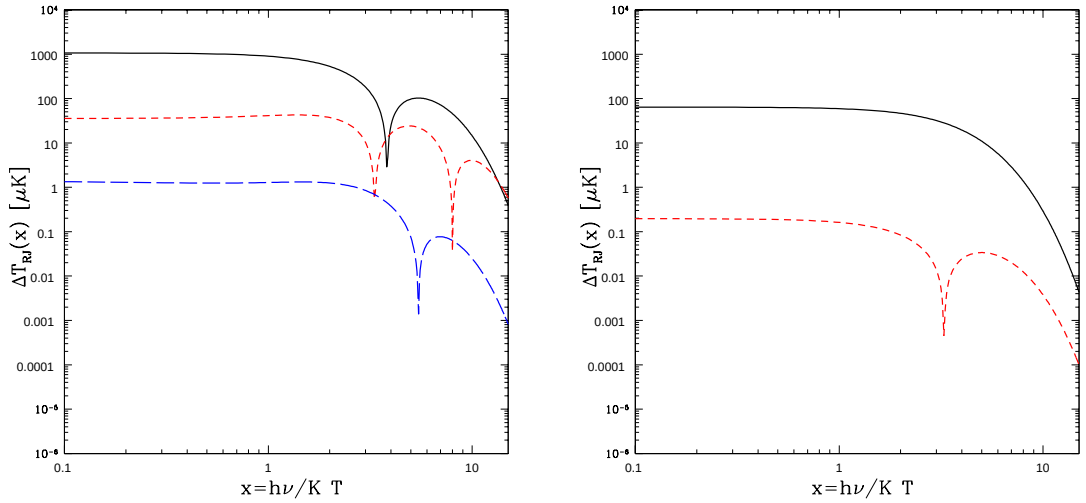
$$\begin{aligned} \frac{\Delta I}{I_0} &= \Delta\tau_{\rm T} \frac{x e^x}{e^x - 1} \big[\theta_e F_0^T(x) + \theta_e^2 F_1^T(x) \\ &\qquad + \beta_b F_0^K(\mu) + \beta_b^2 F_1^K(x,\mu) + \theta_e \beta_b F_1^{TK}(x,\mu) + \cdots \big] \ . \end{aligned} \tag{4.66}$$

Here F_0^T is the well known thermal SZ distortion piece, given in Eqn. (4.4), and F_1^T is the first relativistic correction to the thermal effect:

$$\begin{aligned} F_0^T(x) &= -4 + x \coth\left(\frac{x}{2}\right) \\ F_1^T(x) &= e^{\frac{3x}{2}} \left[5\left(-1+e^x\right)^3\right]^{-1} \Big[x \left(-235 + 77\,x^2\right) \cosh\left(\frac{x}{2}\right) \\ &\quad + x \left(235 + 7\,x^2\right) \cosh\left(\frac{3\,x}{2}\right) \\ &\quad - 8 \left(-25 + 42\,x^2 + \left(25 + 21\,x^2\right) \cosh(x)\right) \sinh\left(\frac{x}{2}\right) \Big] \ . \end{aligned} \tag{4.67}$$

The frequency dependence obtained here agrees with that obtained by Sazonov & Sunyaev (1998a,b); Itoh et al. (2000); Challinor et al. (2000).

The terms F_0^K and F_1^K are the lowest order kinematic effect and the first rela-



(a) Thermal effect (b) Kinematic effect

Figure 4-4: The frequency dependence of the thermal and kinematic SZ effects and their first relativistic corrections (as RJ brightness temperature distortions, as defined in Eqn. (4.6)). The solid lines in each plot show the lowest order effect, and the short dashed lines show the first relativistic correction. The long dashed line in plot (a) shows the "interference" term F_1^{TK}.

tivistic correction respectively:

$$\begin{aligned} F_0^K(\mu) &= \mu \ , \\ F_1^K(x,\mu) &= -1 - \mu^2 + \frac{x\,(3+11\,\mu^2)\,\coth(\frac{x}{2})}{20} \ . \end{aligned} \tag{4.68}$$

which agree with the forms in Sazonov & Sunyaev (1998a,b); Itoh et al. (2000); Challinor et al. (2000). The "interference" term between the thermal and kinematic effects is:

$$F_1^{TK}(x,\mu) = \mu\,\frac{[-45+14\,x^2+(45+7\,x^2)\,\cosh(x)-47\,x\,\sinh(x)]}{10\,\sinh^2\,(x/2)} \ . \tag{4.69}$$

The thermal and kinematic effects, their relativistic corrections, and the interference term are plotted for representative cluster parameters in Fig. 4-4. These were computed for a cluster with electrons at temperature $k_B T_e = 10$ keV, a bulk flow velocity $V_b = 1000$ km s^{-1} at an angle cosine $\mu = 1/\sqrt{2}$ to the line of sight, and an optical depth to scattering of $\tau_{\rm T} = 0.01$. (Note that the dips in the curves are zero crossings).

Computing the polarization magnitude of the final lab frame matrix, we find:

$$\Pi(x,\mu) = \Delta\tau_{\rm T}\,\beta_b^2\,\left(1-\mu^2\right)\left[F_0^P(x)+F_1^P(x)\theta_e+O(\theta_e^2)\right] \ , \tag{4.70}$$

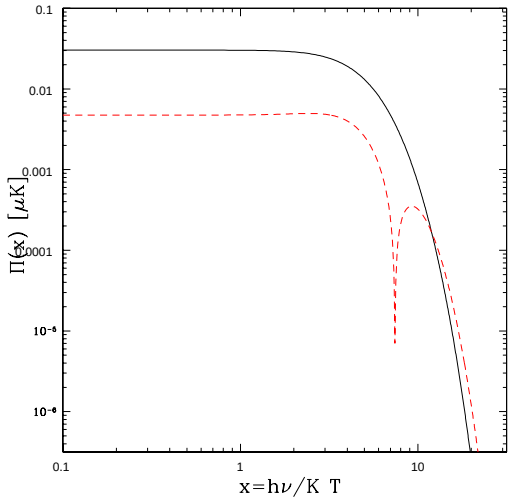


Figure 4-5: Polarization magnitude generated by scattering of CMB monopole (solid line), and the first relativistic correction (dashed line), as RJ brightness temperature distortions, in the case of a concentrated cluster with electrons at temperature $k_B T_e = 10$ keV, a bulk flow velocity $V_b = 1000$ km s^{-1} perpendicular to the line of sight, and an optical depth to scattering of $\tau_{\rm T} = 0.01$.

where

$$F_0^P(x) \;=\; \frac{e^x\,(1+e^x)\,x^2}{20\,(-1+e^x)^2}\;, \tag{4.71}$$

and

$$F_1^P(x) \;=\; \frac{e^{\frac{5\,x}{2}}\,x^2}{10\,(-1+e^x)^4}\left[\left(-4+11\,x^2\right)\,\cosh\left(\frac{x}{2}\right)+\left(4+x^2\right)\,\cosh\left(\frac{3\,x}{2}\right)\right.$$
$$\left.-8\,x\,\left(3\,\sinh\left(\frac{x}{2}\right)+\sinh\left(\frac{3x}{2}\right)\right)\right]\;. \tag{4.72}$$

As $\theta_e \to 0$ this reduces to the cold electron result Eqn. (4.41). The frequency dependence of these results for a cluster with typical parameters is shown in Fig. 4-5.

4.3 Double scattering effects

Here we discuss the polarization generated by double scattering of the CMB in a galaxy cluster. The first scattering generates a radiation field with the SZ distortions described in §4.2, which in general is anisotropic. On a further scattering of this radiation field, the quadrupole component of the anisotropy generates polarization, which is of order $\theta_e \tau_{\rm T}^2$ and $V_b \tau_{\rm T}^2$ for the anisotropy associated with the thermal and kinematic effects respectively (for a cluster with average optical depth $\tau_{\rm T}$).

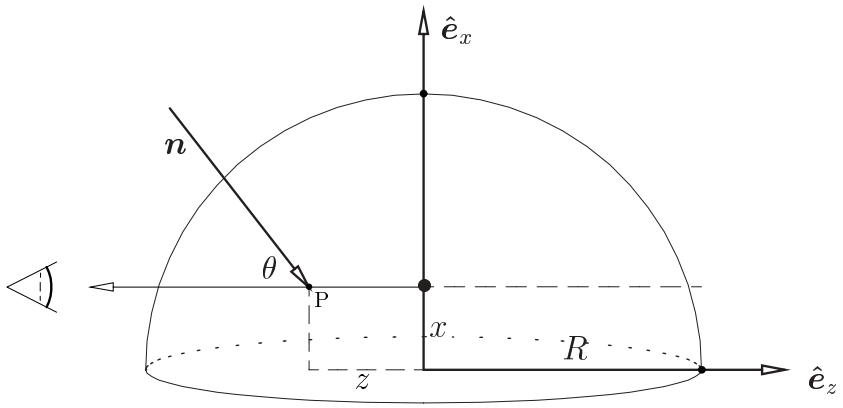


Figure 4-6: The coordinate system used to describe the generation of polarization by double Thomson scattering.

We now derive expressions for the Stokes parameters of the double scattered radiation field for a general cluster gas density and velocity distribution. The scattering geometry is illustrated in Fig. 4-6. The line of sight through the cluster is taken to be in the $-\boldsymbol{z}$ direction. The observed polarization matrix is then a function of the projected coordinates $\boldsymbol{x} = (x, y)$. The electrons have number density n_e, temperature θ_e and bulk velocity $\boldsymbol{V}_b = c\boldsymbol{\beta}_b$, which are all functions of position in the cluster. We assume the incident singly scattered radiation is unpolarized. We consider only the anisotropy introduced by the lowest order thermal and kinematic SZ effects.

First, consider the anisotropic part of the radiation field due to the thermal SZ effect incident at the point P indicated on the diagram. To lowest order in the electron temperature, this is given by the line integral of the intensity distortions contributed by a first scattering of the CMB by electrons distributed along the incident ray:

$$\frac{\Delta I_T(\boldsymbol{n}, \nu)}{I_0} = \left(\frac{\tilde{x} e^{\tilde{x}}}{e^{\tilde{x}} - 1} \right) F_0^T(\tilde{x})\, \sigma_T \int_{\boldsymbol{n}} dl\; n_e \theta_e \;, \tag{4.73}$$

where $\int_{\boldsymbol{n}}$ denotes the line integral along the incident ray $\boldsymbol{n}$ passing through P, with dl the distance element along the ray, and we use $\tilde{x} \equiv h\nu / k_B T_{\rm CMB}$ to distinguish from the cluster projected coordinate vector $\boldsymbol{x}$.

For example if the cluster is spherical with uniform density, the intensity of the singly scattered radiation radiation field incident on the point P along the incident photon direction $\boldsymbol{n}$ is proportional to the distance to the edge of the sphere along direction $-\boldsymbol{n}$ from P. Thus there is an anisotropy of the local radiation field at a given point in the cluster due to the differing path lengths to the sphere boundary along the rays converging on this point. The frequency dependence F_0^T is given in Eqn. (4.67).

Similarly, the anisotropy in the singly scattered radiation field produced by the

kinematic SZ effect yields the incident intensity at point P

$$\frac{\Delta I_K(\boldsymbol{n},\nu)}{I_0} = \left(\frac{\tilde{x}e^{\tilde{x}}}{e^{\tilde{x}}-1}\right)\sigma_T \int_{\boldsymbol{n}} dl\ n_e\,\boldsymbol{\beta}_b\cdot\boldsymbol{n}\ . \tag{4.74}$$

The double scattered radiation field can now be obtained from the kinetic equation with a gain term of the form given in Eqn. (3.44), for scattering of unpolarized radiation in the electron rest frame and Thomson limit, provided we ignore the motion of the scattering electrons at P. This approach is sufficient for calculation of the lowest order double scattering effect. We thus obtain the following intensity polarization matrix of the radiation scattered into direction $\boldsymbol{n'}$ at point P:

$$\begin{aligned} I_{ij}(\nu,\boldsymbol{n}_s) &= \frac{3}{16\pi}\Delta\tau \int d\Omega\ \Delta I(\boldsymbol{n},\nu)\,[(\delta_{ij}-n_in_j) \\ &\quad -n'_in'_j[1+(\boldsymbol{n'}\cdot\boldsymbol{n})^2]+(n'_in_j+n'_jn_i)(\boldsymbol{n'}\cdot\boldsymbol{n})\big]\ , \end{aligned} \tag{4.75}$$

where $d\Omega$ is the solid angle element about $\boldsymbol{n}$, $\Delta\tau$ is the optical depth of the scattering electrons at point P, and ΔI stands for either ΔI_T or ΔI_K.

The total observed intensity matrix at the observer is given by setting $\boldsymbol{n'}=-\boldsymbol{z}$, and integrating along the line of sight. It is convenient to express the result in terms of a Stokes basis along the line of sight direction. We choose the basis $\boldsymbol{e}_1=\boldsymbol{x},\ \boldsymbol{e}_2=\boldsymbol{y}$. Using polar coordinates in which $\boldsymbol{n}=(\cos\varphi\sin\theta,\sin\varphi\sin\theta,\cos\theta)$, we obtain:

$$\begin{aligned} Q'(\boldsymbol{x}) &= I_{xx}-I_{yy} = -\frac{3\sigma_T}{16\pi}\int n_e(z)dz\int d\Omega\ \Delta I(\boldsymbol{n},\nu)\cos(2\varphi)\sin^2\theta\ d\Omega\ , \\ U'(\boldsymbol{x}) &= -2I_{xy} = \frac{3\sigma_T}{16\pi}\int n_e(z)dz\int d\Omega\ \Delta I(\boldsymbol{n},\nu)\sin(2\varphi)\sin^2\theta\ d\Omega\ . \end{aligned} \tag{4.76}$$

This result agrees with that obtained by Sazonov & Sunyaev (1999). We computed the double scattering effects numerically in the case of a spherical cluster of radius R with uniform electron density n_e. We will just briefly outline the numerical procedure. The electron density, temperature and bulk velocity of a homogeneous approximately spherical cluster were defined on a cubic grid (with 64^3 points), and a square array of bins on the x-y face set up. Then a numerical integration of Eqns. (4.76) was performed by stepping along the line of sight parallel to the z-axis passing through each bin in increments Δz, computing the Stokes parameters of the radiation scattered into the line of sight at each element Δz. At each point along the line of sight, a large number of rays were shot to the edge of the cube along which the intensity anisotropies Eqn. (4.73) and Eqn. (4.74) were integrated. Line integrals were performed by numerical interpolation of the grid point values.

In Fig. 4-7, we plot the polarization magnitude generated by double scattering of the thermal SZ component in a homogeneous spherical cluster, as a function of the distance along a projected diameter of the sphere. The polarization magnitude is divided by the dimensionless electron temperature θ_e and τ_C^2, where τ_C is the characteristic optical depth ($n_e\sigma_{\rm T}R$). For typical cluster parameters, the maximum polarization magnitude has a temperature $\Pi\approx 0.2\mu$K. The double scattering component

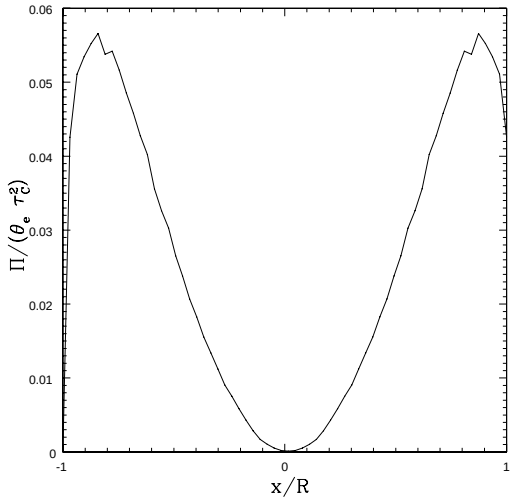


Figure 4-7: Polarization magnitude generated by double scattering of the thermal SZ distortion in a spherical cluster of radius R with uniform electron density n_e.

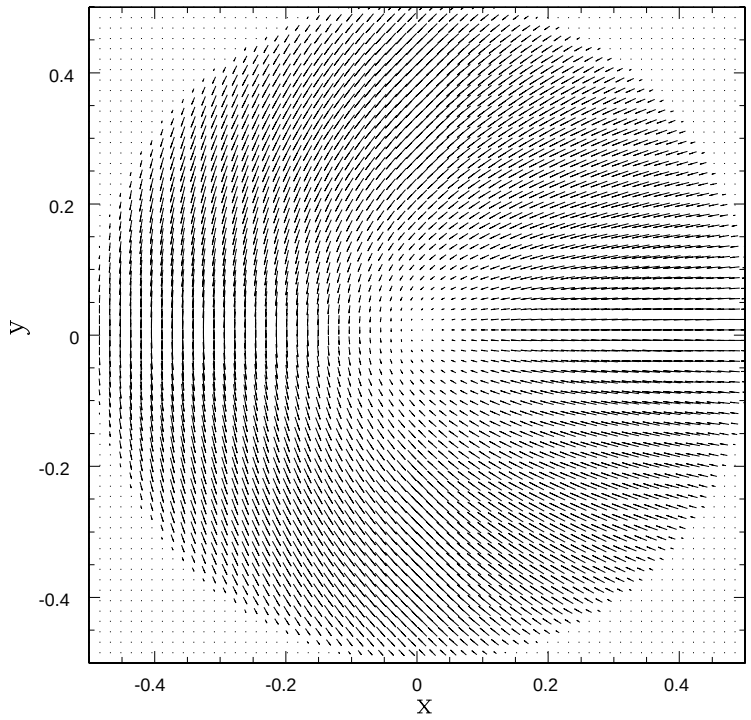
of the polarization tends to dominate over the effects due to single scattering. The polarization magnitude falls to zero at the center as it obviously must by symmetry. The polarization vector field in this case is, by symmetry, either radial or tangential depending on whether the frequency is above or below the zero of the thermal SZ distortion.

In Fig. 4-8 we show the polarization vector field and magnitude produced by double scattering of the kinematic SZ component, again in the case of a homogeneous spherical cluster, with a uniform bulk velocity in the (negative) x-direction (perpendicular to the line of sight). The headless arrows in the plot show the plane of polarization, and their length is proportional to the polarization magnitude. In this case, the polarization vector field depends on the direction of the bulk velocity on the sky. It is perhaps not intuitively obvious that the polarization magnitude should fall to zero at the cluster center, but this is indeed the case as is apparent from the figure. This was noted previously by Sazonov & Sunyaev (1999).

4.4 Monte Carlo simulation of the SZE

In this section we describe a general Monte Carlo procedure which yields the unpolarized and polarized SZ distortions. This has been implemented in the work of (Molnar & Birkinshaw, 1999; Sazonov & Sunyaev, 1998b,a), but they did not consider the procedure for polarized photons.

Monte Carlo simulations can be used to solve radiative transfer problems by generating large numbers of simulated photons at specified sources and allowing them to propagate, scatter, and be absorbed in a manner consistent with the known probability distributions for these processes, recording parameters of interest for each



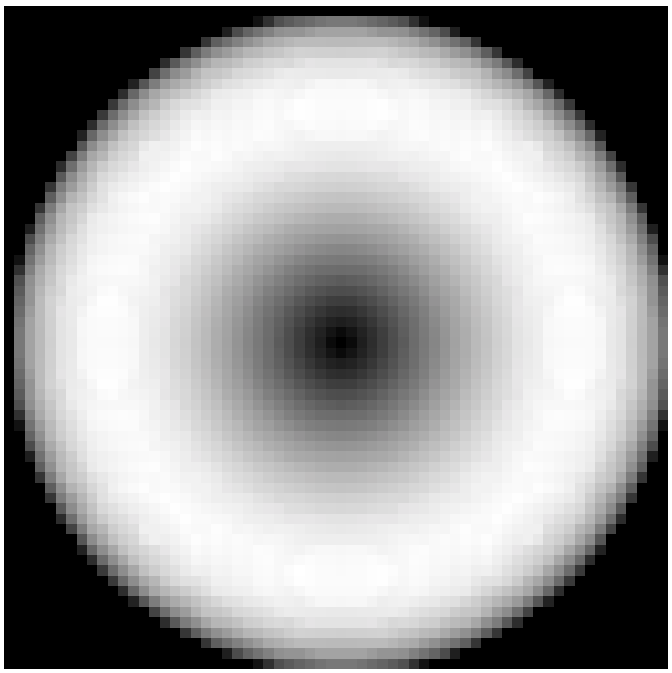


Figure 4-8: Polarization field (with projected coordinates x and y in units of the cluster diameter) of the double scattered component generated in a homogeneous cluster with a constant bulk velocity along the x-axis in the figure (towards negative x), and its magnitude (linear scale).

photon as the simulation proceeds (see for example Yusef-Zadeh et al. (1984); Code & Whitney (1995); Wood et al. (2001); Baes & Dejonghe (2002)).

For sufficiently large numbers of photons, the mean of the tabulated quantities should approach the solution of the radiative transfer equations (for a detailed discussion of the Monte Carlo simulation of Compton scattering of unpolarized photons, without restriction to the Thomson limit, see Pozdnyakov et al. (1983) and Gorecki & Wilczewski (1984)).

To simulate the single scattering SZ effects considered in §4.2, we need to design a Monte Carlo procedure which will converge to the solution of the relativistic master equation (3.73). This is achieved as follows. We generate unpolarized photons in lab frame, with frequency drawn from the blackbody distribution function of photon energies, and a random (uniformly distributed) direction. The polarization matrix of this photon is constructed. An electron velocity is generated by working first in the frame comoving with the electron bulk velocity (given by a general velocity vector in lab). In this frame, the electron velocity magnitude is drawn from a relativistic Maxwellian distribution, and the direction selected randomly. Then the electron momentum is boosted from the comoving frame into the lab frame. Then, in order to account for the flux factor appearing in Eqn. (3.73), the sampled electron is rejected or accepted according to whether a rejection crieterion is met which simulates the effect of this factor. Once a photon and electron momentum and photon polarization matrix in lab have been generated, the photon momentum and polarization tensor are boosted into the rest frame of the electron. In this frame, the probability distribution of the direction of the scattered photon is given by the phase function for Thomson scattering, assuming the Thomson limit. In this limit, the frequency of the scattered photon may be set to that of the incident photon (in the rest frame). Once a scattered direction has been sampled from the phase function, the polarization matrix of the scattered photon is determined. The rest frame photon momentum and polarization matrix is then boosted back into lab frame, where the photon frequency, direction and polarization matrix elements are recorded.

Having described the procedure in general terms, we now explain how each of the steps is implemented numerically. First of all, we need to sample photon energies from the distribution of photon *number* (*not* energy, since the Monte Carlo procedure builds up a histogram of photon number counts in energy bins):

$$p_\nu(x)dx = \frac{1}{2\xi(3)}\frac{x^2 dx}{e^x - 1} \ . \tag{4.77}$$

To sample from this distribution we compute the cumulative function

$$F(x) = \int_0^x dx' \ p_\nu(x') \ . \tag{4.78}$$

The sampled energy x is obtained by generating a uniform random deviate [1] ξ and

[1]Note that by *uniform random deviate* we mean a random variable distributed uniformly on the interval $[0, 1]$.

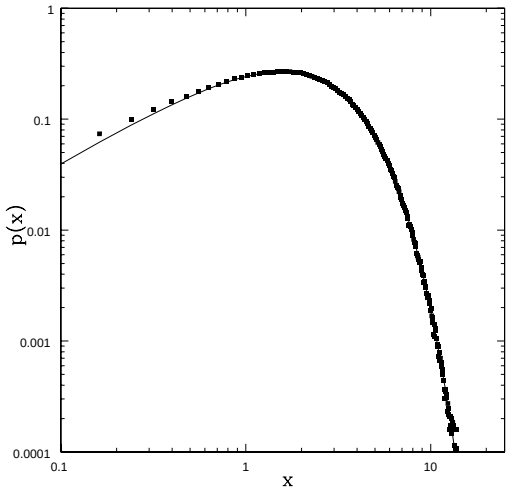


Figure 4-9: The p.d.f. of photon number obtained by inversion sampling of the photon number distribution (filled squares). The true p.d.f (solid line) is superimposed. 2×10^6 photons were generated and binned in 256 energy bins (linearly spaced) in the interval $x \in [0.005, 20]$.

solving the equation

$$F(x) = \xi \ . \tag{4.79}$$

Computation time is cut down by performing an accurate integration of (4.78) before the simulation and storing the values of F at about 100 values of x in $[0, 20]$. Then the inversion is performed during the simulation using Newton's method to solve Eqn. (4.79) with $F(x)$ constructed by interpolation between these stored points (using cubic splines). The result of this procedure is shown in Fig. 4-9. The photon direction vector in the lab frame, $\boldsymbol{n}_i$, is chosen randomly (by setting the polar angles to $\cos\theta_i = 2\xi_1 - 1$, $\phi_i = 2\pi\xi_2$ where ξ_1, ξ_2 are uniform random deviates).

Now we consider the sampling of the electron velocity from a relativistic Maxwellian. It follows from the form of the distribution function Eqn. (4.57) that the p.d.f of electron velocities in the comoving frame (in which the average electron velocity vanishes) is:

$$p_e(\beta_e)d\beta_e \ = \ \frac{\gamma^5 \beta_e^2 e^{-(\gamma-1)/\theta_e} d\beta_e}{\theta_e \ e^{1/\theta_e} \ K_2(\frac{1}{\theta_e})} \ , \tag{4.80}$$

where $\theta_e = k_B T/m_e c^2$ is the dimensionless electron temperature, $\gamma = 1/\sqrt{1-\beta_e^2}$, and K_2 is the modified Bessel function of the second order and second kind. Since we are dealing with small values of θ_e, it is useful to note that for $\theta_e \ll \frac{1}{2}$,

$$\theta_e \ e^{1/\theta_e} \ K_2\left(\frac{1}{\theta_e}\right) \approx \sqrt{\frac{\pi\theta_e^3}{2}} \ . \tag{4.81}$$

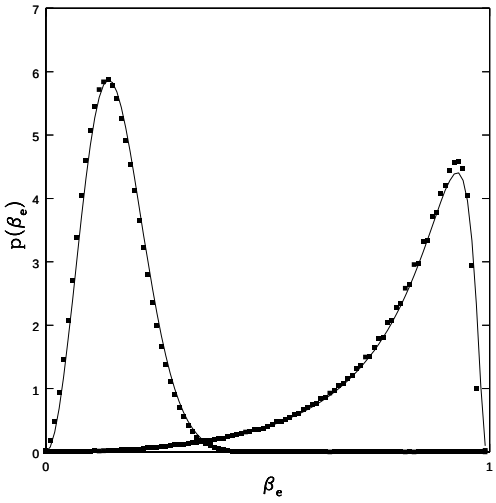


Figure 4-10: The p.d.f. of electron velocities (filled squares) obtained by inversion sampling of a relativistic Maxwellian distribution, for dimensionless electron temperatures $\Theta = 0.03$ (leftmost curve) and $\Theta = 0.5$ (rightmost curve). Here we have binned 10^6 samples in 100 β_e bins. The true p.d.f (solid line) is superimposed.

Sampling from this distribution may be achieved by a procedure similar to that used to sample from the Planck distribution. The result is illustrated in figure 4-10.

The direction of the electron velocity vector in the comoving frame, $\boldsymbol{n}'_e$, is chosen randomly. Having sampled the electron velocity parameter β_e and direction vector $\boldsymbol{n}'_e$ in the local rest frame of the electron gas, the electron velocity vector $\boldsymbol{v}_e$ in lab frame (moving with a specified bulk velocity $-\boldsymbol{V}_b$ with respect to the comoving frame) is determined by a Lorentz transformation. Then to account for the flux factor, this electron is either accepted or rejected by choosing a uniform random deviate ξ and testing whether

$$\frac{1}{2}\left(1 - \boldsymbol{v}_e \cdot \boldsymbol{n}_i\right) \leq \xi \ . \tag{4.82}$$

If this criterion is satisified, the electron is accepted, and we procede to make the transformation to the rest frame. Otherwise, it is rejected, and a new electron is sampled.

Having described how to sample the photon and electron momenta in the lab frame, we now describe some of the details required to simulate the scattering process. To apply our Thomson scattering procedure we need to transform the photon momentum and polarization matrix into a frame in which the electron is at rest before the scattering (and also at rest after the scattering, in the Thomson limit). Quantities in the electron rest frame are denoted with primes. The momentum of the photon before scattering in the electron rest frame is found by applying Lorentz

transformation matrices to obtain $p_i^{\mu'} = \Lambda^{\mu'}{}_{\mu}\, p_i^{\mu}$ where

$$\begin{aligned} \Lambda^0{}_0 &= \gamma = 1/\sqrt{1-\beta_e^2}, \qquad \Lambda^0{}_i = \Lambda^i{}_0 = -\gamma\beta_e^i \ , \\ \Lambda^i{}_j &= (\gamma-1)\frac{\beta_e^i\beta_{ej}}{\beta_e^2} + \delta^i{}_j \ . \end{aligned} \tag{4.83}$$

The polarization matrix is transformed into the electron rest frame using the matrix technique described in §2.4. The transformed polarization matrix thus obtained is used to sample the direction of the Thomson scattered photon $\boldsymbol{n}_s'$ in the electron rest frame. Transforming back yields the polarization matrix and momentum of the scattered photon in the lab frame. Given an incident photon with 4×4 normalized polarization matrix $\boldsymbol{\phi}_i(\vec{p}_i, \vec{v}_l)$ in lab frame, the explicit steps are as follows. The physical polarization matrix in the rest frame is computed, in lab frame coordinates:

$$\boldsymbol{\phi}_i(\vec{p}_i, \vec{v}_e) = \boldsymbol{P}(\vec{p}_i, \vec{v}_e)\boldsymbol{\phi}_i(\vec{p}_i, \vec{v}_l)\boldsymbol{P}(\vec{p}_i, \vec{v}_e)^T \ . \tag{4.84}$$

This is Lorentz transformed into rest frame coordinates:

$$\boldsymbol{\phi}_i'(\vec{p}_i, \vec{v}_e) = \boldsymbol{\Lambda}(\boldsymbol{\beta}_e)\boldsymbol{\phi}_i(\vec{p}_i, \vec{v}_e)\boldsymbol{\Lambda}(\boldsymbol{\beta}_e)^T \ , \tag{4.85}$$

where $\boldsymbol{\Lambda}$ has components given in equation (4.83). Rejection sampling of the scattered photon direction $\boldsymbol{n}_s'$. is now performed in the rest frame using the polarization matrix $\boldsymbol{\phi}_i'$. The p.d.f of the scattered photon direction in the rest frame, $\boldsymbol{n}_s'$, is given by the Thomson phase function $\Phi[\boldsymbol{n}_s', \boldsymbol{\phi}_i']$ (see Eqn. (3.19)). In order to sample $\boldsymbol{n}_s'$ from the phase function, we use the rejection method. The polar angles of the scattered photon are denoted (μ_s', ϕ_s'). This procedure is based on the method described by Code & Whitney (1995). They use a Stokes vector approach, and have to rotate into a particular polarization basis at each scattering. The advantage of the polarization matrix approach described here is that we do not need to specify a particular choice of basis. To apply the rejection sampling procedure, we need to find a proposal density function $Q(\mu_s', \phi_s')$ from which we can easily generate samples, such that there exists a constant c such that for all μ_s', ϕ_s':

$$cQ(\mu_s', \phi_s') > \Phi(\mu_s', \phi_s') \ , \tag{4.86}$$

and which is normalized:

$$\int \frac{d\Omega_s'}{4\pi}\, Q(\mu_s', \phi_s') = 1 \ . \tag{4.87}$$

This is most easily achieved by finding an upper bound Φ_+ on the phase function $\Phi[\boldsymbol{n}_s', \boldsymbol{\phi}_i']$ for all possible directions $\boldsymbol{n}_s'$, and setting $Q = 1,\ c = \Phi_+$ so that Q is simply uniform. We then sample from $Q(\mu_s', \phi_s')$ to obtain a trial scattered photon direction $\boldsymbol{n}_t' = (\mu_t', \phi_t')$, and generate a uniform random deviate y from the interval $[0, cQ(\mu_t', \phi_t')]$. We now evaluate $\Phi(\mu_t', \phi_t')$ (suppressing the other arguments here) and accept or reject the sample (μ_t', ϕ_t') by comparing the value of y with the value of $\Phi(\mu_t', \phi_t')$. If $y > \Phi(\mu_t', \phi_t')$, the sample (μ_t', ϕ_t') is rejected. The process is repeated until $y \leq \Phi(\mu_t', \phi_t')$. To determine an upper bound Φ_+, first we note that the phase

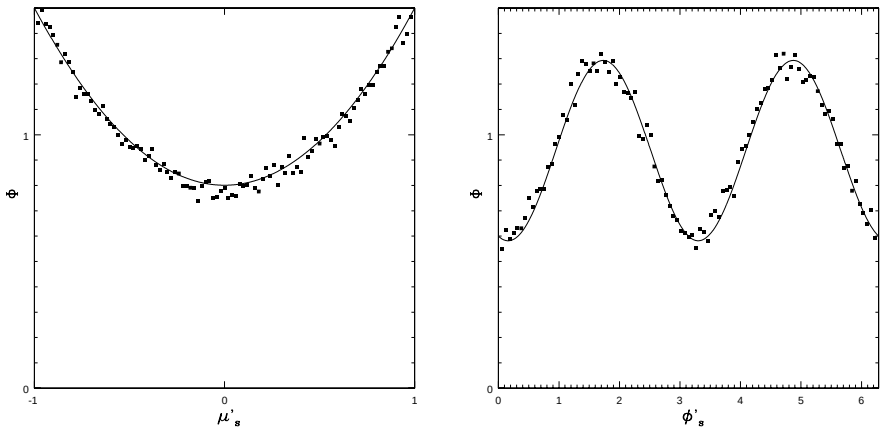


Figure 4-11: Joint p.d.f of the scattered photon angles computed by the rejection method compared to the true phase function. The incident photon has $\mu_i = 1.0$, and polarization matrix given by Eqn. (3.20), with $Q/I = 0.6,\ U/I = 0.5$. The left hand figure is a slice of the function $\Phi(\mu'_s, \phi'_s)$ at $\phi'_s = 0$, the right hand figure at $\mu'_s = 0$. Samples have been binned in a 100×100 grid in the (μ'_s, ϕ'_s) plane. In this simulation there were 10^7 samples in total.

function can always be written in the form (3.21) (in the case $V = 0$) by rotating the coordinate system with some unitary rotation matrix $\boldsymbol{R}$ so that the photon points along the polar axis. This does not affect the phase function, since

$$\mathrm{Tr}\left[\boldsymbol{\phi}'_i \boldsymbol{P}(\boldsymbol{n}'_s)\right] = \mathrm{Tr}\left[\boldsymbol{R}^{-1}\boldsymbol{\phi}'_i \boldsymbol{R}\boldsymbol{R}^{-1}\boldsymbol{P}(\boldsymbol{n}'_s)\boldsymbol{R}\right] \ . \tag{4.88}$$

Then considering the requirement for a physical polarization matrix $I^2 \geq Q^2 + U^2$, it is easy to see that the form of the phase function in equation (3.21) implies $0 \leq \Phi \leq 3/2$. Thus we take the upper bound to be $\Phi_+ = 3/2$. With this value of Φ_+, the ratio of accepted points to rejected points is $2/3$. In Figure 4-11 we show the result of sampling from the phase function for a polarized incident beam using the rejection procedure described, compared to the exact phase function. With the sampled scattered photon direction $\boldsymbol{n}'_s$, the physical (normalized) polarization matrix of the Thomson scattered beam in rest frame coordinates is determined:

$$\boldsymbol{\phi}'_s(\vec{p}_s, \vec{v}_e) = \boldsymbol{P}(\boldsymbol{n}'_s)\boldsymbol{\phi}'_i \boldsymbol{P}(\boldsymbol{n}'_s)^T \ . \tag{4.89}$$

The scattered matrix in lab frame coordinates is computed:

$$\boldsymbol{\phi}_s(\vec{p}_s, \vec{v}_e) = \boldsymbol{\Lambda}(-\boldsymbol{\beta})\boldsymbol{\phi}'_s \boldsymbol{\Lambda}(-\boldsymbol{\beta})^T \ . \tag{4.90}$$

Finally, the physical scattered matrix in lab frame coordinates is computed:

$$\boldsymbol{\phi}_s(\vec{p}_s, \vec{v}_l) = \boldsymbol{P}(\vec{p}_s, \vec{v}_l)\boldsymbol{\phi}_s(\vec{p}_s, \vec{v}_e)\boldsymbol{P}(\vec{p}_s, \vec{v}_l)^T \ . \tag{4.91}$$

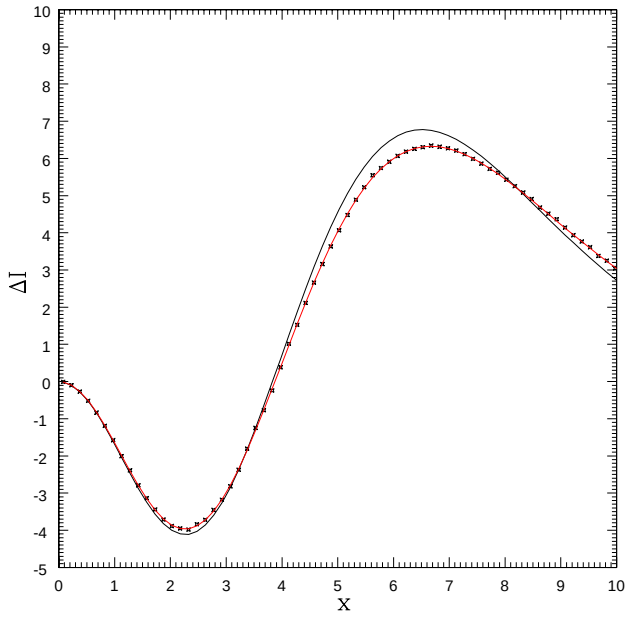


Figure 4-12: Monte Carlo simulation of thermal SZ intensity distortion, with ΔI in units of i_0. The electron temperature is $\theta_e = 0.01$ The points are the results of the simulation. The line which passes through the points is the analytic result with relativistic corrections to order θ_e^3. The other line is the Kompaneets approximation.

Then a photon count is added to the energy bin corresponding to the scattered photon energy, and a count subtracted from the energy bin of the incident photon. After a large number of scatterings have been simulated, the resulting histogram of counts is normalization to account for the total number of photons simulated and the bin widths. This yields the photon number distortion, which can easily be converted to the intensity or brightness temperature distortion.

The polarization matrix in lab frame may similarly be computed by adding the elements of the normalized polarization matrix of each scattered photon to the appropriate bin of photon angles and frequency, and normalizing after the simulation.

The result of using this procedure to compute the thermal SZ intensity distortion is shown in Fig. 4-12. In this simulation [2], the bulk velocity was set to zero, and the electron temperature sampled from a relativistic Maxwellian with temperature $\theta_e = 0.01$. The points show the result of the Monte Carlo simulation, and the curve which passes through them is the analytic result for the thermal SZ distortion correct to $O(\theta_e^3)$ (from Itoh et al. (1998)). The other curve is the lowest order thermal SZ distortion obtained from the Kompaneets equation. The Monte Carlo computation is in full agreement with the analytic results.

[2]The code used to perform this simulation was written by E. Bertschinger

Chapter 5

Cluster polarization as a cosmological probe

The CMB radiation incident on galaxy clusters has an intrinsic intensity quadrupole Q_2 created by inhomogeneity at the surface of last scattering. On a single scattering of the CMB in a galaxy cluster with typical line of sight optical depth τ_C, this anisotropy generates polarization of order $Q_2\tau_C$. In this chapter we explore how a measurement of this polarization signal would allow an estimate of the CMB quadrupole at non-zero redshift.

This is of interest because it would potentially allow us to get around the restriction of cosmic variance. To elaborate, at $z = 0$ we only have one CMB sky to observe, and the observed CMB temperature anisotropy is a projected snapshot of the perturbation to the photon distribution function at the time of last scattering. The theoretical predictions for the angular power spectrum of the CMB anisotropies on our sky are obtained by evolving the transfer functions for matter and radiation, with a Boltzmann code such as CMBFAST, and ensemble averaging over realizations of the primordial potential with a given power spectrum of Fourier modes (Bond & Efstathiou, 1987; Ma & Bertschinger, 1995; Seljak & Zaldarriaga, 1996; Zaldarriaga & Seljak, 1997). With the transfer functions one may obtain the angular power spectrum according to an observer at any epoch desired. The theoretical predictions yield the variance of the Gaussian p.d.f from which the temperature anisotropy is drawn.

But since we are limited to the data obtained from a single point of view, we can measure only $(2l + 1)$ independent real data points for each spherical harmonic mode of the CMB on our sky, to compare with the ensemble average prediction of the variance. There is thus an intrinsic fractional sample variance of the harmonic C_l of $2/(2l + 1)$ (see section §5.1), which severely limits comparison with the ensemble averaged theory at low l. This restriction is important because it limits the accuracy of measurements of the power spectrum of the inflationary potential on the largest scales. The theoretical predictions thus obtained for the CMB power spectra are fundamentally limited by this sample variance, commonly termed the *cosmic variance*.

However some information about the CMB sky at non-zero redshift may be obtained via measurements of the polarization of the radiation emitted by hot electrons in the intra-cluster gas of galaxy clusters. Thomson scattering of the $l = 2$ part of the

CMB anisotropy generates a secondary polarization anisotropy which depends on the spherical harmonic components a_{2m} as seen by an observer at the cluster, producing a characteristic pattern of polarizations on the sky (Sazonov & Sunyaev, 1999). This signal produced by a cluster is sensitive to the density perturbations on a last scattering surface different to our own. Thus in principle, this allows one to make more accurate comparison to the theoretical predictions for CMB angular power spectra at low l than allowed by the cosmic variance limit, as first pointed out by Kamionkowski & Loeb (1997).

However the observed signals are correlated if the comoving separation between the clusters is less than a certain correlation length. Thus it turns out that one cannot reduce the sample variance by more than roughly the number of regions available which produce uncorrelated signals (which grows rapidly with redshift). This correlation information is contained in the generalized correlation functions of the CMB temperature anisotropy coefficients, $\langle a_{lm}(\boldsymbol{x},\tau)a^*_{l'm'}(\boldsymbol{x}',\tau')\rangle$, which contain all of the statistical information (assuming Gaussianity) about the variation of the a_{lm} coefficients as the observation point and associated last scattering surface change. This is a generalization of the angular power spectrum.

We now outline the structure of this chapter. In §5.1 and §5.2 we derive the two-point generalized correlation functions of the spherical harmonic coefficients, assuming a Gaussian primordial perturbation spectrum, and working in the large scale limit (appropriate for working with the harmonics $l = 2$) in which only the Sachs-Wolfe effects need to be incorporated in the transfer functions. Having obtained the correlation functions of the a_{lm}, in §5.3, we derive expressions for the Stokes parameters Q, U (defined in an appropriate all-sky basis) of the CMB radiation scattered into the line-of-sight by the cluster gas, in terms of the local a_{lm} at the cluster, for a general line-of-sight. Then in §5.4 we consider the the statistical variation of these Stokes parameters with the comoving position of the cluster. We construct a simple estimator for the time dependent $l = 2$ harmonic of the angular power spectrum, $C_2(\tau)$, and show that this estimator beats the cosmic variance limit (ignoring the effect of noise). In §5.5 there is a discussion and summary. Note that we restrict the discussion to the case of a flat FRW universe throughout, for simplicity.

5.1 Generalized correlation functions

The CMB power spectra change with the comoving spatial position and redshift of the observer at a given redshift, and also with redshift along the same line of sight. Nearby observers however see similar CMB maps and power spectra - the (primary) CMB power spectra measured by separated observers are strongly correlated if the space and time separations are small enough. Here we analyze how the statistical variation with position may be characterized completely with a set of correlation functions of the photon distribution function. The coefficients at any given point are all independent, but there are spatial and temporal correlations between any pair of coefficients at different points. The joint p.d.f of any set of the coefficients is Gaussian (assuming that the initial conditions were Gaussian) and is thus described completely by the covariance function, which we derive.

Restricting attention to scalar modes in flat FRW space, the metric may be written in conformal Newtonian form

$$ds^2 = a^2(\tau)\left\{-(1+2\psi)d\tau^2 + (1-2\phi)\delta_{ij}dx^i dx^j\right\} \ . \tag{5.1}$$

In a flat FRW space, we are free to manipulate the comoving 3-vectors of events as if we were dealing with position vectors in Minkowski space (see e.g. Peacock (1999), p.71, Eqn. (3.19)). The comoving spatial 3-vector is denoted $\boldsymbol{x}$, and the radial comoving coordinate $|\boldsymbol{x}|$ denoted χ. Henceforth when computing distances and scalar products we neglect the perturbations ϕ and ψ and work in flat space; their overall effect on photon paths and spacetime separations is a higher order effect than we need not consider here.

The fractional CMB temperature perturbation $\Delta(\boldsymbol{x}, \hat{\boldsymbol{n}}, \tau) = \Delta T/T_{\rm CMB}$ is a function of the spatial position of the observer $\boldsymbol{x}$, the conformal time of observation τ, and the direction of "view" $\hat{\boldsymbol{n}}$ (Ma & Bertschinger, 1995) [1]. Of course, one can work with cosmic time, or redshift, if desired, by a change of variables. In order to expand the sky temperature in spherical harmonics $a_{lm}(\boldsymbol{x}, \tau)$, we must define a conventional polar coordinate system everywhere. It is simplest to use the convention that the polar axis is taken to be the same at each point. Then general expressions may be derived without explicitly specifying the polar coordinate system. The fractional temperature anisotropy can be expanded

$$\Delta(\boldsymbol{x}, \hat{\boldsymbol{n}}, \tau) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm}(\boldsymbol{x}, \tau)Y_{lm}(\hat{\boldsymbol{n}}) \ . \tag{5.2}$$

(Note that the a_{lm} as defined here are dimensionless). The general two point correlation of the temperature anisotropy is

$$\langle \Delta(\boldsymbol{x}, \hat{\boldsymbol{n}}, \tau)\Delta(\boldsymbol{x}', \hat{\boldsymbol{n}}', \tau')\rangle = \sum_{lm}\sum_{l'm'} C_{lml'm'}(\boldsymbol{x}, \tau, \boldsymbol{x}', \tau')Y_{lm}(\hat{\boldsymbol{n}})Y^*_{l'm'}(\hat{\boldsymbol{n}}') \ , \tag{5.3}$$

[1] Note that $\hat{\boldsymbol{n}}$ is opposite to the direction of the momentum of the observed photons.

where we have defined the generalized CMB correlation functions

$$C_{lml'm'}(\boldsymbol{x},\tau,\boldsymbol{x}',\tau') \equiv \langle a_{lm}(\boldsymbol{x},\tau)a^*_{l'm'}(\boldsymbol{x}',\tau')\rangle \ . \tag{5.4}$$

We refer to the functions with different values of l, m, l', m' as separate correlation functions. In fact the whole set of functions $C_{lml'm'}(\boldsymbol{x},\tau,\boldsymbol{x}',\tau')$ form the covariance function of the Gaussian random process from which the photon distribution function, defined at all points in space and observed in all directions, is sampled (see section 5.2). If the a_{lm} are produced by underlying Gaussian random fields, then this function contains all of the information extractable in principle from the (ensemble averaged) primary CMB. To characterize non-Gaussian processes (for example correlations of secondary anisotropies) we require higher order products. Given a set of cosmological parameters, these correlation functions may be computed using the photon transfer function which describes the physics of the propagation of CMB photons from the last scattering surface to the cluster.

We will be interested only in the case where both sets of spacetime coordinates lie on our past light cone, in order that we are computing only quantities which are directly observable. (It is possible in principle that the anisotropy seen by observers at spacetime points inside our past light cone might be obtained indirectly by observation of multiply scattered light, but we do not consider that here.) Choosing ourselves to be at the spatial origin of the comoving coordinate system, at conformal time τ_0 (the age of the universe in conformal time), the events at $\boldsymbol{x}$, $\boldsymbol{x}'$ occurred at conformal times $\tau = \tau_0 - \chi$, $\tau' = \tau_0 - \chi'$ respectively. The correlation function may therefore be written as a function of spatial variables only, $C_{lml'm'}(\boldsymbol{x},\boldsymbol{x}')$.

The generalized correlation functions obey some simple symmetry relations:

$$\begin{aligned} C_{l,-m,l',-m'}(\boldsymbol{x},\boldsymbol{x}') &= (-1)^{m+m'}C^*_{lml'm'}(\boldsymbol{x},\boldsymbol{x}') \\ C_{lml'm'}(\boldsymbol{x}',\boldsymbol{x}) &= C^*_{l'm'lm}(\boldsymbol{x},\boldsymbol{x}') \ . \end{aligned} \tag{5.5}$$

When we need to specify the coordinate system we will take the polar axis to be the $\boldsymbol{e}_z$ axis at each point, with the $\phi = 0$ plane normal to the $\boldsymbol{e}_y$ axis. Then the direction vector $\hat{\boldsymbol{n}}$ refers to the polar coordinates (θ,ϕ) about the the $\boldsymbol{e}_z$ axis. Of course, there is no physical significance in the choice of the polar axis for the spherical harmonics. If desired, one could obtain more general expressions for the spherical harmonic coefficients defined about any polar axis, using the Wigner rotation matrices $D^l_{mm'}$ (Rose, 1995), which relate coefficients defined with respect to primed and unprimed left handed coordinate systems as follows

$$a_{lm}(\boldsymbol{x},\tau) = \sum_{m'} D^l_{mm'}a'_{lm'}(\boldsymbol{x},\tau) \ . \tag{5.6}$$

Now we derive the relationship between the generalized correlation functions and the CMB transfer function. In a flat FRW space, the temperature anisotropy may be Fourier expanded in comoving wavenumber $\boldsymbol{k}$ on the three-dimensional hyper-surface

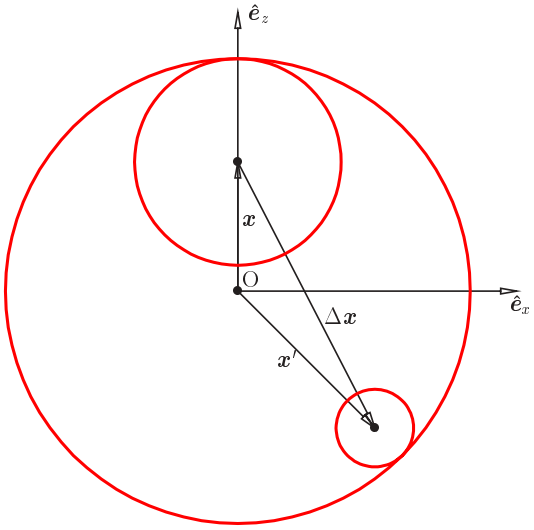


Figure 5-1: Comoving coordinate system for CMB covariance function, in flat space. All points in the plane shown lie on the past light cone of the observer at O. The outermost solid circle indicates our last scattering surface. The solid circles about the observers at positions $\boldsymbol{x}$ and $\boldsymbol{x}'$ (and conformal times $\tau = \tau_0 - |\boldsymbol{x}|$ and $\tau' = \tau_0 - |\boldsymbol{x}|$ respectively) indicate their last scattering surfaces, which are smaller since recombination occurred in the less distant past according to them. Note that any orientation of the points $\boldsymbol{x}$ and $\boldsymbol{x}'$ in space can be rotated into this plane.

of constant τ, Σ_τ,

$$\Delta(\boldsymbol{x}, \hat{\boldsymbol{n}}, \tau) = \int d^3k \; e^{i\boldsymbol{k}\,\cdot\,\boldsymbol{x}} \Delta(\boldsymbol{k}, \hat{\boldsymbol{n}}; \Sigma_\tau) \; , \tag{5.7}$$

where $\Delta(\boldsymbol{k}, \hat{\boldsymbol{n}}; \Sigma_\tau)$ is the Fourier transform associated with this hyper-surface only (Ma & Bertschinger, 1995).

Now since each Fourier mode evolves independently in linear theory, and corresponds in the case of scalar perturbations to a plane wave perturbation which has azimuthal symmetry about $\boldsymbol{k}$, $\Delta(\boldsymbol{k}, \hat{\boldsymbol{n}}; \Sigma_\tau)$ depends only on $\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{n}}$ and $|\boldsymbol{k}|$ and may therefore be expanded in Legendre polynomials:

$$\Delta(\boldsymbol{k}, \hat{\boldsymbol{n}}; \Sigma_\tau) = \sum_{l=0}^{\infty} (-i)^l (2l+1) \Delta_l(\boldsymbol{k}, \tau) P_l(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{n}}) \; . \tag{5.8}$$

The $(-i)^l$ is included by convention to be consistent with most other authors. It is convenient to use the addition theorem at this point to express the Legendre polynomials in spherical harmonics, giving

$$\Delta(\boldsymbol{x}, \hat{\boldsymbol{n}}, \tau) = 4\pi \int d^3k \; e^{i\boldsymbol{k}\,\cdot\,\boldsymbol{x}} \sum_{l=0}^{\infty} (-i)^l \; \Delta_l(\boldsymbol{k}, \tau) \sum_{m=-l}^{l} Y^*_{lm}(\hat{\boldsymbol{k}}) Y_{lm}(\hat{\boldsymbol{n}}) \; . \tag{5.9}$$

Employing the orthogonality relation for spherical harmonics $\int d\Omega\ Y^*_{lm}(\hat{\boldsymbol{n}})Y_{l'm'}(\hat{\boldsymbol{n}}) = \delta_{l'l}\delta_{m'm}$ (where $d\Omega$ is the integral over solid angle elements centered about direction $\hat{\boldsymbol{n}}$) yields

$$\begin{aligned} a_{lm}(\boldsymbol{x},\tau) &= \int d\Omega\ \Delta(\boldsymbol{x},\hat{\boldsymbol{n}},\tau)Y^*_{lm}(\hat{\boldsymbol{n}}) \\ &= (-i)^l\ 4\pi \int d^3k\ e^{i\boldsymbol{k}\,\boldsymbol{\cdot}\,\boldsymbol{x}}\Delta_l(\boldsymbol{k},\tau)Y^*_{lm}(\hat{\boldsymbol{k}})\ . \end{aligned} \tag{5.10}$$

The correlation function may now be written as

$$\begin{aligned} C_{lml'm'}(\boldsymbol{x},\boldsymbol{x}') &= (-i)^{l-l'}\ (4\pi)^2 \int d^3k\ d^3k' e^{i\boldsymbol{k}\,\boldsymbol{\cdot}\,\boldsymbol{x}} e^{-i\boldsymbol{k'}\,\boldsymbol{\cdot}\,\boldsymbol{x'}} \\ &\qquad \langle\Delta_l(\boldsymbol{k},\tau)\Delta^*_{l'}(\boldsymbol{k}',\tau')\rangle Y^*_{lm}(\hat{\boldsymbol{k}})Y_{l'm'}(\hat{\boldsymbol{k}}')\ . \end{aligned} \tag{5.11}$$

Now since the Boltzmann equation which governs the evolution of the CMB anisotropy $\Delta_l(\boldsymbol{k},\tau)$ is independent of $\hat{\boldsymbol{k}}$, the $\hat{\boldsymbol{k}}$ dependence comes entirely from the initial conditions, and we may write $\Delta_l(\boldsymbol{k},\tau)$ in terms of the CMB transfer function $\Delta_l(k,\tau)$ which is defined by (Ma & Bertschinger, 1995):

$$\Delta_l(\boldsymbol{k},\tau) = \phi_i(\boldsymbol{k})\Delta_l(k,\tau)\ , \tag{5.12}$$

where $\phi_i(\boldsymbol{k})$ is the initial potential perturbation and $\Delta_l(k,\tau)$ is real [2]. By the assumption of translational invariance $\phi_i(\boldsymbol{k})$ has a two-point correlation function of the form

$$\langle\phi_i(\boldsymbol{k})\phi^*_i(\boldsymbol{k}')\rangle = P_\phi(k)\ \delta^3(\boldsymbol{k}-\boldsymbol{k}')\ , \tag{5.13}$$

where $P_\phi(k)$ is the power spectrum of the primordial (post-inflationary) gravitational potential fluctuations. Then we may write

$$\langle\Delta_l(\boldsymbol{k},\tau)\Delta^*_{l'}(\boldsymbol{k}',\tau')\rangle = \Delta_l(k,\tau)\Delta_{l'}(k',\tau')\ P_\phi(k)\ \delta^3(\boldsymbol{k}-\boldsymbol{k}')\ . \tag{5.14}$$

If we evaluate $C_{lml'm'}(\boldsymbol{x},\boldsymbol{x}')$ for $\boldsymbol{x} = \boldsymbol{x}'$ (and $\tau = \tau'$), the covariance matrix is diagonal and the familiar orthogonality relation follows

$$C_{lml'm'}(\boldsymbol{x},\boldsymbol{x}) = (4\pi)^2\delta_{l'l}\delta_{m'm}\int k^2dk\ \Delta^2_l(k,\tau)P_\phi(k) \equiv \delta_{l'l}\delta_{m'm}C_l(\tau)\ , \tag{5.15}$$

where $C_l(\tau) \equiv \langle|a_{lm}(\boldsymbol{x},\tau)|^2\rangle$ is the ensemble average of the l harmonic of the CMB power spectrum according to an observer at conformal time τ. Thus at any given point all of the a_{lm} are independent random variables. Using the addition theorem,

[2] Reality of $\Delta_l(k,\tau)$ is ensured by the choice of including the $(-i)^l$ factor in Eqn. (5.8); see Ma & Bertschinger (1995).

we obtain the usual real-space angular correlation function, at any epoch

$$\langle \Delta(\boldsymbol{x}, \hat{\boldsymbol{n}}, \tau) \Delta(\boldsymbol{x}, \hat{\boldsymbol{n}}', \tau) \rangle \;=\; \frac{1}{4\pi} \sum_l (2l+1) C_l(\tau) P_l(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}') \; . \tag{5.16}$$

With $\boldsymbol{x} \neq \boldsymbol{x}'$, $\tau \neq \tau'$ we obtain a more general expression for the correlation function:

$$\begin{aligned} C_{lml'm'}(\boldsymbol{x}, \boldsymbol{x}') \;&=\; (-i)^{l-l'} \; (4\pi)^2 \int d^3k \; e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \\ &\qquad \Delta_l(k,\tau) \Delta_{l'}(k,\tau') \; P_\phi(k) \; Y^*_{lm}(\hat{\boldsymbol{k}}) \; Y_{l'm'}(\hat{\boldsymbol{k}}) \; . \end{aligned} \tag{5.17}$$

The symmetries stated in equation (5.5) may be verified from this expression.

We may perform the angular part of the d^3k integral in Eqn. (5.17) by expanding the plane wave piece in spherical Bessel functions (valid in flat space). Defining $\Delta\boldsymbol{x} \equiv \boldsymbol{x} - \boldsymbol{x}'$, $\Delta\hat{\boldsymbol{x}} \equiv (\boldsymbol{x} - \boldsymbol{x}')/|\boldsymbol{x} - \boldsymbol{x}'|$,

$$\begin{aligned} \exp\left(i\boldsymbol{k} \cdot \Delta\boldsymbol{x}\right) \;&=\; \sum_{l''=0}^{\infty} (2l''+1) i^{l''} P_{l''}(\hat{\boldsymbol{k}} \cdot \Delta\hat{\boldsymbol{x}}) j_{l''}(k|\Delta\boldsymbol{x}|) \\ &=\; 4\pi \sum_{l''=0}^{\infty} i^{l''} j_{l''}(k|\Delta\boldsymbol{x}|) \sum_{m''=-l''}^{l''} Y^*_{l''m''}(\hat{\boldsymbol{k}}) Y_{l''m''}(\Delta\hat{\boldsymbol{x}}) \; , \end{aligned} \tag{5.18}$$

where we used the addition theorem to separate the $\hat{\boldsymbol{k}}$ and $\Delta\hat{\boldsymbol{x}}$ dependence. Then the correlation function becomes

$$\begin{aligned} C_{lml'm'}(\boldsymbol{x}, \boldsymbol{x}') \;&=\; (4\pi)^3 \int k^2 dk \; \Delta_l(k,\tau) \Delta_{l'}(k,\tau') \; P_\phi(k) \; (-i)^{l-l'} \; (-1)^m \\ &\times \sum_{l''m''} i^{l''} \; j_{l''}(k|\Delta\boldsymbol{x}|) \; Y_{l''m''}(\Delta\hat{\boldsymbol{x}}) \int d\Omega_k \; Y_{l,-m}(\hat{\boldsymbol{k}}) \; Y_{l'm'}(\hat{\boldsymbol{k}}) Y^*_{l''m''}(\hat{\boldsymbol{k}}) \; . \end{aligned} \tag{5.19}$$

The angular integral of the product of three spherical harmonics is expressible in terms of the Clebsch-Gordon coefficients $C(l_1, l_2, l_3 \; ; m_1, m_2, m_3)$ (see e.g. Rose (1995)),

$$\begin{aligned} &\int d\Omega_k \; Y_{lm}(\hat{\boldsymbol{k}}) \; Y_{l'm'}(\hat{\boldsymbol{k}}) Y^*_{l''m''}(\hat{\boldsymbol{k}}) \\ &\qquad = \sqrt{\frac{(2l+1)(2l'+1)}{4\pi(2l''+1)}} \; C(l,l',l'' \; ;0,0,0) C(l,l',l'' \; ;m,m',m'') \end{aligned} \tag{5.20}$$

The Clebsch-Gordon coefficients are non-zero only if $m + m' = m''$, and l, l', l'' satisfy the *triangle condition* that l'' be equal to one of $l + l', l + l' - 1, \cdots, |l - l'|$. The sum

over l'' therefore reduces to a finite sum. We have

$$C_{lml'm'}(\boldsymbol{x},\boldsymbol{x}') = \ (-i)^{l-l'}\ (-1)^m\ (4\pi)^3\sqrt{\frac{(2l+1)(2l'+1)}{4\pi}}$$
$$\times\sum_{l''=|l-l'|}^{l+l'} C(l,l',l''\,;0,0,0)C(l,l',l''\,;-m,m',m'-m)$$
$$\times\frac{i^{l''}}{\sqrt{2l''+1}}\ K_{l,l',l''}(|\Delta\boldsymbol{x}|,\tau,\tau')\ Y^*_{l'',m'-m}(\Delta\hat{\boldsymbol{x}})\ , \tag{5.21}$$

where all of the physical information is contained in the kernel

$$K_{l,l',l''}(|\Delta\boldsymbol{x}|,\tau,\tau') \equiv \int k^2dk\ \Delta_l(k,\tau)\Delta_{l'}(k,\tau')\ P_\phi(k)\ j_{l''}(k|\Delta\boldsymbol{x}|)\ . \tag{5.22}$$

A useful fact is that the coefficient $C(l,l',l''\,;0,0,0)$ vanishes unless $l+l'+l''$ is even. Thus for example if $l=l'$, only even values of l'' occur in the sum.

Now we review the calculation of the cosmic variance of the spherical harmonic coefficients. First, let us establish the statistical properties of the coefficients a_{lm}. Assuming Gaussianity of the primordial potential, it follows that all of the various N-point joint probability distribution functions for the a_{lm} are (complex) multivariate Gaussians, that is

$$P\left[a_{l_1m_1}(\boldsymbol{x}_1),\cdots,a_{l_Nm_N}(\boldsymbol{x}_N)\right] \ = \ \frac{1}{\sqrt{(2\pi)^N|\det[\boldsymbol{R}]|}}$$
$$\times\exp\left[-\frac{1}{2}\sum_{ij}^{N} a_{l_im_i}(\boldsymbol{x}_i)R^{-1}_{ij}a^*_{l_jm_j}(\boldsymbol{x}_j)\right]\ , \tag{5.23}$$

where the covariance matrix $\boldsymbol{R}$ is given by (note that each index labels both the set of values l_i,m_i and also the point $\boldsymbol{x}_i$ in three dimensional space):

$$R_{ij} = \langle a_{l_im_i}(\boldsymbol{x}_i)a^*_{l_jm_j}(\boldsymbol{x}_j)\rangle = C_{l_im_il_jm_j}(\boldsymbol{x}_i,\boldsymbol{x}_j)\ . \tag{5.24}$$

It is easiest to see why the a_{lm} coefficients have a multivariate Gaussian distribution by considering all the cumulants $\langle a_{l_1m_1}(\boldsymbol{x}_1)\cdots a^*_{l_Mm_M}(\boldsymbol{x}_M)\rangle_C$ for any integer M (and any permutation of conjugated and un-conjugated factors). Using equation (5.11) we can express these in terms of integrals of products of M factors of ϕ_i. Then assuming ϕ_i is a Gaussian random field, and applying Wick's theorem to obtain the ensemble average of the product, one finds that all of the cumulants vanish except for the the two point functions. The unique distribution which has this property is the multivariate Gaussian above.

Thus given a cosmological model we have the p.d.f of the ensemble from which the $a_{lm}(\boldsymbol{x})$ are drawn, and the associated ensemble average angular power spectrum harmonics $C_l(\tau)$. Then given a set of observations $a^o_{lm}(\boldsymbol{x})$, these ensemble average predictions are compared to the observed quantities $C^o_l(\boldsymbol{x}) = \sum_{m=-1}^{l}|a^o_{lm}(\boldsymbol{x})|^2/(2l+1)$

(in practice instrument noise is taken into account). The mean square difference between the experimental observations and the ensemble averaged theory is then characterized by the cosmic variance

$$\langle [C_l^o(\boldsymbol{x}) - C_l(\boldsymbol{x})]^2 \rangle = \langle C_l^o(\boldsymbol{x})^2 \rangle - C_l(\tau)^2 \ . \tag{5.25}$$

Expanding this gives

$$\langle C_l^o(\boldsymbol{x})^2 \rangle \ = \ \sum_{mm'} (2l+1)^{-2} \langle |a_{lm}^o(\boldsymbol{x})|^2 |a_{lm'}^o(\boldsymbol{x})|^2 \rangle \ . \tag{5.26}$$

To evaluate the right hand side, we need to compute the ensemble average of the product of four a_{lm}'s:

$$\begin{aligned} \langle a_{lm_1}(\boldsymbol{x},\tau) a^*_{lm_2}(\boldsymbol{x},\tau) a_{l'm_3}(\boldsymbol{x}',\tau') a^*_{l'm_4}(\boldsymbol{x}',\tau') \rangle = (4\pi)^4 \int d^3k_1 d^3k_2 d^3k_3 d^3k_4 \\ e^{i(\boldsymbol{k}_1 - \boldsymbol{k}_2)\cdot \boldsymbol{x}} e^{i(\boldsymbol{k}_3 - \boldsymbol{k}_4)\cdot \boldsymbol{x}'} \langle \Delta_l(\boldsymbol{k}_1,\tau) \Delta_l^*(\boldsymbol{k}_2,\tau) \Delta_{l'}(\boldsymbol{k}_3,\tau') \Delta_{l'}^*(\boldsymbol{k}_4,\tau') \rangle \\ Y^*_{lm_1}(\hat{\boldsymbol{k}}_1) Y_{lm_2}(\hat{\boldsymbol{k}}_2) Y^*_{l'm_3}(\hat{\boldsymbol{k}}_3) Y_{l'm_4}(\hat{\boldsymbol{k}}_4) \ . \end{aligned} \tag{5.27}$$

The ensemble average in the integrand may be expanded

$$\begin{aligned} \langle \Delta_l(\boldsymbol{k}_1,\tau) \Delta_l^*(\boldsymbol{k}_2,\tau) \Delta_{l'}(\boldsymbol{k}_3,\tau') \Delta_{l'}^*(\boldsymbol{k}_4,\tau') \rangle \ &= \ \Delta_l(k_1,\tau) \Delta_l(k_2,\tau) \Delta_{l'}(k_3,\tau') \Delta_{l'}(k_4,\tau') \\ &\quad \langle \phi_i(\boldsymbol{k}_1) \phi_i^*(\boldsymbol{k}_2) \phi_i(\boldsymbol{k}_3) \phi_i^*(\boldsymbol{k}_4) \rangle \ . \end{aligned} \tag{5.28}$$

Assuming that the primordial potential is a Gaussian random field, the expectation value of the potentials may be expanded via Wick's theorem as

$$\begin{aligned} \langle \phi_i(\boldsymbol{k}_1) \phi_i^*(\boldsymbol{k}_2) \phi_i(\boldsymbol{k}_3) \phi_i^*(\boldsymbol{k}_4) \rangle = \langle \phi_i(\boldsymbol{k}_1) \phi_i^*(\boldsymbol{k}_2) \rangle \langle \phi_i(\boldsymbol{k}_3) \phi_i^*(\boldsymbol{k}_4) \rangle \\ + \langle \phi_i(\boldsymbol{k}_1) \phi_i(\boldsymbol{k}_3) \rangle \langle \phi_i^*(\boldsymbol{k}_2) \phi_i^*(\boldsymbol{k}_4) \rangle \\ + \langle \phi_i(\boldsymbol{k}_1) \phi_i^*(\boldsymbol{k}_4) \rangle \langle \phi_i^*(\boldsymbol{k}_2) \phi_i(\boldsymbol{k}_3) \rangle \ . \end{aligned} \tag{5.29}$$

Then on using $\langle \phi_i(\boldsymbol{k}) \phi_i^*(\boldsymbol{k}') \rangle = P_\phi(k) \ \delta^3(\boldsymbol{k} - \boldsymbol{k}')$, $\langle \phi_i(\boldsymbol{k}) \phi_i(\boldsymbol{k}') \rangle = P_\phi(k) \ \delta^3(\boldsymbol{k} + \boldsymbol{k}')$, due to the second term in the equation above we get quantities like $Y_{lm}(-\hat{\boldsymbol{k}})$ inside the integrand. In polar coordinates this is equal to $Y_{lm}(\pi - \theta, \phi + \pi) = (-1)^l Y_{lm}(\theta, \phi)$ (the transformation of the associated Legendre functions under reflection through the origin gives $(-1)^{l+m}$, and the complex exponential gives another $(-1)^m$). Thus we get

$$\begin{aligned} \langle a_{lm_1}(\boldsymbol{x},\tau) a^*_{lm_2}(\boldsymbol{x},\tau) a_{l'm_3}(\boldsymbol{x}',\tau') a^*_{l'm_4}(\boldsymbol{x}',\tau') \rangle = \delta_{m_1 m_2} \delta_{m_3 m_4} C_l(\tau) \ C_{l'}(\tau') \\ + C_{lm_1 l'm_4}(\boldsymbol{x},\boldsymbol{x}') C^*_{lm_2 l'm_3}(\boldsymbol{x},\boldsymbol{x}') + C_{lm_1 l',-m_3}(\boldsymbol{x},\boldsymbol{x}') C^*_{lm_2 l',-m_4}(\boldsymbol{x},\boldsymbol{x}') \ . \end{aligned} \tag{5.30}$$

With $m_1 = m_2 = m$, and $m_3 = m_4 = m'$, and allowing $l \neq l'$ for the moment, we find

$$\begin{aligned} \langle a_{lm}(\boldsymbol{x},\tau) a^*_{lm}(\boldsymbol{x},\tau) a_{l'm'}(\boldsymbol{x}',\tau') a^*_{l'm'}(\boldsymbol{x}',\tau') \rangle \\ = C_l(\tau) \ C_{l'}(\tau') + |C_{lml'm'}(\boldsymbol{x},\boldsymbol{x}')|^2 + |C_{lml',-m'}(\boldsymbol{x},\boldsymbol{x}')|^2 \ . \end{aligned} \tag{5.31}$$

Now setting $l' = l$, $\boldsymbol{x} = \boldsymbol{x}'$, and using $C_{lml'm'}(\boldsymbol{x}, \boldsymbol{x}) = \delta_{ll'}\delta_{mm'}C_l(\tau)$, we obtain for the cosmic variance associated with each harmonic, $\langle C_l^o(\boldsymbol{x})^2\rangle$

$$\langle C_l^o(\boldsymbol{x})^2\rangle - C_l(\tau)^2 \;=\; \frac{2}{2l+1}\, C_l(\tau)^2 \;. \tag{5.32}$$

This quantity captures how much we can expect the measured power spectra to differ from the ensemble average.

5.2 Transfer functions

To compute $C_{lml'm'}$ for a given cosmological model we need the CMB transfer function $\Delta_l(k,\tau)$. On the large angular scales accessible via the polarization technique, the only significant effects responsible for the temperature anisotropy which need to be included in the transfer function are the Sachs-Wolfe (SW) and integrated Sachs-Wolfe (ISW) effects (Sachs & Wolfe, 1967). The SW effect is the anisotropy due to the temperature fluctuations on the last scattering surface, and the associated time dilation effect. The ISW effect arises because, at late times, as the universe is making the transition from the matter dominated phase into the vacuum dominated phase, the fluctuations in the gravitational potential - on scales still in the linear regime - are still evolving with redshift. As photons fall into and climb out of this time changing potential they are red-shifted and thus a temperature anisotropy is generated.

We first consider the transfer function of the SW effect. This is computed by ignoring the physics on scales comparable to the acoustic horizon at the time of recombination, and retaining only the large scale effects. In this limit, the anisotropy is produced solely by the variation in potential ϕ (and the consequent gravitational redshift and time dilation effects on the photons) and photon density δ_γ across the last scattering shell, ignoring the small scale acoustic waves which give rise to the acoustic peaks in the angular power spectrum. Using the line-of-sight integration method (Seljak & Zaldarriaga, 1996), the SW temperature anisotropy is given in real space by

$$\Delta(\boldsymbol{x}, \boldsymbol{n}, \tau) \;=\; \int_0^\tau d\chi'\, \dot{\zeta}(\tau - \chi')\, \left[\frac{1}{4}\delta_\gamma(\chi'\hat{\boldsymbol{n}}) + \phi(\chi'\hat{\boldsymbol{n}})\right] \;. \tag{5.33}$$

Here χ' is the comoving distance measured along the past light cone of the observer at $(\boldsymbol{x}, \tau)$, in the direction $\boldsymbol{n}$. The visibility function is defined by $\zeta(\tau) \equiv e^{-\tau_{\rm T}(\tau)}$, with Thomson optical depth $\tau_{\rm T}(\tau) = \int_0^\tau d\chi' a(\tau - \chi') n_e(\tau - \chi')\sigma_{\rm T}$ (here and elsewhere a dot means a derivative with respect to conformal time τ).

In the Sachs-Wolfe approximation (valid on scales much larger than the acoustic horizon), and assuming adiabatic initial conditions, a perturbative analysis of the equations of motion shows that $\delta_\gamma = -\frac{8}{3}\phi$, and that in Fourier space the evolution of the potential is given by $\phi(\boldsymbol{k}, \tau) = \frac{9}{10}\phi_i(\boldsymbol{k})$ (see, for example, Bashinsky & Bertschinger (2002); Padmanabhan (2003)). The factor of 9/10 accounts for the evolution of the transfer function between radiation and matter domination (in the

case of adiabatic initial conditions). Decomposing the plane waves (working here in flat space) into spherical waves, we obtain

$$\Delta_l(k,\tau) \;=\; \frac{3}{10}\int_0^\tau d\chi'\;\dot{\zeta}(\tau-\chi')\;j_l(k\chi')\;. \tag{5.34}$$

The visibility function ζ contains the physics of recombination. It rises rapidly during recombination from 0 to 1, with derivative sharply peaked about the time of recombination, τ_r (Dodelson & Jubas, 1995). The effect of the finite thickness of the last scattering shell can only influence the radiation field on rather small scales, so for times well after recombination, and for low l, we may assume that recombination occurred instantaneously at time τ_r. In this limit we may take ζ to be a delta function centered on τ_r, and the transfer function reduces to

$$\Delta_l(k,\tau) = \frac{3}{10} j_l[k(\tau-\tau_r)]\;. \tag{5.35}$$

Taking the usual power law form $P_\phi(k) = Ak^{n-4}$ ($n=1$ gives the scale-invariant Harrison-Zeldovich spectrum) with an arbitrary amplitude A (with dimensions of $(c/H_0)^{n-1}$), with the transfer function in equation (5.35) the integral in equation (5.15) may be done analytically, yielding the well known Sachs-Wolfe expression for the angular power spectrum at low l and $n<3$ (see for example Peacock (1999)):

$$\begin{aligned} C_l(\tau) &= A(4\pi)^2\left(\frac{3}{10}\right)^2\int_0^\infty k^2dk\;k^{n-4}j_l^2\left[k(\tau-\tau_r)\right] \\ &= A(4\pi)^2\left(\frac{3}{10}\right)^2\pi 2^{n-4}\frac{\Gamma(3-n)\Gamma(\frac{2l+n-1}{2})}{\Gamma^2(\frac{4-n}{2})\Gamma(\frac{2l+5-n}{2})}\;(\tau-\tau_r)^{1-n}\;. \end{aligned} \tag{5.36}$$

Note that if $n=1$, this expression has no time dependence. This is a manifestation of the scale invariance property of the $n=1$ case.

In the general case including the ISW effect (and assuming a flat universe), the CMB transfer function is given by a generalization of Eqn. (5.33), the line of sight integral:

$$\begin{aligned} \Delta(\boldsymbol{x},\boldsymbol{n},\tau) &= \int_0^\tau d\chi'\;\dot{\zeta}(\tau-\chi')\;\left[\frac{1}{4}\delta_\gamma(\chi'\hat{\boldsymbol{n}})+\phi(\chi'\hat{\boldsymbol{n}})\right] \\ &\qquad + \int_0^\tau d\chi'\;\zeta(\tau-\chi')\;2\dot{\phi}(\chi'\hat{\boldsymbol{n}})\;. \end{aligned} \tag{5.37}$$

In linear theory, the growth of the amplitude of the potential perturbations is governed by the growth function $D_+(\tau)$ of the dark matter perturbation

The evolution of the potential perturbation in the adiabatic case is then given by $\phi(\boldsymbol{k},\tau) = \frac{9}{10}\phi_i(\boldsymbol{k})D_+(\tau)/a(\tau)$.

In the case of a flat universe with only non-relativistic matter and vacuum energy, the solution for the growth function, normalized to $D_+ = a$ at early times, has the

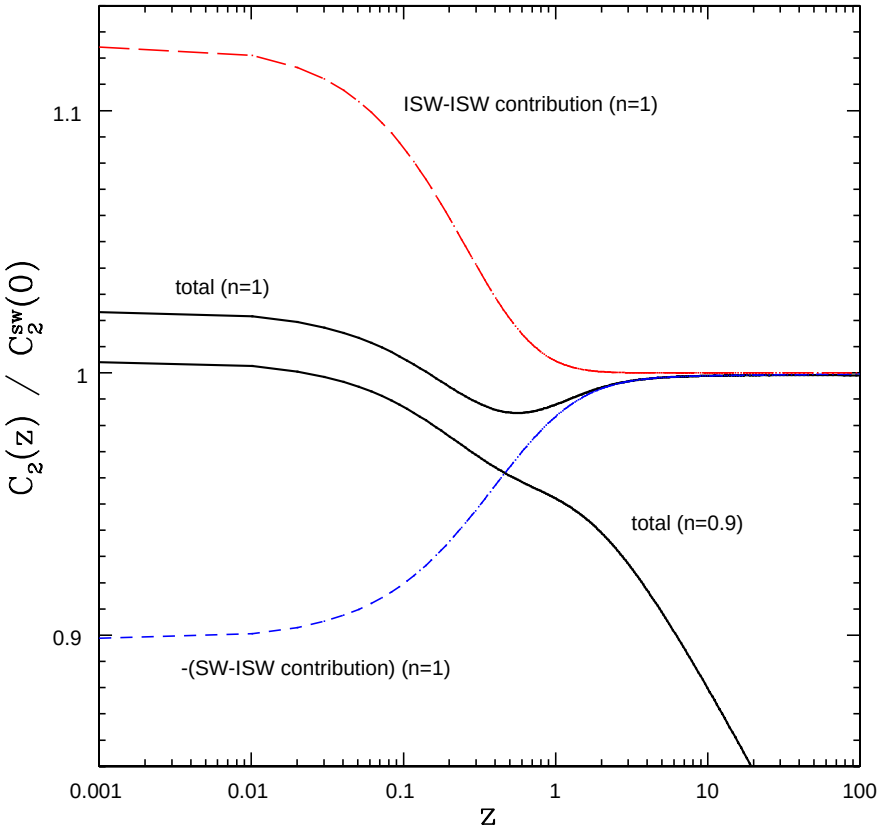


Figure 5-2: A plot of $C_2(z)$, including both the SW and ISW effects, normalized to the value of the SW contibution at $z = 0$. The growth function was computed in a ΛCDM model with $(\Omega_\Lambda, \Omega_m) = (0.7, 0.3)$. The upper solid black curve is for the case of power spectral index $n = 1$, and the lower solid black curve for $n = 0.9$. The dashed upper and lower curves curves show the contributions to the $n = 1$ case from the ISW-ISW term and the SW-ISW interference term respectively. These two tend to cancel. (Note that the interference term is negative, and its magnitude is plotted here).

simple form (Heath, 1977; Seljak, 1995):

$$D_+(\tau) = \frac{5}{2}\Omega_m \frac{\sqrt{\Omega_m + \Omega_\Lambda a(\tau)^3}}{a(\tau)^{3/2}} \int_0^{a(\tau)} X^{3/2}(a')\, da$$
$$\text{where} \qquad X(a) = \frac{a}{\Omega_m + \Omega_\Lambda a^3} \ . \tag{5.38}$$

In the instantaneous recombination approximation, this leads to the following CMB transfer function

$$\Delta_l(k,\tau) = \frac{3}{10} j_l\left[k(\tau - \tau_r)\right] + \frac{9}{5}\int_0^{\tau-\tau_r} d\chi'\ j_l(k\chi')\ \left.\frac{\partial}{\partial\tau}\frac{D_+}{a}\right|_{\tau-\chi'} , \tag{5.39}$$

where the time derivative in the integrand is evaluated at time $\tau - \chi'$.

In the kernel $K_{ll'l''}$, the product of two transfer function appears. Thus there are three terms, a contribution entirely from the SW effect, an "interference" contribution from both the SW and ISW effects, and a contribution entirely from the ISW effect. Since the ISW part of the transfer function is usually negative, the interference term tends to cancel the third term. This is illustrated in Fig. 5-2, which shows the redshift dependence of the CMB harmonic C_2 in a ΛCDM model with $(\Omega_\Lambda, \Omega_m) = (0.7, 0.3)$, and various values of the power spectral index n.

5.3 Scattering of CMB quadrupole

In this section we consider the generation of polarization by scattering of the intrinsic CMB quadrupole from electrons in a galaxy cluster, which we idealize as a concentrated clump of stationary electrons with a Thomson optical depth $\tau_{\rm C}$. The optical depth will vary as the line of sight moves across the cluster, but here we ignore this and take the cluster to lie on a specific angular point on the sky. We will assume that the optical depths of each cluster are uncorrelated with each other.

The incident CMB has an intrinsic intensity quadrupole Q_2 created by inhomogeneity at the surface of last scattering. A single scattering of the quadrupole anisotropy generates polarization of order $Q_2\tau_{\rm C}$. We compute the Stokes parameters of the scattered radiation as a function of the cluster position on the sky. The Stokes parameters of the radiation scattered into the line of sight to the cluster are functions of the quadrupole anisotropy in the local CMB radiation field at the cluster.

This is characterized by the coefficients $a_{2m}(\boldsymbol{x})$ of the spherical harmonic expansion of the fractional temperature anisotropy of the radiation field, which are functions of the spatial position of the cluster in comoving coordinates denoted $\boldsymbol{x}$ (see §5.1 for a description of this coordinate system). The direction vector of the line of sight from the observer to the cluster is $\hat{\boldsymbol{x}}$. The coordinate system used is illustrated in Fig. 5-3.

In terms of the general set of coefficients $a_{lm}(\boldsymbol{x})$, we may write the brightness temperature of the incident CMB radiation field at the cluster as a function of the

direction vector $\hat{\boldsymbol{n}}$ of the incoming photon *as viewed* from the cluster:

$$I(\hat{\boldsymbol{n}}, \boldsymbol{x}) = T_{\rm CMB}(\tau) \sum_{lm} a_{lm}(\boldsymbol{x}) Y_{lm}(\hat{\boldsymbol{n}}) \ , \tag{5.40}$$

where $T_{\rm CMB}$ is given in Eqn. 4.1, and $\tau = \tau_0 - |\boldsymbol{x}|$ is the conformal time of the scattering events. Here $I(\hat{\boldsymbol{n}}, \boldsymbol{x})$ is the brightness temperature (do not confuse it with the specific intensity). Thus since the primary anisotropy is blackbody, there is no frequency dependence in I. We may suppress the frequency dependence of the (brightness temperature) polarization matrix and associated Stokes parameters of the scattered photon also, since there is no energy transfer to the photons in the case of scattering from stationary electrons. Since the primary CMB anisotropies are blackbody, the intensity matrix of the scattered photon is also blackbody. We assume that the incident CMB radiation is unpolarized, which is sufficient to compute the lowest order polarization signal generated by the quadrupole anisotropy (there are relativistic corrections to the effect described here in the case of a cluster with a peculiar velocity with respect to the CMB, as discussed by Challinor & van Leeuwen (2002), which turn out to be negligible).

The brightness temperature polarization matrix I_{ij} of the radiation scattered into the line of sight can be computed using the kinetic equation with a gain term of the form given in Eqn. (3.44), for scattering of unpolarized radiation in the electron rest frame and Thomson limit. The matrix I_{ij} in direction $-\hat{\boldsymbol{x}}$, in terms of the brightness temperature of the CMB incident at the cluster, is given by

$$\begin{aligned} I_{ij}(\boldsymbol{x}) \ &= \ \frac{3\tau_C}{16\pi} \int d\Omega' \ I(\hat{\boldsymbol{n}}, \boldsymbol{x}) \left[(\delta_{ij} - \hat{n}_i \hat{n}_j) \right. \\ &\quad \left. - \hat{x}_i \hat{x}_j [1 + (\hat{x} \cdot \hat{n})^2] + (\hat{x}_i \hat{n}_j + \hat{x}_j \hat{n}_i)(\hat{x} \cdot \hat{n}) \right] \ , \end{aligned} \tag{5.41}$$

where $d\Omega'$ is the solid angle element about the $\hat{\boldsymbol{n}}$ direction. [3]

We now define a polarization basis to define the Stokes parameters of the radiation incident at the observer from a cluster in any direction on the sky. We denote the polarization basis vectors as $\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2$ and leave these unspecified for the moment. The Stokes parameters measured in the $\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2$ basis at our position due to scattering in the cluster at comoving position $\boldsymbol{x}$ are then (note that the terms in the second line of equation (5.41) do not contribute):

$$\begin{aligned} I(\boldsymbol{x}) + Q(\boldsymbol{x}) \ &= \ 2 I_{ij}(\boldsymbol{x}) \ \hat{e}_{1,i} \hat{e}_{1,j} \ , \\ I(\boldsymbol{x}) - Q(\boldsymbol{x}) \ &= \ 2 I_{ij}(\boldsymbol{x}) \ \hat{e}_{2,i} \hat{e}_{2,j} \ , \\ U(\boldsymbol{x}) \ &= \ 2 I_{ij}(\boldsymbol{x}) \ \hat{e}_{1,i} \hat{e}_{2,j} \ . \end{aligned} \tag{5.42}$$

On substitution of Eqn. (5.40) we find

[3] The primed solid angle element $d\Omega'$ is associated with the unprimed direction vector $\hat{\boldsymbol{n}}$ since we wish to reserve $d\Omega$ for the polar angles of the cluster on the sky.

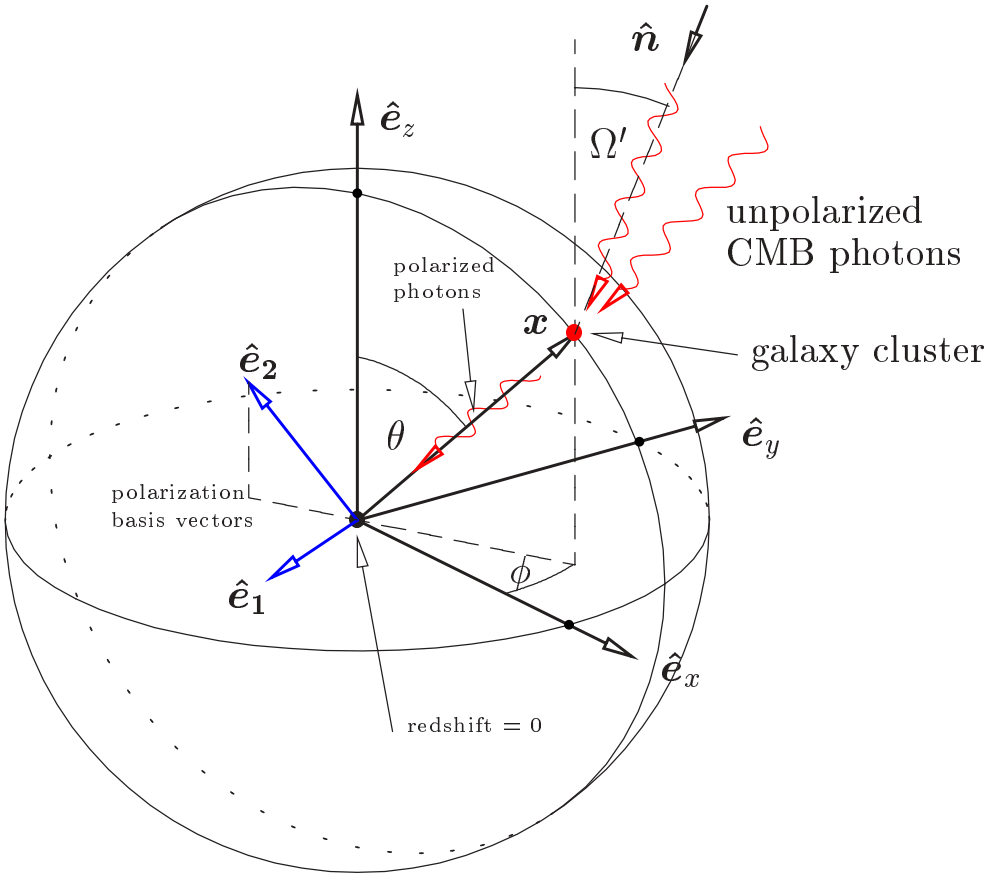


Figure 5-3: Illustrates the coordinate system used to describe the generation of polarization by Thomson scattering. The CMB incident on a cluster at comoving position $\boldsymbol{x}$, which we approximate as unpolarized, is Thomson scattered and re-radiated by free electrons in the cluster, producing partially polarized radiation scattered into the line of sight. At the observer position (redshift $z = 0$) the radiation is decomposed into Stokes parameters with the polarization basis vectors $\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2$ indicated (defined in equation (5.52). The CMB radiation incident on the cluster at $\boldsymbol{x}$ is decomposed into spherical harmonics defined with respect to polar coordinates θ and ϕ. Only the $l = 2$ harmonics of the incident radiation field generate polarization.

$$
\begin{aligned}
Q(\boldsymbol{x}) &= \frac{3\tau_C}{16\pi} T_{\rm CMB}(\tau) \sum_{lm} a_{lm}(\boldsymbol{x}) \int d\Omega' \; Y_{lm}(\hat{\boldsymbol{n}}) \; \left[(\hat{\boldsymbol{e}}_2 \cdot \hat{\boldsymbol{n}})^2 - (\hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{n}})^2 \right] \; , \\
U(\boldsymbol{x}) &= -\frac{3\tau_C}{8\pi} T_{\rm CMB}(\tau) \sum_{lm} a_{lm}(\boldsymbol{x}) \int d\Omega' \; Y_{lm}(\hat{\boldsymbol{n}}) (\hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{n}}) (\hat{\boldsymbol{e}}_2 \cdot \hat{\boldsymbol{n}}) \; . \qquad (5.43)
\end{aligned}
$$

To perform the angular integral we need to expand the integrands in equations (5.43) in spherical harmonics by expressing $\hat{\boldsymbol{n}}$ in polar coordinates. In polar coordinates about the $\hat{\boldsymbol{z}}$ axis, the Cartesian components of the direction vectors are taken to be:

$$
\begin{aligned}
\hat{\boldsymbol{n}} &= (\hat{n}_x, \hat{n}_y, \hat{n}_z) = (\sin\theta' \cos\phi', \sin\theta' \sin\phi', \cos\theta') \; , \\
\hat{\boldsymbol{x}} &= (\hat{x}_x, \hat{x}_y, \hat{x}_z) = (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta) \; . \qquad (5.44)
\end{aligned}
$$

We use the following trick [4]. Any spherical harmonic can be expanded in terms of the complex quantities $(z_1, z_2, z_3) = (\sin\theta' e^{i\phi'}, \sin\theta' e^{-i\phi'}, \cos\theta')$.

In terms of these functions we may write $\hat{\boldsymbol{n}} = ((z_1 + z_2)/2, i(z_2 - z_1)/2, z_3)$. Note that $z_2 = z_1^*, \; z_3^* = z_3, \; z_1 z_2 = 1 - z_3^2$. It is also convenient to define

$$
\begin{aligned}
\hat{\boldsymbol{e}}_+ &\equiv \hat{\boldsymbol{e}}_x + i \hat{\boldsymbol{e}}_y \; , \\
\hat{\boldsymbol{e}}_- &\equiv \hat{\boldsymbol{e}}_x - i \hat{\boldsymbol{e}}_y \; . \qquad (5.45)
\end{aligned}
$$

Then for instance we have

$$
\hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{n}} = \frac{z_1}{2} \; \hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{e}}_- + \frac{z_2}{2} \; \hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{e}}_+ + z_3 \; \hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{e}}_z \qquad (5.46)
$$

The $l = 2$ spherical harmonics may be written as functions quadratic in the z's as follows:

$$
\begin{aligned}
Y_{2,0} &= \sqrt{\frac{5}{4\pi}} \left(\frac{3}{2} z_3^2 - \frac{1}{2} \right) , \\
Y_{2,1} &= -\sqrt{\frac{15}{8\pi}} \; z_1 z_3, \qquad Y_{2,-1} = \sqrt{\frac{15}{8\pi}} \; z_2 z_3 \\
Y_{2,2} &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \; z_1^2, \qquad Y_{2,-2} = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \; z_2^2 \; . \qquad (5.47)
\end{aligned}
$$

Then expanding the integrands in equation (5.43) in the z's using expressions like (5.46) and comparing [5] with the expressions for Y_{2m} above, we find the following

[4] See for example the discussion of spherical harmonics in Byron & Fuller (1992)

[5] In performing this calculation, it is necessary to use the relation $z_1 z_2 = 1 - z_3^2$ to eliminate one of the coefficients (z_1, z_2, z_3).

manifestly real result for the integrands of Q and U:

$$\begin{aligned}(\hat{\boldsymbol{e}}_2\cdot\hat{\boldsymbol{n}})^2-(\boldsymbol{e}_1\cdot\hat{\boldsymbol{n}})^2 &= \sqrt{\frac{8\pi}{5}}\sum_{m=-2}^{2}Q_m(\hat{\boldsymbol{x}})Y_{2,m}(\hat{\boldsymbol{n}})\ ,\\ -2(\hat{\boldsymbol{e}}_1\cdot\hat{\boldsymbol{n}})(\boldsymbol{e}_2\cdot\hat{\boldsymbol{n}}) &= \sqrt{\frac{8\pi}{5}}\sum_{m=-2}^{2}U_m(\hat{\boldsymbol{x}})Y_{2,m}(\hat{\boldsymbol{n}})\ .\end{aligned}\tag{5.48}$$

The coefficients Q_m, U_m appearing in this expression are the following functions of the arbitrary polarization basis vectors chosen by the observer, which in turn are functions of the cluster direction on the sky (so Q_m, U_m are written as functions of the cluster direction vector $\hat{\boldsymbol{x}}$, which will become explicit once a polarization basis is chosen):

$$\begin{aligned}Q_0(\hat{\boldsymbol{x}}) &= -\frac{1}{\sqrt{2}}\left[(\hat{\boldsymbol{e}}_1\cdot\hat{\boldsymbol{e}}_z)^2-(\boldsymbol{e}_2\cdot\hat{\boldsymbol{e}}_z)^2\right]\ ,\\ Q_1(\hat{\boldsymbol{x}}) &= \frac{1}{\sqrt{3}}\left[(\hat{\boldsymbol{e}}_-\cdot\hat{\boldsymbol{e}}_1)(\hat{\boldsymbol{e}}_z\cdot\hat{\boldsymbol{e}}_1)-(\hat{\boldsymbol{e}}_-\cdot\hat{\boldsymbol{e}}_2)(\boldsymbol{e}_z\cdot\hat{\boldsymbol{e}}_2)\right]\ ,\\ Q_2(\hat{\boldsymbol{x}}) &= \frac{1}{2\sqrt{3}}\left[(\hat{\boldsymbol{e}}_-\cdot\hat{\boldsymbol{e}}_2)^2-(\hat{\boldsymbol{e}}_-\cdot\hat{\boldsymbol{e}}_1)^2\right]\ ,\end{aligned}\tag{5.49}$$

and

$$\begin{aligned}U_0(\hat{\boldsymbol{x}}) &= -\sqrt{2}(\hat{\boldsymbol{e}}_1\cdot\hat{\boldsymbol{e}}_z)(\boldsymbol{e}_2\cdot\hat{\boldsymbol{e}}_z)\ ,\\ U_1(\hat{\boldsymbol{x}}) &= \frac{1}{\sqrt{3}}\left[(\hat{\boldsymbol{e}}_-\cdot\boldsymbol{e}_1)(\hat{\boldsymbol{e}}_z\cdot\hat{\boldsymbol{e}}_2)+(\hat{\boldsymbol{e}}_-\cdot\hat{\boldsymbol{e}}_2)(\boldsymbol{e}_z\cdot\hat{\boldsymbol{e}}_1)\right]\ ,\\ U_2(\hat{\boldsymbol{x}}) &= -\frac{1}{\sqrt{3}}(\hat{\boldsymbol{e}}_-\cdot\boldsymbol{e}_1)(\hat{\boldsymbol{e}}_-\cdot\hat{\boldsymbol{e}}_2)\ .\end{aligned}\tag{5.50}$$

Also, we define the quantities with negative m by the relations

$$\begin{aligned}Q_{-m}(\hat{\boldsymbol{x}}) &\equiv (-1)^m Q_m^*(\hat{\boldsymbol{x}})\ ,\\ U_{-m}(\hat{\boldsymbol{x}}) &\equiv (-1)^m U_m^*(\hat{\boldsymbol{x}})\ .\end{aligned}\tag{5.51}$$

Here we see explicitly that only the quadrupole anisotropy in the incident unpolarized radiation field generates a polarization signal, which is a well known fact that is not obvious in the matrix formalism. [6]

We specialize now to a particular choice of of polarization basis vectors. Defining a direction vector $\hat{\boldsymbol{l}}$, then a suitable choice is

$$\hat{\boldsymbol{e}}_1=\frac{\hat{\boldsymbol{x}}\times\hat{\boldsymbol{e}}_z}{\sqrt{1-\mu^2}}\ ,\quad \hat{\boldsymbol{e}}_2=\frac{\hat{\boldsymbol{e}}_z-\mu\hat{\boldsymbol{x}}}{\sqrt{1-\mu^2}}\ ,\tag{5.52}$$

[6] Since the integrand of the kinetic equation is quadratic in $\hat{n}_i$, it *is* obvious that some combination of the monopole, dipole and quadrupole generates polarization, but not that *only* the quadrupole contributes.

where $\mu = \hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{e}}_z = \cos\theta$, so that the Stokes parameters Q, U are defined with respect to the plane containing the $\hat{\boldsymbol{e}}_z$ axis and the photon direction $-\hat{\boldsymbol{x}}$ (see Fig. 5-3).

Then the coefficients Q_m, U_m are

$$\begin{aligned} Q_0(\hat{\boldsymbol{x}}) &= \frac{1}{\sqrt{2}} \sin^2\theta \ , \\ Q_1(\hat{\boldsymbol{x}}) &= \frac{1}{\sqrt{3}} \cos\theta \sin\theta e^{-i\phi} \ , \\ Q_2(\hat{\boldsymbol{x}}) &= \frac{1}{2\sqrt{3}} (1+\cos^2\theta) e^{-2i\phi} \ , \end{aligned} \tag{5.53}$$

and

$$\begin{aligned} U_0(\hat{\boldsymbol{x}}) &= 0 \ , \\ U_1(\hat{\boldsymbol{x}}) &= \frac{i}{\sqrt{3}} \sin\theta e^{-i\phi} \ , \\ U_2(\hat{\boldsymbol{x}}) &= \frac{i}{\sqrt{3}} \cos\theta e^{-2i\phi} \ , \end{aligned} \tag{5.54}$$

where θ is the polar angle between $\boldsymbol{x}$ and the $\hat{\boldsymbol{e}}_z$ axis and ϕ is the azimuthal angle between the projection of $\boldsymbol{x}$ on the $(\hat{\boldsymbol{e}}_x, \hat{\boldsymbol{e}}_y)$ plane and the $\hat{\boldsymbol{e}}_x$ axis. The Stokes parameters may now finally be expressed as a linear combination of the $a_{2m}(\boldsymbol{x})$, with polar axis $\hat{\boldsymbol{e}}_z$. Angular integration picks out the five coefficients a_{2m} of the primary anisotropy:

$$\begin{aligned} Q(\boldsymbol{x}) &= \tau_C P_0 \sum_{m=-2}^{2} Q_m(\hat{\boldsymbol{x}}) a_{2m}(\boldsymbol{x}) \ , \\ U(\boldsymbol{x}) &= \tau_C P_0 \sum_{m=-2}^{2} U_m(\hat{\boldsymbol{x}}) a_{2m}(\boldsymbol{x}) \ , \end{aligned} \tag{5.55}$$

where $P_0 \equiv \frac{3}{4\sqrt{10\pi}} T_{\rm CMB}(\tau)$. Note that this depends on conformal time, but the fractional distortion in the Stokes parameters is redshift independent (as is the case for all of the SZ effects). Note also that

$$Q_m + iU_m = \frac{4}{3}\sqrt{\frac{3\pi}{5}}\ {}_2Y^*_{2m} \ . \tag{5.56}$$

Thus we may write

$$Q(\boldsymbol{x}) \pm iU(\boldsymbol{x}) = -\frac{6}{20}\sqrt{\frac{2}{3}} \tau_C T_{\rm CMB}(\tau) \sum_{m=-2}^{2} {}_{\pm 2}Y_{2m}(\hat{\boldsymbol{x}})\ a_{2m}(\boldsymbol{x}) \ , \tag{5.57}$$

where ${}_sY_{lm}$ are the "spin-weighted spherical harmonics" of spin s (see e.g. Hu & White (1997)). This form is familiar from the all-sky calculations of the CMB polarization

Boltzmann hierarchy (Zaldarriaga & Seljak, 1997).

It is useful to give here the transformation from our notation to the so-called "Berkeley notation" for the CMB quadrupole, which is the conventional notation used in the literature. This is defined with reference to great circle of the plane of the Milky Way on the sky (the galactic equator). Aligning our z–axis with the normal to this plane, the galactic coordinates are related to our polar coordinate system by the replacement in $\boldsymbol{x}$ of $\cos\theta$ by $\sin b$, where b is the usual galactic latitude with respect to the galactic plane, and the replacement of ϕ by l, the galactic longitude. The dependence of the quadrupole part $I_Q(l,b)$ of the CMB temperature anisotropy is written in terms of five (real) quantities Q_a ($a \in \{1,\cdots,5\}$), where (Lubin et al., 1983; Bennett et al., 1992; Kogut et al., 1996):

$$\begin{aligned} I_Q(l,b) &= Q_1(3\sin^2 b - 1)/2 + Q_2 \sin 2b\ \cos l \\ &\quad + Q_3 \sin 2b\ \sin l + Q_4 \cos^2 b\ \cos 2l + Q_5 \cos^2 b\ \sin 2l\ . \end{aligned} \tag{5.58}$$

Expressing this in terms of the $l = 2$ spherical harmonics, we find that the relationship between the complex a_{lm} and the Q_i is as follows:

$$\begin{aligned} a_{20} &= \sqrt{\frac{4\pi}{5}} Q_1\ , \\ a_{21} &= -\sqrt{\frac{8\pi}{15}}\,(Q_2 - iQ_3)\ , \\ a_{22} &= \sqrt{\frac{8\pi}{15}}\,(Q_4 - iQ_5)\ . \end{aligned} \tag{5.59}$$

The quadrupole can also be written in terms of a symmetric traceless tensor Q_{ij}, $I_Q(l,b) = \sum_{ij} x_i x_j\, Q_{ij}$, where (Sazonov & Sunyaev, 1999),

$$Q_{ij} = \begin{bmatrix} Q_4 - Q_1/2 & Q_5 & Q_2 \\ Q_5 & -Q_4 - Q_1/2 & Q_3 \\ Q_2 & Q_3 & Q_1 \end{bmatrix}\ . \tag{5.60}$$

In terms of the Q_a, the Stokes parameters generated by scattering of the local quadrupole become:

$$\begin{aligned} Q(\boldsymbol{x}) &= \frac{3}{20} T_{\rm CMB}(\tau) \Big[Q_1 \sin^2\theta - \frac{2}{3}\sin 2\theta\,(Q_2\cos\phi + Q_3 \sin\phi) \\ &\qquad \frac{2}{3}(1+\cos^2\theta)\,(Q_4 \cos 2\phi - Q_5 \sin 2\phi) \Big]\ . \\ U(\boldsymbol{x}) &= \frac{\sqrt{3}}{5} T_{\rm CMB}(\tau) \Big[-\sin\theta\,(Q_2 \sin\phi + Q_3\cos\phi) \\ &\qquad + \cos\theta\,(Q_4\cos 2\phi - Q_5 \cos 2\phi) \Big]\ . \end{aligned} \tag{5.61}$$

Finally in this section, we quote the following properties of Q_m, U_m, which are needed in §5.4 (these are derived using the explicit forms in Eqns. (5.53) and (5.54)):

$$\sum_{m=-2}^{2} |Q_m(\hat{\boldsymbol{x}}_i)|^2 = \sum_{m=-2}^{2} |U_m(\hat{\boldsymbol{x}}_i)|^2 = \frac{2}{3} \ ,$$
$$\sum_{m=-2}^{2} Q_m(\hat{\boldsymbol{x}}_i) U_m^*(\hat{\boldsymbol{x}}_i) = 0 \ , \tag{5.62}$$

and

$$\sum_{m=-2}^{2} (-1)^m Q_m(\hat{\boldsymbol{x}}_i) Q_m^*(\hat{\boldsymbol{x}}_i) = \frac{2}{3}\left[1 - \frac{1}{2}\sin^2 2\theta_i\right] \ ,$$
$$\sum_{m=-2}^{2} (-1)^m U_m(\hat{\boldsymbol{x}}_i) U_m^*(\hat{\boldsymbol{x}}_i) = \frac{2}{3}\cos 2\theta_i \ ,$$
$$\sum_{m=-2}^{2} (-1)^m Q_m(\hat{\boldsymbol{x}}_i) U_m^*(\hat{\boldsymbol{x}}_i) = 0 \ . \tag{5.63}$$

5.4 Statistics of the cluster polarization signal

Now we consider the information obtainable from measurements of the CMB polarization signal (due to scattering of the CMB quadrupole) from galaxy clusters at various redshifts and lines of sight. Assuming that the redshifts of each cluster can be obtained, this allows mapping, in principle, of a particular linear combination of a_{lm} over a significant portion of our past light cone. Galaxy clusters at similar redshifts and on lines of sight separated by small angles will produce polarization signals with Stokes parameters which are strongly correlated. Widely separated clusters produce uncorrelated signals — and it is the combination of these uncorrelated signals that beats cosmic variance. In similar work, Cooray & Baumann (2003) constructed estimators of $C_2(\tau)$ (taking into account the kinematic SZ contamination of the polarization signal also), but did not consider the effect of statistical variation in the polarization signal. This variation was considered by Seto & Sasaki (2000), but they did not include the ISW effect.

Using Eqn. (5.55), the two-point correlation function $\langle Q(\boldsymbol{x}) Q(\boldsymbol{x}')\rangle$, of the Stokes parameters, as defined in the basis Eqn. (5.52), due to two clusters at general comoving

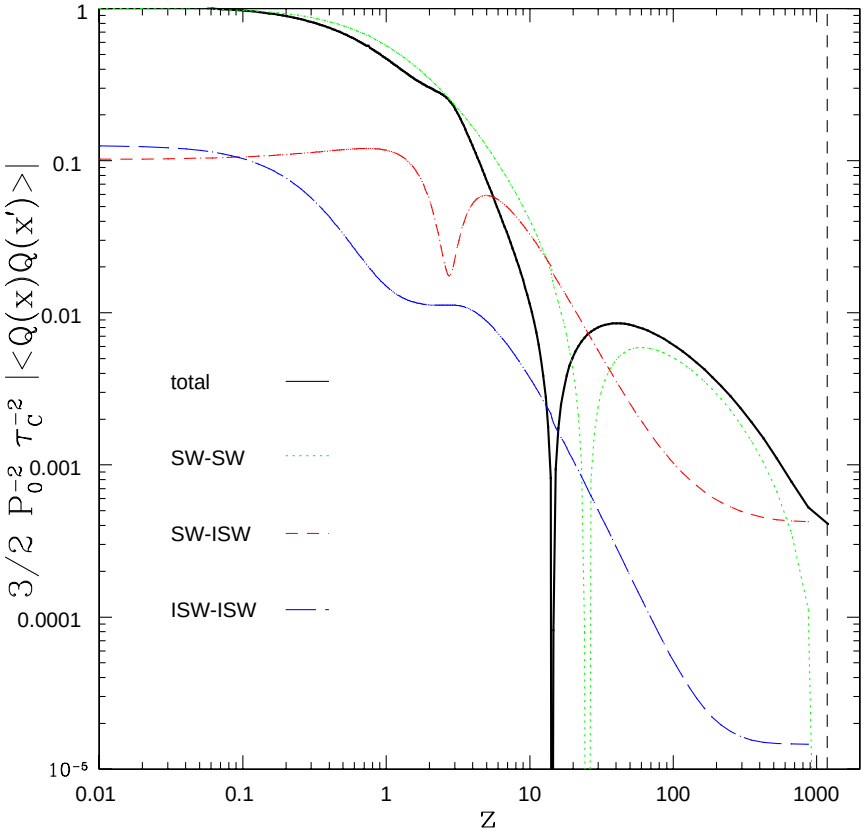


Figure 5-4: Normalized magnitude of the two-point correlation of the Stokes Q, $\frac{3}{2}P_0^{-2}\tau_C^{-2}|\langle Q(\boldsymbol{x})Q(\boldsymbol{x}')\rangle|$, where $\boldsymbol{x}$ is taken to be at redshift $z = 0$ and $\boldsymbol{x}'$ is a point at redshift z in the plane orthogonal to $\hat{\boldsymbol{z}}$. The growth function was computed with cosmological parameters $\Omega_m = 0.35$, $\Omega_\Lambda = 0.65$, $n = 1$. (Note that the dip in the interference term, which occurs at $z \approx 3$, is not a zero crossing).

positions $\boldsymbol{x}, \boldsymbol{x}'$ with Thomson optical depths τ_C, τ_C' is given by [7]

$$\langle Q(\boldsymbol{x})Q(\boldsymbol{x}')\rangle = \tau_C\tau_C' P_0^2 \sum_{m,m'=-2}^{2} Q_m(\hat{\boldsymbol{x}})Q^*_{m'}(\hat{\boldsymbol{x}}')\, C_{2m2m'}(\boldsymbol{x},\boldsymbol{x}') \ , \tag{5.64}$$

and similarly for $\langle U(\boldsymbol{x})U(\boldsymbol{x}')\rangle$ and $\langle Q(\boldsymbol{x})U(\boldsymbol{x}')\rangle$.

The two-point correlation function $\langle Q(\boldsymbol{x})Q(\boldsymbol{x}')\rangle$, for points lying on the same line of sight, is shown in Fig. 5-4, which was computed with cosmological parameters $\Omega_m = 0.35$, $\Omega_\Lambda = 0.65$, $n = 1$. The solid black curve is the total correlation, and the other curves show the contributions from the SW and ISW terms, and their interference term. Note that the interference term is negative - its magnitude is shown here. Only the surface Sachs-Wolfe contribution has a genuine zero crossing. The vertical dotted line indicates the time of (instantaneous) recombination. Note that the SW part of the correlation passes through zero at redshift ≈ 10, since at redshifts higher than this $\boldsymbol{x}'$ is in a region of the universe separated from the origin by a comoving distance greater than $c/(2H_0)$, and thus the $l = 2$ correlations die off rapidly.

As $\boldsymbol{x} \to \boldsymbol{x}'$, $\langle a_{2m}(\boldsymbol{x})a^*_{2m'}(\boldsymbol{x}')\rangle \to C_2(\tau)\delta_{mm'}$, therefore

$$\begin{aligned} \langle Q(\boldsymbol{x})^2\rangle &= \langle U(\boldsymbol{x})^2\rangle = \frac{2}{3}\tau_C^2 P_0^2\ C_2(\tau) \ , \\ \langle Q(\boldsymbol{x})U(\boldsymbol{x})\rangle &= 0 \ . \end{aligned} \tag{5.65}$$

Thus the ensemble average polarization magnitude due to the scattering of the CMB quadrupole is:

$$\langle \Pi(\boldsymbol{x})^2\rangle \equiv \frac{\langle Q(\boldsymbol{x})^2\rangle + \langle U(\boldsymbol{x})^2\rangle}{T_{\rm CMB}(\tau)^2} = \frac{4}{3}\tau_C^2 \left(\frac{P_0}{T_{\rm CMB}(\tau)}\right)^2\ C_2(\tau) \ . \tag{5.66}$$

The quadrupole Q_2 is conventionally defined by $C_2 = (4\pi/5)Q_2^2$. Thus the root mean square polarization magnitude is given by (recall $P_0 \equiv \frac{3}{4\sqrt{10\,\pi}} T_{\rm CMB}(\tau)$)

$$\langle P\rangle \equiv \sqrt{\langle \Pi(\boldsymbol{x})^2\rangle} = \frac{\sqrt{6}}{10}\tau_C Q_2 \ , \tag{5.67}$$

as obtained by Cooray & Baumann (2003). In a ΛCDM model, $\langle P\rangle \approx 5\tau_C\,\mu$K (at zero redshift), so the magnitude of this signal is comparable to that of the other SZ

[7] These correlation functions of the Stokes parameters of the polarization signals from each cluster are obviously not rotationally invariant quantities — they depend on the particular choice of basis used. We note in passing that a correlation function of the polarization signals which does not depend on the choice of basis may be obtained using polarization matrices. Given the normalized polarization matrix of the signal from each cluster, $\boldsymbol{\phi}(\boldsymbol{x})$, the polarization part may be constructed by defining $\widetilde{\boldsymbol{\phi}}(\boldsymbol{x}) \equiv \boldsymbol{\phi}(\boldsymbol{x}) - \boldsymbol{P}(\boldsymbol{x})/2$. Then the correlation function $\langle {\rm Tr}[\widetilde{\boldsymbol{\phi}}(\boldsymbol{x})\widetilde{\boldsymbol{\phi}}(\boldsymbol{x}')]\rangle$ is independent of basis and contains all the relevant information. We do not develop this further here, but note that in future work it would be preferable to use this matrix approach rather than the Stokes parameter based approach.

polarization effects.

Now the relations Eqn. (5.65) suggest unbiased estimators $\widehat{C}_2^Q(\tau)$ and $\widehat{C}_2^U(\tau)$ of $C_2(\tau)$, given the measured Stokes parameters of N clusters at the same redshift $z(\tau)$ on lines of sight $\hat{\boldsymbol{x}}_i$ with optical depths $\tau_{C,i}$ $(i = 1, \cdots, N)$, which beat the cosmic variance limit:

$$\begin{aligned}
\widehat{C}_2^Q(\tau) &\equiv \frac{3}{2} P_0^{-2} \sum_i W_i \left[(Q(\boldsymbol{x}_i) + N_i)^2 - \sigma_P^2 \right] , \\
\widehat{C}_2^U(\tau) &\equiv \frac{3}{2} P_0^{-2} \sum_i W_i \left[(U(\boldsymbol{x}_i) + N_i)^2 - \sigma_P^2 \right] .
\end{aligned} \tag{5.68}$$

We have added to the polarization signal from each cluster a random instrument noise component N_i, assumed to be Gaussian with brightness temperature variance σ_P for both Q and U, in which case $\langle N_i \rangle = 0$, $\langle N_i^2 \rangle = \sigma_P^2$, $\langle N_i^2 \, N_j^2 \rangle = \sigma_P^4 (1 + 2\delta_{ij})$. For the mean of these estimators to equal $C_2(\tau)$, the weights W_i must be chosen to satisfy

$$\sum_i W_i \, \tau_{C,i}^2 = 1 . \tag{5.69}$$

We will consider the simple choice

$$W_i = \frac{\tau_{C,i}^{n-2}}{\sum_j \tau_{C,j}^n} . \tag{5.70}$$

Note that $n = 0$ would be a bad choice since it gives more weight to clusters producing weaker signals, reducing the signal to noise. A better choice is the uniform weighting $n = 2$.

The cosmic variance limit on these estimators is determined by the variances (with X, X' indicating either Q or U)

$$\begin{aligned}
\mathrm{Var}\, \widehat{C}_2^X(\tau) &= \frac{9}{4} P_0^{-4} \sum_{ij} W_i \, W_j \left[\langle (X(\boldsymbol{x}_i) + N_i)^2 (X(\boldsymbol{x}_j) + N_j)^2 \rangle \right. \\
&\qquad \left. - \sigma_P^2 \langle X(\boldsymbol{x}_i) + N_i)^2 \rangle - \sigma_P^2 \langle X(\boldsymbol{x}_j) + N_j)^2 \rangle + \sigma_P^4 \right] .
\end{aligned} \tag{5.71}$$

The sum over i, j may be broken into a contribution from clusters at the same location, $\mathrm{Var}_1(X)$, a contribution from clusters at seperate locations, $\mathrm{Var}_2(X, X)$, and the noise variance, $N(W_i)$:

$$\mathrm{Var}\, \widehat{C}_2^X(\tau) = \mathrm{Var}_1(X) + \mathrm{Var}_2(X, X) + N(W_i) , \tag{5.72}$$

where

$$\begin{aligned}
\mathrm{Var}_1(Q) &= P_0^{-4} C_2(\tau)^2 \sum_i W_i^2 \tau_{C,i}^4 \left[1 + (1 - \frac{1}{2} \sin^2 2\theta_i)^2 \right] , \\
\mathrm{Var}_1(U) &= P_0^{-4} C_2(\tau)^2 \sum_i W_i^2 \tau_{C,i}^4 \left[1 + \cos^2 2\theta_i \right] ,
\end{aligned} \tag{5.73}$$

and

$$\mathrm{Var}_2(X,X') = \frac{9}{2}\sum_{i,\,j>i} W_i\,W_j\,\tau_{C,i}^2\tau_{C,j}^2\,\left|\sum_{m_1m_2} X_{m_1}(\hat{\boldsymbol{x}}_i)X'^{*}_{m_2}(\hat{\boldsymbol{x}}_j)\,C_{2m_12m_2}(\boldsymbol{x}_i,\boldsymbol{x}_j)\right|^2$$
$$+\,\frac{9}{2}\sum_{i,\,j>i} W_i\,W_j\;\;\tau_{C,i}^2\tau_{C,j}^2\left|\sum_{m_1m_2}(-1)^{m_2}X_{m_1}(\hat{\boldsymbol{x}}_i)X'^{*}_{m_2}(\hat{\boldsymbol{x}}_j)\,C_{2m_12,m_2}(\boldsymbol{x}_i,\boldsymbol{x}_j)\right|^2\;, \qquad (5.74)$$

and

$$N(W_i) = \frac{9}{2}P_0^{-4}\sum_i W_i^2\left[\frac{4}{3}P_0^2\tau_{C,i}^2C_2(\tau)\sigma_P^2+\sigma_P^4\right]\;. \qquad (5.75)$$

Here we used the 4-point correlation function from Eqn. (5.31), the relations in equations (5.62), and the fact that $C_{2m2m'}(\boldsymbol{x}_i,\boldsymbol{x}_i)=\delta_{mm'}C_2(\tau)$. Note that the variances are functions of the angular positions of the clusters on the sky, which is due to the specific choice of polarization basis for the Stokes parameters.

In the simplest case all clusters have the same optical depth τ_C, and the weights given in equation (5.70) are all equal, $W_i=\tau_C^{-2}/N$. This gives

$$N(W_i) = \frac{9}{2N}P_0^{-2}C_2(\tau)^2\left[\frac{4}{3}\left(\frac{\sigma_P^2}{\tau_C^2C_2(\tau)}\right)+\left(\frac{\sigma_P^2}{\tau_C^2C_2(\tau)}\right)^2\right] \qquad (5.76)$$

If all of the correlations $C_{2m2,m'}(\boldsymbol{x}_i,\boldsymbol{x}_j)$ with $i\neq j$ vanish, then only the first terms on the right hand sides of equations (5.73) remain. If however the cluster positions $\boldsymbol{x}_i$ are close enough that the these correlations approach $C_2(\tau)$, then the $O(N^2)$ terms in the $\mathrm{Var}_2(X,X')$ terms combine to swamp the first terms. Thus in order to beat cosmic variance by a factor of $O(N^{-1/2})$ we need N sets of clusters which are mutually uncorrelated (as pointed out in a qualitative discussion by Kamionkowski & Loeb (1997)).

The number of uncorrelated regions available increases as the redshift increases, since the comoving region surrounding each cluster outside of which the polarization is approximately uncorrelated with that produced by the cluster is smaller at higher redshift. This is because smaller comoving scales contribute to the $l=2$ harmonic of the CMB on the sky at higher redshift, i.e. $C_2(\tau<\tau_0)$ depends on fluctuations of smaller scale than $C_2(\tau_0)$.

At very low redshift, any cluster will be correlated with any other, and we get back the usual cosmic variance constraint. In other words, we can beat the cosmic variance on $C_2(\tau<\tau_0)$, but not on $C_2(\tau_0)$, today's quadrupole.

To demonstrate the reduction in cosmic variance, we first combine the Q and U measurements to obtain an improved estimator of $C_2(\tau)$. Taking a linear combination of $\widehat{C}_2^Q,\widehat{C}_2^U$ yields an improved estimator,

$$\widehat{C}_2^P(\tau) \;\equiv\; \alpha\widehat{C}_2^Q(\tau)+(1-\alpha)\widehat{C}_2^U(\tau)\;, \qquad (5.77)$$

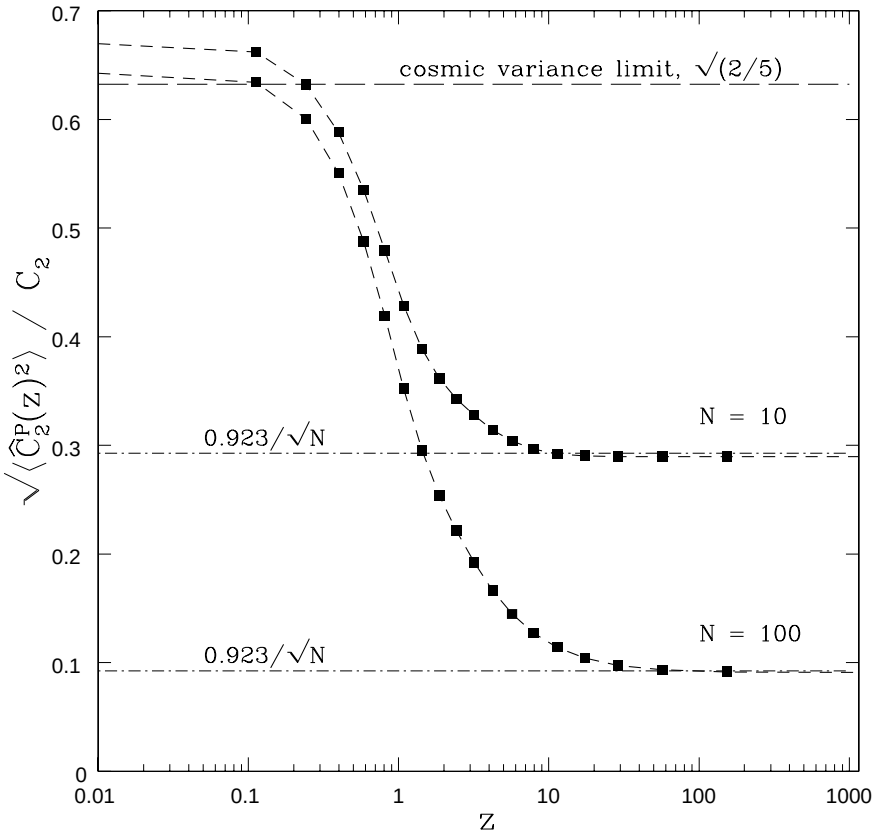


Figure 5-5: Reduction in cosmic variance with clusters at redshift z.

with $0 \leq \alpha \leq 1$. This has variance

$$\begin{aligned} \mathrm{Var}\, \widehat{C}_2^P(\tau) &= \alpha^2 \mathrm{Var}\, \widehat{C}_2^Q(\tau) + (1-\alpha)^2 \mathrm{Var}\, \widehat{C}_2^U(\tau) \\ &\quad + 2\alpha(1-\alpha)\, \mathrm{Cov}(\widehat{C}_2^Q, \widehat{C}_2^U)\ . \end{aligned} \tag{5.78}$$

It is easy to see that the covariance $\mathrm{Cov}(\widehat{C}_2^Q, \widehat{C}_2^U)$ is zero because of the relations in equations (5.62), which imply $\mathrm{Var}_2(Q,U) = 0$. Thus the optimum value of α is trivially $\alpha = 1/2$.

In Fig. 5-5 we show the limits of the accuracy to which C_2 may be measured via polarization measurements from a set of clusters all at the same redshift, but sprinkled in random directions on the sky (in cosmology $\Omega_m = 0.35$, $\Omega_\Lambda = 0.65$, power spectral index $n = 1$). The filled squares indicate $\widehat{C}_2^P$, the variance of the average estimator for C_2, $(\widehat{C}_2^Q + \widehat{C}_2^U)/2$, for 10 and 100 clusters as indicated. This is expressed as a fraction of the Sachs-Wolfe contribution to C_2, which is independent of redshift for $n = 1$. The noise term is not included here.

Only the Sachs-Wolfe contribution to the estimator is shown here. [8] The horizontal dashed line at $\sqrt{2/5} \approx 0.63$ indicates the cosmic variance limit on a single CMB sky given in Eqn. (5.32).

As $z \to 0$, the estimator variance slightly exceeds the cosmic variance limit given a single CMB sky — clusters with overlapping correlation spheres as $z \to 0$ are no more useful than a direct measurement of the quadrupole on our sky by e.g. WMAP. In fact our estimator of C_2 is worse than a direct measure — but an optimal estimator could be constructed which would yield all of the a_{2m} at $z \approx 0$ from cluster measurements at various points on the sky.

As $z \to \infty$, the estimator variance approaches $0.923/\sqrt{N}$. This may be derived by averaging the estimator variances in Eqn. (5.73) over angles in the $N \to \infty$ limit, assuming the signals from each cluster are uncorrelated in this high redshift limit.

At $z = 2$, with 100 clusters, the cosmic variance limit on determination of C_2 has been reduced from about 60% of the ensemble average to 30%. Thus with a reasonably large number of clusters this estimator allows much more precise constraint of C_2 than is possible with one CMB sky.

[8] The general computation with the ISW effect included was considerably more difficult to achieve numerically, so we computed the simplified case with only the SW contribution. But note that the ISW constribution will become negligible at sufficiently high z, before Ω_Λ begins to dominate.

5.5 Discussion

We have developed a statistical theory of the part of the polarization signal in the CMB in the direction of galaxy clusters produced by scattering of the CMB temperature quadrupole. We have shown explicitly that it is possible to beat the cosmic variance limits using indirect information about the last scattering surfaces of distant observers obtained via these polarization measurements, and thus constrain the C_2 harmonic of the CMB as a function of redshift with greater statistical accuracy. A more accurate measurement of $C_2(\tau)$ would provide a test of the low quadrupole amplitude observed by WMAP, thereby making a key test of standard cosmology.

In the standard ΛCDM model the ISW effect produces a small $(1-2\%)$ bump in CMB harmonic $C_2(\tau)$, which is swamped by the high cosmic variance at low redshift. However the ISW effect produces a significant feature in the two-point correlation function of the Stokes parameters which might be detectable. Detection of the ISW effect would provide additional information about the acceleration of the universe and the dark energy. We also note that this method is a rather sensitive probe of deviations from the scale-invariant power spectrum, because if $n \neq 1$ then the Sachs-Wolfe contribution to $C_2(\tau)$ either grows or decays rapidly with conformal time.

The procedure for getting around the cosmic variance limit that we have outlined is something of an idealization. We assumed that we have the polarization signals from many clusters at the same redshift, and we also ignored noise. In practice, separating the quadrupole signal from the other SZ signals will be a major experimental challenge. However the signal to noise may be increased to some extent by combining signals from clusters nearby in direction and redshift, since the signal from sufficiently nearby clusters is strongly correlated.

Given actual data of cluster polarizations, optical depths, and redshifts, the optimal procedure for constraining the cosmological parameters $\{\Omega_m,\ \Omega_\Lambda, n, w\}$ (w parametrizes the equation of state of the dark energy) involves maximizing the likelihood function for the parameters. This may be constructed from the joint p.d.f of the quantities $Q(\boldsymbol{x}_i)$. By the same argument used for the a_{lm}, this p.d.f has Gaussian form:

$$P\left[Q(z_1,\hat{\boldsymbol{x}}_1),\cdots,Q(z_N,\hat{\boldsymbol{x}}_N)|\Omega_m,\Omega_\Lambda,n,w\right] = \frac{1}{\sqrt{(2\pi)^N|\det[\boldsymbol{R}]|}}\exp\left[-\frac{1}{2}\sum_{ij}^{N}Q(\boldsymbol{x}_i)R_{ij}^{-1}Q(\boldsymbol{x}_j)\right] \quad , \quad (5.79)$$

with covariance $R_{ij} = \langle Q(\boldsymbol{x}_i)Q(\boldsymbol{x}_j)\rangle$ (and similarly for the general case with Q and U). The comoving vectors are known functions of the parameters given the redshifts and direction vectors, $\boldsymbol{x}_i = X(z_i,\Omega_m,\Omega_\Lambda,w)\hat{\boldsymbol{x}}_i$.

Now we consider to what extent the this method can put better constraints on the primordial power spectrum. In principle the error bar at wavenumber corresponding to $l = 2$ in the primordial power spectrum can be reduced by roughly the same factor as the measurement of the harmonic $C_2(\tau)$ (see Eqn. (5.81)). The $l = 2$ harmonic of the CMB anisotropy at redshift $z = 0$ contains contributions from a broad

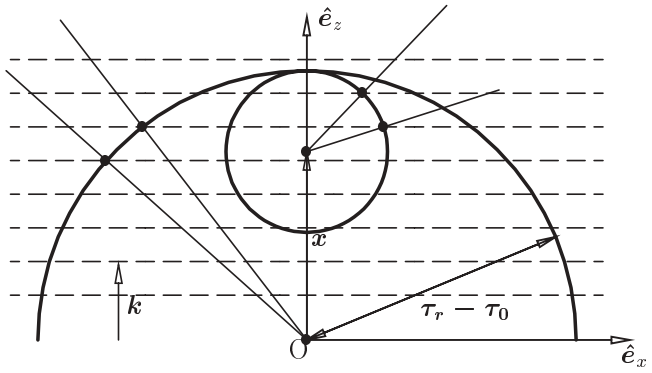


Figure 5-6: Illustrates that the angular scale of a given comoving k-mode subtended on the CMB sky of an observer at high redshift is greater than the angular scale of the same k-mode on the CMB sky of an observer at low redshift.

range of wavenumbers of order c/H_0. But at higher redshifts the CMB quadrupole probes the potential on a smaller last scattering surface than at redshift $z = 0$. Thus measurements at higher redshift constrain the primordial power spectrum at higher wavenumber (see Fig. 5-6). The comoving scale on which power is probed by our method at a conformal time τ is given roughly by comoving wavenumber $k_l = l/[c(\tau_r - \tau_0)]$ (with $l = 2$).

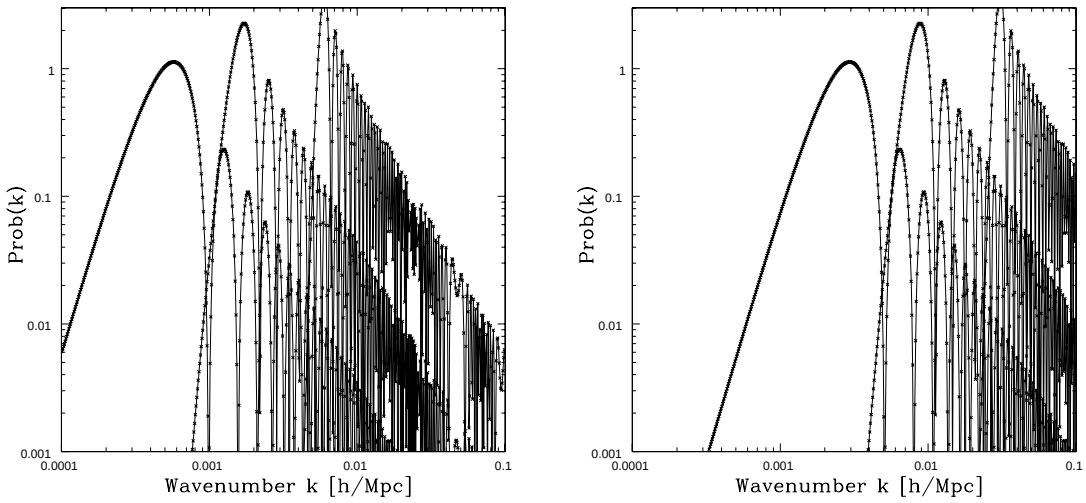
To obtain a measure of the wavenumber being probed, there is useful method (Tegmark & Zaldarriaga, 2002) for relating the CMB power spectrum at a given epoch directly to the primordial power spectrum, given the CMB transfer functions. This method yields an estimate of the range of wavenumbers k which contribute to each l value. The power in the $l = 2$ mode at time τ yields the *window function*:

$$\mathcal{P}(k|\tau, l) = \frac{(4\pi)^2 \Delta_l^2(k,\tau) P_\phi(k)\, k^3}{C_l(\tau)}, \qquad \int_{-\infty}^{\infty} \mathcal{P}(k|\tau, l)\, d\ln k = 1\ , \tag{5.80}$$

where the transfer function is computed with some assumed model. For estimation of the power at k modes corresponding to $l < 20$, the Sachs-Wolfe transfer function is sufficient. The window function obtained using the Sachs-Wolfe transfer function is shown for the model $h^2\Omega_m = 0.141, \Omega_\Lambda = 0.71, h = 0.7, n = 1$ at two redshifts in Fig. 5-7. The k value is then plotted at the median k_i of the window function with an error bar extending from the 20th to the 80th percentile (which corresponds to full-width-half- maximum in the Gaussian case). The range of wavenumbers contributing to the $l = 2$ harmonic at a given redshift is shown in Fig. 5-8. An estimate of the primordial power at this wavenumber is then

$$\hat{P}_\phi(k_i|\tau, l) \;=\; \frac{k_i \widehat{C}_l(\tau)}{\int_{-\infty}^{\infty} k \Delta_l^2(k,\tau)\, k^3\, d\ln k}\ . \tag{5.81}$$

This gives a reasonable estimate because the transfer function Δ_l is quite stronly peaked at $k_l = l/[c(\tau_r - \tau_0)]$. Crudely, $\Delta_l^2(k,\tau) \approx A\delta(k - l(\tau - \tau_r)^{-1})$, yielding the



(a) $z = 0$ (b) $z = 20,\ |\boldsymbol{x}| = 0.805$

Figure 5-7: The window function defined in equation (5.80), at two different redshifts. From left to right, the curves are for multipoles $l = 2, 8, 32$.

ensemble average $\langle \hat{P}_\phi(k) \rangle \approx (k_i/k_l) P_\phi(k) \approx P_\phi(k)$. The fractional variance in the estimator $\hat{P}_\phi$ is the same as the fractional variance in the estimator $\widehat{C}_l(\tau)$ obtained from polarization measurements.

Thus at each redshift z we can in principle reduce the cosmic variance error bars in the inferred primordial spectrum by the factor shown in Fig. 5-5 at the range of wavenumber shown in Fig. 5-8. The gain will be greater at smaller scales which correspond to contributions from higher redshift.

It remains to be seen if this technique will be a useful cosmological probe in practice, but we are optimistic in light of the recent progress in the detection of the primary CMB polarization signal (Bennett et al., 2003; Kogut et al., 2003).

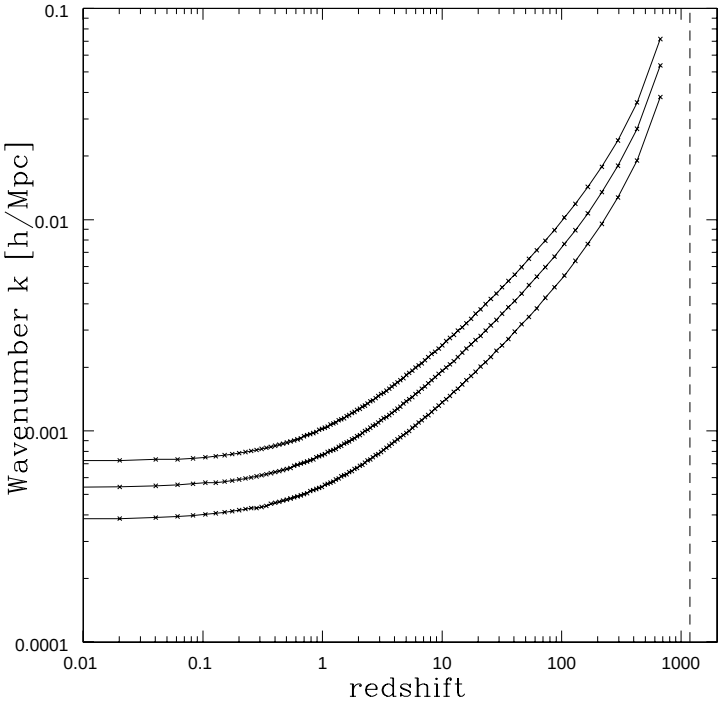


Figure 5-8: The wavenumbers at the 20th, 50th and 80th percentile of the window function for the $l = 2$ harmonic, as a function of redshift. This gives the range of wavenumbers contributing to the $l = 2$ harmonic at a given redshift.

Appendix A

Symmetry of the Klein-Nishina matrix element

It is apparent that the square of the invariant amplitude for the Klein-Nishina formula should be symmetric under the interchange of the initial and final states, but it is written in a way that is very asymmetric. Here we show that it is possible to write the K-N invariant matrix element in a way that is manifestly symmetric between the initial and final states. [1] One might think that the non-invariance of the $\epsilon_i \cdot \epsilon_f$ term is perhaps canceled by the noninvariance of the rest, but it turns out that it is much simpler than that. Each part is separately symmetric.

The matrix element for Compton scattering is usually written in an asymmetric way. One can call the initial and final electron 4-momenta $\vec{q}_i$ and $\vec{p}_f$, and the initial and final photon 4-momenta $\vec{q}_i$ and $\vec{p}_f$. Using conventions for which the Lorentz dot-product is defined as time minus space, one defines $p_1 = \vec{q}_i \cdot \vec{q}_i$ and $p_2 = \vec{p}_f \cdot \vec{q}_i$. That is, in the initial rest frame of the electron, p_1 and p_2 are the initial and final photon energies, multiplied by m_e, the mass of an electron. One also defines polarization vectors for the photons to have zero time-components in this frame, so $\epsilon_i \cdot \vec{q}_i = 0$, $\epsilon_f \cdot \vec{q}_i = 0$. The differential cross section is then

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4m_e^2} \left(\frac{p_2}{p_1}\right)^2 M^2 \ , \tag{1}$$

where

$$M^2 = \frac{p_2}{p_1} + \frac{p_1}{p_2} - 2 + 4(\epsilon_2 \cdot \epsilon_1)^2 \ . \tag{2}$$

The invariant matrix element M^2 should be symmetric under the interchange of initial and final states, $i \leftrightarrow f$, but it does not look that way. However, it really is. To see this, note that conservation of 4-momentum implies that

$$\vec{q}_i + \vec{q}_i = \vec{p}_f + \vec{p}_f \ . \tag{3}$$

Squaring both sides, and using the fact that $\vec{q}_i^{\,2} = \vec{p}_f^{\,2}$ and $\vec{q}_i^{\,2} = \vec{p}_f^{\,2}$, one has immediately

[1] This appendix is based on a private communication from A. Guth

that

$$\vec{\vec{q}}_i \cdot \vec{\vec{q}}_i = \vec{p}_f \cdot \vec{p}_f \ , \tag{4}$$

so in fact p_1 is invariant under $i \leftrightarrow f$. Similarly conservation of 4-momentum implies that

$$\vec{\vec{q}}_i - \vec{p}_f = \vec{p}_f - \vec{\vec{q}}_i \ , \tag{5}$$

and squaring implies that

$$\vec{\vec{q}}_i \cdot \vec{p}_f = \vec{p}_f \cdot \vec{\vec{q}}_i \ , \tag{6}$$

so p_2 is invariant under $i \leftrightarrow f$.

The only remaining problem is the $(\epsilon_2 \cdot \epsilon_1)^2$ term, which is not manifestly invariant, since the ϵ's were both defined to have vanishing 4th components in the initial rest frame of the electron, so $\epsilon_i \cdot \vec{\vec{q}}_i = 0$, $\epsilon_f \cdot \vec{\vec{q}}_i = 0$. One can use an arbitrary gauge for the polarization vectors, however, if one explicitly constructs the gauge transformation satisfying $\epsilon \cdot \vec{\vec{q}}_i = 0$ before calculating the dot product. That is, if ϵ_i does not satisfy $\epsilon_i \cdot \vec{\vec{q}}_i = 0$, then one constructs

$$\epsilon_i' = \epsilon_i - \frac{\epsilon_i \cdot \vec{\vec{q}}_i}{\vec{\vec{q}}_i \cdot \vec{\vec{q}}_i} \vec{\vec{q}}_i \tag{7}$$

and

$$\epsilon_f' = \epsilon_f - \frac{\epsilon_f \cdot \vec{\vec{q}}_i}{\vec{p}_f \cdot \vec{\vec{q}}_i} \vec{p}_f \ , \tag{8}$$

so $\epsilon_i' \cdot \vec{\vec{q}}_i = \epsilon_f' \cdot \vec{\vec{q}}_i = 0$. To continue, it is useful to define a more compact notation. Let

$$A_{\alpha\beta} \equiv \epsilon_\alpha \cdot p_\beta \tag{9}$$

$$B_{\alpha\beta} \equiv \epsilon_\alpha \cdot k_\beta \ . \tag{10}$$

where α and β can be either i or f. Each polarization vector is orthogonal to its corresponding momentum, so $B_{ii} = B_{ff} = 0$. In this notation Eqs. (7) and (8) become

$$\epsilon_i' = \epsilon_i - \frac{A_{ii}}{p_1} \vec{\vec{q}}_i \tag{11}$$

and

$$\epsilon_f' = \epsilon_f - \frac{A_{fi}}{p_2} \vec{p}_f \ . \tag{12}$$

So, for polarization vectors ϵ_i and ϵ_f written in an arbitrary gauge, the equation for M^2 must be written by replacing $\epsilon_i \cdot \epsilon_f$ with

$$\epsilon_i' \cdot \epsilon_f' = \epsilon_i \cdot \epsilon_f - \frac{A_{fi} B_{if}}{p_2} - \frac{A_{ii} B_{fi}}{p_1} + \frac{A_{ii} A_{fi}}{p_1 p_2} \vec{\vec{q}}_i \cdot \vec{p}_f \ . \tag{13}$$

To proceed, we want to use some identities that follow from energy-momentum con-

servation. Dotting both sides of Eq. (3) with ϵ_i, one finds

$$A_{ii} = B_{if} + A_{if} \ , \tag{14}$$

and dotting both sides with ϵ_f (and reversing the sides of the equation) gives

$$A_{ff} = B_{fi} + A_{fi} \ . \tag{15}$$

Since one has 6 dot products — A_{ii}, A_{if}, A_{fi}, A_{ff}, B_{if}, and B_{fi} — and two constraints (Eqs. (14) and (15)), one can eliminate two of the dot products from all expressions. The simplest result seems to arise from eliminating the B's. One also needs to simplify $\vec{\vec{q}}_i \cdot \vec{p}_f$, which can be done by dotting Eq. (3) with $\vec{p}_f$:

$$p_2 + \vec{\vec{q}}_i \cdot \vec{p}_f = p_1 \ ,$$

so

$$\vec{\vec{q}}_i \cdot \vec{p}_f = p_1 - p_2 \ . \tag{16}$$

Finally, substituting into Eq. (13),

$$\begin{aligned} \epsilon_i' \cdot \epsilon_f' &= \epsilon_i \cdot \epsilon_f - \frac{A_{fi}(A_{ii} - A_{if})}{p_2} - \frac{A_{ii}(A_{ff} - A_{fi})}{p_1} + \frac{A_{ii}A_{fi}}{p_1 p_2}(p_1 - p_2) \\ &= \epsilon_i \cdot \epsilon_f + \frac{A_{fi}A_{if}}{p_2} - \frac{A_{ii}A_{ff}}{p_1} \ . \end{aligned} \tag{17}$$

Written in this form, the result is manifestly symmetric under the $i \leftrightarrow f$ interchange, and it is valid for polarization vectors ϵ_i and ϵ_f written in an arbitrary gauge. One can check that the expression vanishes if ϵ_i is replaced by $\vec{\vec{q}}_i$, or if ϵ_f is replaced by $\vec{p}_f$.

Appendix B

Monte Carlo simulation of polarized radiative transfer in an inhomogeneous cloud

Here we consider how to set up a simulation of the Thomson scattering of radiation emergent from a point source through an inhomogenous cloud of electrons within which the source is embedded. This parallels the analytic calculations of Murphy & Chernoff (1993), except with the Monte Carlo method we can easily study the effect of more than one scattering and multiple point sources. A Monte Carlo calculation of Thomson scattering of polarized light in a clumpy cloud using a Stokes parameter approach is described in Kishimoto (1996), and it is instructive to compare the complexity of that scheme to the matrix approach described here.

In the most physical picture of Monte Carlo radiative transfer, simulated photons are generated at some source within the cloud (or injected at the boundary), and allowed to scatter through the cloud until they are absorbed or strike the boundary (Witt, 1977; Code & Whitney, 1995; Gordon et al., 2001). The photons which strike the boundary emerge from the cloud in any direction. But we wish to make a projected image of the scattered liga ht as seen by an observer in the far field in a given direction.

To do this, the flux of photons emerging in a fixed direction can be computed by considering each scattering point as a secondary source, of intensity weighted according to the probability of a photon reaching the observer (Yusef-Zadeh et al., 1984). An image of the intensity and polarization of the radiation escaping along the observer's line of sight can then be formed by accumulating the normalized polarization matrices of escaping photons in a grid of bins lying in the projected plane normal to the line of sight.

To generate an image in the far field, appropriate for example in the case of observations of a galaxy cluster, the correct procedure is to propagate parallel rays from each secondary source. The resulting image will be equivalent to that seen by an observer with a lens in the far field. If the cloud is close enough to the observer, then there is a perspective effect which can be accounted for by projecting rays directly to the observation point rather than along a common viewing direction, but in this case the emergent photons must be weighted by the solid angle subtended by the bin on

the projected face through which the outgoing ray passes.

We now describe the procedure used to generate the images in Fig. 1-1. We assume that Thomson opacity values are provided on a cubic grid, and the line of sight to the observer is parallel to $\boldsymbol{z}$. A photon is generated at the position of the point source (if there are multiple sources, photons are created at each source in numbers proportional to the desired intensity of each source). We choose a random direction for the photon, $\boldsymbol{n}_i$ and assign it the normalized polarization matrix (assuming that the source is emitting unpolarized light) $\boldsymbol{\phi}_i(\boldsymbol{n}_i) = \boldsymbol{P}(\boldsymbol{n}_i)/2$. We sample the optical depth τ to scattering by choosing a random deviate ξ and setting $\tau = -\ln(1-\xi)$ (this correctly models the probability of scattering at a given point along the ray; see e.g. Molnar & Birkinshaw (1999)). The opacity is then integrated along the photon direction until either this optical depth is reached, or the photon hits the cube boundary. The value of the opacity $\kappa(\boldsymbol{x})$ at each point is computed by interpolation of the opacities at the corners of the grid cube surrounding $\boldsymbol{x}$.

If the photon hits the boundary before reaching the pre-decided optical depth, it is discarded. (Thus we discard photons which arrive at the observer without scattering. But in the far field case, only rays that lie exactly along the line of sight would arrive at the observer, so it is only really of interest to compute the scattered radiation field). If the photon reaches the pre-decided optical depth at some point without leaving the cube, we cause it to scatter at that point. First we generate the normalized polarization matrix $\boldsymbol{\phi}_O$ of the photon scattered towards the observer generated by the secondary source at the scattering position:

$$\boldsymbol{\phi}_O(\boldsymbol{z}) = \frac{\boldsymbol{P}(\boldsymbol{z})\boldsymbol{\phi}_i(\boldsymbol{n}_i)\boldsymbol{P}(\boldsymbol{z})}{\mathrm{Tr}\left[\boldsymbol{P}(\boldsymbol{z})\boldsymbol{\phi}_i(\boldsymbol{n}_i)\right]} \ . \tag{B.1}$$

We compute the optical depth τ_O for the photon to be re-emitted and escape from the cube along the z-axis towards the observer, and add the following polarization matrix to the appropriate projected image bin:

$$\boldsymbol{\phi}_B(\boldsymbol{z}) = e^{-\tau_O}\boldsymbol{\phi}_O(\boldsymbol{z}) \ . \tag{B.2}$$

Note that this correctly accounts for the probability of the photon scattering into the line of sight (the phase function). Then a new photon direction $\boldsymbol{n}_s$ is computed, by rejection sampling from the phase function $\Phi(\boldsymbol{n}_s, \boldsymbol{\phi}_i(\boldsymbol{n}_i))$ (given in Eqn. (3.19)). The polarization matrix of the scattered photon is computed:

$$\boldsymbol{\phi}_s(\boldsymbol{n}_s) = \frac{\boldsymbol{P}(\boldsymbol{n}_s)\boldsymbol{\phi}(\boldsymbol{n}_i)\boldsymbol{P}(\boldsymbol{n}_s)}{\mathrm{Tr}\left[\boldsymbol{\phi}(\boldsymbol{n}_i)\boldsymbol{P}(\boldsymbol{n}_s)\right]} \ . \tag{B.3}$$

The photon is allowed to re-scatter multiple times, binning the secondary radiation produced at each scattering, until the photon it leaves the cube. Then a new photon is generated at the source and the whole procedure repeated. The binned values are normalized at the end of the simulation to obtain the observed intensity.

In Fig. B-1 we show the results of applying this procedure to compute the polarized radiation produced by two point sources of equal luminosity, in the same density cube

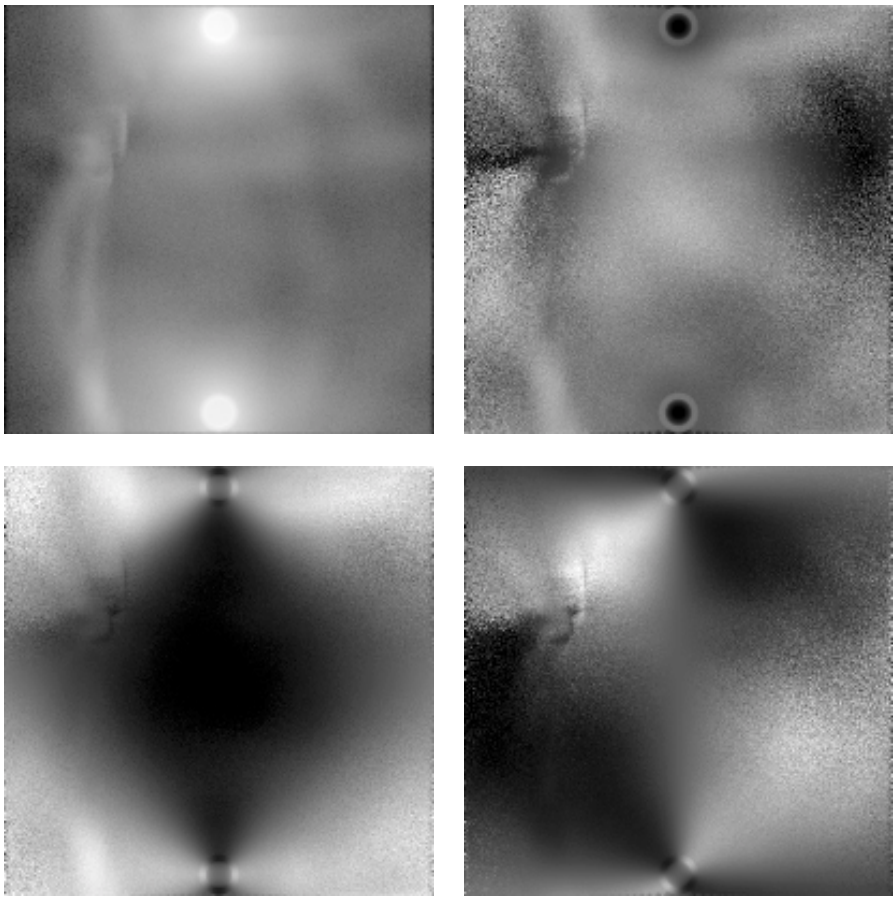


Figure B-1: Radiation Thomson scattered into the line of sight from two point sources embedded in an inhomogeneous cloud of electrons. In the various panels we show: the total intensity I (upper left), the magnitude of the polarization Π (upper right), the Stokes Q parameter (lower left), the Stokes U parameter (lower right). The Stokes parameters are defined with respect to the axes of the projected face. The intensity is plotted on a log scale, the polarization and Stokes parameters on a linear scale.

used to generate Fig. 1-1. The point sources are located at opposite ends of a line passing through the center of the cube perpendicular to the line of sight and parallel to one of the cube sides. 4×10^7 photons were generated, distributed equally between the point sources. The average number of scatterings undergone by each photon in the simulation was 1.6.

Bibliography

Acquista, C., & Anderson, J. L. 1974, ApJ, 191, 567

Audit, E., & Simmons, J. F. L. 1999, MNRAS, 305, L27

Baes, M., & Dejonghe, H. 2002, MNRAS, 335, 441

Barakat, R. 1963, J. Opt. Soc. Am., 53, 317

Bashinsky, S., & Bertschinger, E. 2002, Phys. Rev. D, 65, 123008

Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Spergel, D. N., Tucker, G. S., Wollack, E., Wright, E. L., Barnes, C., Greason, M. R., Hill, R. S., Komatsu, E., Nolta, M. R., Odegard, N., Peirs, H. V., Verde, L., & Weiland, J. L. 2003, astro-ph/0302207

Bennett, C. L., Smoot, G. F., Hinshaw, G., Wright, E. L., Kogut, A., de Amici, G., Meyer, S. S., Weiss, R., Wilkinson, D. T., Gulkis, S., Janssen, M., Boggess, N. W., Cheng, E. S., Hauser, M. G., Kelsall, T., Mather, J. C., Moseley, S. H., Murdock, T. L., & Silverberg, R. F. 1992, ApJ, 396, L7

Bildhauer, S. 1989a, A&A, 219, 25

—. 1989b, Classical and Quantum Gravity, 6, 1171

—. 1990, Classical and Quantum Gravity, 7, 2367

Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton, NJ, Princeton University Press)

Birkinshaw, M. 1999, Phys. Rept., 310, 97

Bond, J. R., & Efstathiou, G. 1987, MNRAS, 226, 655

Born, M., & Wolf, E. 1980, Principles of Optics (Oxford: Pergamon Press)

Breuer, R. A., & Ehlers, J. 1980, Royal Society of London Proceedings Series A, 370, 389

—. 1981, Royal Society of London Proceedings Series A, 374, 65

Britton, M. C. 2000, ApJ, 532, 1240

Byron, F., & Fuller, R. 1992, Mathematics of Classical and Quantum Physics (Dover)

Carozzi, T., Karlsson, R., & Bergman, J. 2000, Phys. Rev. E, 61, 2024

Challinor, A. 2000, Phys. Rev. D, 62, 43004

Challinor, A., & Lasenby, A. 1998, ApJ, 499, 1

Challinor, A., & van Leeuwen, F. 2002, Phys. Rev. D, 65, 103001

Challinor, A. D., Ford, M. T., & Lasenby, A. N. 2000, MNRAS, 312, 159

Chandrasekhar, S. 1960, Radiative Transfer (Dover)

Code, A. D., & Whitney, B. A. 1995, ApJ, 441, 400

Cooray, A., & Baumann, D. 2003, Phys. Rev. D, 67, 63505

Costa, E., Soffitta, P., Bellazzini, R., Brez, A., Lumb, N., & Spandre, G. 2001, Nature, 411, 662

Dautcourt, G., & Rose, K. 1978, Astronomische Nachrichten, 299, 13

Dialetis, D. 1969, J. Opt. Soc. Am., 59, 74

Diego, J. M., Hansen, S. H., & Silk, J. 2003, MNRAS, 338, 796

Dodelson, S., & Jubas, J. M. 1995, ApJ, 439, 503

Gordon, K. D., Misselt, K. A., Witt, A. N., & Clayton, G. C. 2001, ApJ, 551, 269

Gorecki, A., & Wilczewski, W. 1984, Acta Astronomica, 34, 141

Greiner, W., & Reinhardt, J. 1994, Quantum Electrodynamics (Springer)

Groot, S. R. d., van Leeuwen, W. A., & van Weert, C. G. 1980, Relativistic Kinetic Theory: Principles and Applications (North-Holland Publishing Company)

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Hansen, F. K., & Lilje, P. B. 1999, MNRAS, 306, 153

Heath, D. J. 1977, MNRAS, 179, 351

Henney, W. J. 1994, ApJ, 427, 288

Holzapfel, W. L., Arnaud, M., Ade, P. A. R., Church, S. E., Fischer, M. L., Mauskopf, P. D., Rephaeli, Y., Wilbanks, T. M., & Lange, A. E. 1997, ApJ, 480, 449

Hu, W., & White, M. 1997, Phys. Rev. D, 56, 596

Itoh, N., Kohyama, Y., & Nozawa, S. 1998, ApJ, 502, 7

Itoh, N., Nozawa, S., & Kohyama, Y. 2000, ApJ, 533, 588

Jackson, J. D. 1998, Classical Electrodynamics, 3rd Ed.

Kamionkowski, M., & Loeb, A. 1997, Phys. Rev., D56, 4511

Kashlinsky, A., & Atrio-Barandela, F. 2000, ApJ, 536, L67

Kishimoto, M. 1996, ApJ, 468, 606

Kogut, A., Banday, A. J., Bennett, C. L., Gorski, K. M., Hinshaw, G., Smoot, G. F., & Wright, E. I. 1996, ApJ, 464, L5+

Kogut, A., Spergel, D. N., Barnes, C., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Limon, M., Meyer, S. S., Page, L., Tucker, G., Wollack, E., & Wright, E. L. 2003, astro-ph/0302213

Kopeikin, S., & Mashhoon, B. 2002, Phys. Rev. D, 65, 64025

Kosowsky, A. B. 1994, Ph.D. Thesis

Lee, H.-W., Blandford, R. D., & Western, L. 1994, MNRAS, 267, 303

Lightman, A. P., Press, W. H., Price, R. H., & Teukolsky, S. A. 1975, Problem book in relativity and gravitation (Princeton University Press)

Lubin, P. M., Epstein, G. L., & Smoot, G. F. 1983, Physical Review Letters, 50, 616

Ma, C., & Bertschinger, E. 1995, ApJ, 455, 7

Misner, C. W., Thorne, K. S., & Wheeler, J. A. 1973, Gravitation (W.H. Freeman and Co.)

Molnar, S. M., & Birkinshaw, M. 1999, ApJ, 523, 78

—. 2000, ApJ, 537, 542

Murphy, B. W., & Chernoff, D. F. 1993, ApJ, 418, 60

Nagai, D., Kravtsov, A. V., & Kosowsky, A. 2003, ApJ, 587, 524

Nagirner, D. I., & Poutanen, J. 2001, A&A, 379, 664

Nouri-Zonoz, M. 1999, Phys. Rev. D, 60, 24013

Nozawa, S., Itoh, N., & Kohyama, Y. 1998, ApJ, 508, 17

Padmanabhan, T. 2003, Theoretical Astrophysics. Volume III: Galaxies and Cosmology (Cambridge University Press)

Peacock, J. A. 1999, Cosmological physics (Cambridge University Press)

Peterson, J. B., Carlstrom, J. E., Cheng, E. S., Kamionkowski, M., Lange, A. E., Seiffert, M., Spergel, D. N., & Stebbins, A. 1999, astro-ph/9907276

Phillips, P. R. 1995, ApJ, 455, 419

Plebanski, J. 1960, Physical Review, 118, 1396

Pozdnyakov, L. A., Sobel, I. M., & Syunyaev, R. A. 1983, Astrophysics and Space Physics Reviews, 2, 189

Refregier, A., & Teyssier, R. 2002, Phys. Rev. D, 66, 43002

Rephaeli, Y. 1995a, ARA&A, 33, 541

—. 1995b, ApJ, 445, 33

Rose, M. E. 1995, Elementary Theory Of Angular Momentum (Dover)

Rybicki, G. B., & Lightman, A. P. 1979, Radiative Processes in Astrophysics (Wiley)

Sachs, R. K., & Wolfe, A. M. 1967, ApJ, 147, 73

Sazonov, S. Y., & Sunyaev, R. A. 1998a, ApJ, 508, 1

—. 1998b, Astronomy Letters, 24, 553

—. 1999, MNRAS, 310, 765

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Springer-Verlag)

Seljak, U. 1995, MIT Ph.D. Thesis

Seljak, U., Burwell, J., & Pen, U. 2001, Phys. Rev. D, 63, 63001

Seljak, U., & Zaldarriaga, M. 1996, ApJ, 469, 437

Seto, N., & Sasaki, M. 2000, Phys. Rev. D, 62, 123004

Simmons, J., & Guttman, M. 1970, States, Waves and Photons: A Modern Introduction to Light (Addison-Wesley)

Skrotskii, G. 1957, Dokl. Akad. Nauk SSSR, 114, 73

Stedman, G. E., & Pooke, D. M. 1982, Phys. Rev. D, 26, 2172

Stone, J. M. 1963, Radiation and Optics (McGraw-Hill)

Sunyaev, R. A., & Zeldovich, I. B. 1980a, ARA&A, 18, 537

—. 1980b, MNRAS, 190, 413

Swindell, W. E. 1975, Benchmark Papers in Optics: Polarized Light (Dowden, Hutchinson, and Ross)

Synge, J. L. 1957, The Relativistic Gas (North-Holland Publishing Company)

Tegmark, M., & Zaldarriaga, M. 2002, Phys. Rev. D, 66, 103508

Thorne, K. S. 1981, MNRAS, 194, 439

Wiener, N. 1930, Acta Math, 55, 117

Witt, A. N. 1977, ApJS, 35, 1

Wolf, E. 1959, Nuovo Cimento, 13, 1165

Wood, K., et al. 2001, astro-ph/0107060

Yusef-Zadeh, F., Morris, M., & White, R. L. 1984, ApJ, 278, 186

Zaldarriaga, M., & Seljak, U. 1997, Phys. Rev. D, 55, 1830

Zeldovich, Y. B., & Sunyaev, R. A. 1969, Ap&SS, 4, 301