Title

Statistics of Cosmic Microwave Background and Lorentz Boosts

Author

Choi, Chi Po

Publication Date

2024

Peer reviewed|Thesis/dissertation

Statistics of Cosmic Microwave Background and Lorentz Boosts

By

CHI PO CHOI

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Ethan Anderes, Chair

Jie Peng

Thomas Lee

Committee in Charge

2024

To infinity and beyond

Contents

Chi Po Choi
September 2024
Statistics

Statistics of Cosmic Microwave Background and Lorentz Boosts

Abstract

This dissertation focuses on improving the estimation of the peculiar velocity within the Cosmic Microwave Background (CMB) by introducing a novel method based on the Lorentz Lie group and the Newton Method. We We begin with an introduction to the CMB and its mathematical representation using Isotropic Gaussian Random Fields. We then discuss the limitations of the Quadratic Estimator in accurately capturing the effects of peculiar velocity. To address these limitations, we propose a Maximum Likelihood Estimation (MLE) approach that leverages the Lorentz Lie group framework and utilizes the Newton Method to reduce estimation errors. The effectiveness of this new estimator is demonstrated through computer simulations, which show a reduction in Root-Mean-Square Error for peculiar velocity estimation in the CMB. This research contributes to a deeper understanding and more accurate measurement of peculiar velocity within the CMB context.

Acknowledgments

I thank everyone who has supported me throughout my PhD journey. I am particularly indebted to Professor Ethan Anderes for his invaluable guidance and mentorship. I also thank my friends and fellow mathematicians I met in the Chinese University of Hong Kong Mathematics Department, especially Man Chun Lee and Kwok Kun Kwong, for their invaluable assistance with the proof of the Flow formulation of Lorentz Boosts.

CHAPTER 1

Introduction

1.1. Outline of Dissertation

This dissertation aims to explore the detailed relationship between statistical models of the Cosmic Microwave Background (CMB) and the effects of Lorentz Boosts on these models. The work is divided into six main chapters. Chapter 1 provides an overview and outlines the motivation for this dissertation. Chapter 2 focuses on the CMB, explaining the statistical models used in this area. This chapter directly relates to the "Statistics of Cosmic Microwave Background" part of the thesis title. Chapter 3 explores Lorentz Boosts and how they change our understanding and observations of the CMB, connecting with the second component of the thesis title "Lorentz Boost". Chapter 4 is about the Quadratic Estimator, a key statistical tool for analyzing CMB data. This chapter details the statistical methods commonly used by physicists in the study of the CMB. Chapter 5 brings in the concept of Lie Group Formulation, offering a more advanced math framework that enriches the statistical models talked about in earlier chapters. The final chapter, Chapter 6, presents numerical experiments that assess the validity of the theories and methods discussed in the preceding chapters. Each chapter aims to tackle specific research questions using a mix of mathematical modeling, statistical analysis, and numerical experiments.

1.2. Outline of Chapter 2: Cosmic Microwave Background

The chapter starts with a basic introduction to the CMB, setting the stage for more detailed topics that come later in the dissertation. It highlights the importance of the CMB in science and explains key features of the CMB, such as its near-uniformity and the subtle temperature variations observed across the sky.

The chapter explores the mathematical and statistical models employed in the study of the CMB. It explains how Isotropic Gaussian Random Fields are used to describe the CMB and goes into the mathematical frameworks behind it, including Gaussian Random Fields on a sphere and

spherical harmonics. A simplified model of the CMB on a circle is also presented to aid comprehension.

The chapter then addresses the peculiar velocity problem on CMB. It introduces how the movement of an observer can affect what we see in the CMB. This includes the relativistic light aberration and relativistic Doppler effect. It then goes over how to measure these effects using CMB data and concludes by discussing the difficulties in making these measurements.

The latter part of the chapter introduces the mathematical frameworks used in the statistical models for studying the CMB. It discusses functional spaces on circles and spheres, Fourier series, spherical harmonics, and Gaussian Random Fields. The chapter ends by discussing some details of the mathematics of the toy circular model.

1.3. Outline of Chapter 3: Lorentz Boosts

The chapter begins with an introductory section, setting the stage for the discussions on Lorentz Boosts and their relation to the CMB. The chapter provides a comprehensive understanding of Lorentz Boosts, from the basics to more advanced topics including their mathematical formulations and physical implications, particularly in the context of the CMB.

The chapter covers two important physical concepts: Relativistic Light Aberration and the Relativistic Doppler Effect. These are key ideas from special relativity that help explain the changes we observe due to relative motion. To make the subject easier to understand, the chapter introduces a simplified model of the CMB on a circle to examine the effects of Lorentz Boosts, before moving on to more advanced spherical formulations.

Moreover, the chapter introduces the concept of "Flow Formalism" to examine Lorentz Boosts. It starts with an introduction to the concept of Flow and then explains how rotation can be viewed in this formalism. This formalism is used to reframe the understanding of light aberration and Lorentz Boosts in general manifolds. This section provides a generalized theorem for Lorentz Boosts based on this formalism. It also discusses Lorentz Boosts on a sphere, exploring the representation of Lorentz Boosts in spherical harmonics basis. It also discusses the unitarity of the Lorentz Boost Operator.

Finally, the chapter concludes by a section that talks about the relationship between the dipole in the CMB and the newly proposed concepts Lorentz Beats.

1.4. Outline of Chapter 4: Quadratic Estimator

The chapter starts with an introduction that offers a broad view of what Quadratic Estimators are and why they matter. It emphasizes their importance in the context of the CMB as well as other statistical challenges. The chapter aims to provide a clear understanding of Quadratic Estimators, discussing their use in least squares problems, their connection to the Fisher Scoring Algorithm, and their specialized application for measuring peculiar velocities in the CMB.

The chapter examines Quadratic Estimators within the framework of least squares problems. This section details how these estimators can be used to address optimization problems that are often encountered in statistical analysis in physics. Following this, the chapter explores the relationship between the Quadratic Estimator and the Fisher Scoring Algorithm. This section demonstrates the equivalence of the Fisher Scoring Algorithm and Quadratic Estimators.

The chapter concludes by diving into the mathematical aspects of using Quadratic Estimators to measure peculiar velocities in the CMB. It is divided into two main parts: Least Square Formulation and Fisher Scoring Algorithm. The first part, Least Square Formulation, outlines how the least squares method can be applied to estimate the peculiar velocity in the CMB, establishing the mathematical framework. The second part, Fisher Scoring Algorithm, focuses on applying this algorithm to the least squares formulation for estimating peculiar velocities in the CMB.

1.5. Outline of Chapter 5: Lie Group Formulation

The chapter introduces Lie groups and discusses their relevance to the dissertation's focus, particularly in the context of the CMB and statistical models. The chapter provides an overview of Lie groups, emphasizing their importance in the study of covariance matrices and their connection to advanced statistical techniques such as maximum likelihood estimation and the newly proposed Newton-Wigner Iteration method.

The chapter describes the space of covariance matrices. It explores how covariance matrices can be formulated in the framework of Lie groups, providing a mathematical foundation for the discussions that follow. The chapter introduces the Lorentz Lie Group and discuss the mathematical

properties and importance of the Lorentz group, relating it to concepts to formulate statistical problem on estimating peculiar velocity in CMB.

The chapter discusses Maximum Likelihood over Lie Groups, outlining how Lie group techniques can be applied when working with maximum likelihood estimation. This includes novel ways to perform these estimations and the advantages of using the Lie group formalism.

The chapter concludes with a section on the newly proposed Newton-Wigner Iteration, describing the concept of Wigner rotation related to Lorentz Boosts, explaining its relevance and how it fits into iterative Newton method applied to Maximum Likelihood estimation of peculiar velocity in the CMB.

1.6. Outline of Chapter 6: Numerical Experiments

The chapter presents numerical experiments to investigate the effects of Lorentz Boosts on the CMB and to assess the accuracy of the proposed estimators using Root-Mean-Square (RMS) Error evaluations.

The chapter presents illustrations of the peculiar velocity effect and numerically demonstrates the advantages and limitations of the Lie Group formulation of Lorentz Boosts on the CMB.

It also explores the impact of Lorentz Boosts on the variance of Gaussian Random Fields through numerical experiments. This includes discussions on the smoothing effect observed in the spectrum of Gaussian Random Fields and the truncation of Lorentz Boosts.

The chapter concludes with a numerical study on the Root-Mean-Square Error of the new estimator, covering both theoretical and empirical aspects. It displays the Root-Mean-Square errors obtained from the numerical experiments, allowing for comparison with the theoretical estimates.

CHAPTER 2

Cosmic Microwave Background

2.1. Introduction

This chapter explores the Cosmic Microwave Background (CMB), first observed by Penzias and Wilson [**PW65**], and later explained in more detail by Dicke et al. [**DPRW65**]. The CMB is a leftover signal from the Big Bang, and it provides important information for studying the early universe.

The chapter is organized as follows: Section 2.2 starts with a detailed review of the basic science related to the CMB. This includes an introduction to the main theories used to study the CMB. In Section 2.3, we look at the mathematical ways to describe the CMB. We focus on the isotropic Gaussian Random Field, a key formulation used to explain the statistical features of the CMB. This section defines this concept and explore its role in modeling the CMB. Section 2.4 examines the Lorentz Boost, which is a combination of the Doppler effect and light aberration, both of which are vital for understanding how we observe the CMB. We also introduce a simple, one-dimensional example of the CMB to help explain statistical methods. We further discuss the Doppler effect and light aberration within more complex models, using a circular model as a specific example.

2.2. Science of the Cosmic Microwave Background

This section introduces Cosmology, focusing on the CMB and its importance as evidence for the Big Bang theory. We will also discuss how statisticians can contribute to Cosmology, especially in data analysis and statistical model development.

Cosmology, part of astrophysics, explores the large-scale evolution of the universe. The CMB, a faint radiation that fills the cosmos, is thought to be a leftover from the early universe. It helps us understand the early conditions and processes of the universe.

Statisticians can contribute to cosmology by analyzing the vast amount of data from telescopes and other instruments to extract valuable information. Their work with the fluctuations in the CMB can help us understand the universe's fundamental properties.

2.2.1. Nobel Prizes Associated with the Cosmic Microwave Background. The field of CMB research has been honored with the Nobel Prize in Physics several times, showing its importance in astrophysics. The CMB was first proposed in 1948 and later confirmed in 1964. This discovery by Arno Penzias and Robert Woodrow Wilson earned them a joint Nobel Prize in Physics in 1978.

Further investigations have continued to deepen our understanding of the CMB. In 1992, the COBE DMR experiment detected variations in the CMB's temperature. This discovery revealed the complexity of the CMB data, highlighting the importance of studying both temperature and polarization maps.

These careful observations were recognized with the 2006 Nobel Prize in Physics awarded to John C. Mather and George F. Smoot. Their work with the COBE satellite revealed key characteristics of the CMB radiation, playing a vital role in turning cosmology into a precise, quantitative discipline.

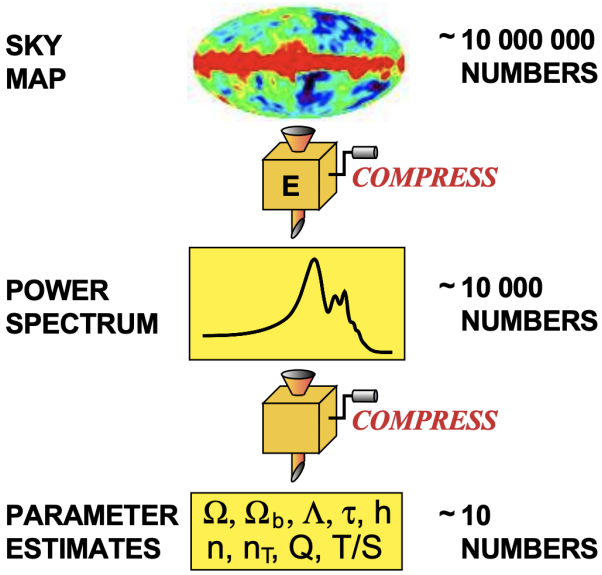
2.2.2. The Uniformity and Anisotropy of the Cosmic Microwave Background. The CMB radiation is the afterglow of the Big Bang. It spreads across the entire sky, showing remarkably uniform properties. Its temperature, consistent with blackbody radiation, provides a unique map of the sky. The temperature of the CMB is nearly uniform, with small fluctuations superimposed. These fluctuations can be modeled using a Gaussian random field.

Anisotropy refers to the variations in temperature of the CMB in different directions. The temperature map on the sphere is considered isotropic if the covariance between any two points depends solely on the angular distance between them. This covariance function can be decomposed using Legendre polynomials, and the variances become the coefficients of this decomposition. This sequence, known as the power spectrum in physics, is essential for analyzing the CMB.

Illustration adapted from [Teg97]. This pipeline visually showcases the process of extracting critical information from the vast amount of data in the CMB observation. Starting with the

SKY MAP, which consists of approximately 10 million numbers representing various observations of the CMB, the data undergoes a compression process to distill the information into the POWER SPECTRUM, reducing the data size to around 10,000 numbers. This spectrum gives a clearer picture of the energy distribution across different frequencies. From here, another compression step takes place to derive the most crucial cosmological parameters, boiling the data down to roughly 10 key numbers. These parameters, such as Ω, Ω_b, Λ, and others, offer insights into fundamental aspects of the universe's structure and behavior, serving as a foundation for many cosmological theories and discussions.

FIGURE 2.1. Illustration adapted from [**Teg97**]. This pipeline visually showcases the process of extracting critical information from the vast amount of data in the CMB observation. Starting with the SKY MAP, which consists of approximately 10 million numbers representing various observations of the CMB, the data undergoes a compression process to distill the information into the POWER SPECTRUM, reducing the data size to around 10,000 numbers. This spectrum gives a clearer picture of the distribution across different frequencies. From here, another compression step takes place to derive the most crucial cosmological parameters, boiling the data down to roughly 10 key numbers. These parameters, such as Ω, Ω_b, Λ, and others, offer insights into fundamental aspects of the universe's structure and behavior, serving as a foundation for many cosmological theories and discussions.



The objective of measuring the temperature map is to estimate the spectrum of the underlying isotropic Gaussian Random Field, allowing us to infer the cosmological parameters. This process requires rigorous statistical methods.

The dipole anisotropy is primarily attributed to the relativistic Doppler effect and light aberration, both arising from the Milky Way's motion relative to the CMB rest frame. The anisotropy introduced by this peculiar velocity goes beyond what can be modeled by an isotropic Gaussian random field.

The relativistic light aberration effect on the celestial sphere can be modeled as a conformal mapping. When combined with the SO(3) rotation on the celestial sphere, it forms a space of diffeomorphisms isomorphic to the Lorentz group. The term "Lorentz Boost" describes the effect of pure relativistic light aberration without rotation.

A primary goal of this dissertation is to investigate the impact of Lorentz Boosts on the Isotropic Gaussian Random Field model of the CMB, thereby enhancing our understanding of both the CMB itself and the early universe.

2.3. Observational Model of CMB

This section discusses the modeling of the CMB as an isotropic Gaussian Random Field on the sphere. We investigate the statistical properties of the CMB as a Gaussian Random Field, which facilitates inference on cosmological models. We consider the CMB without the impact of peculiar velocity as the "primordial CMB".

2.3.1. Isotropic Gaussian Random Field. We describe the isotropic Gaussian Random Field (GRF) on the sphere $\mathbb{S}^2$, a mathematical model essential for understanding the cosmic microwave background in terms of statistics.

2.3.1.1. *Gaussian Random Field on the Sphere.* A Gaussian Random Field is a random function defined on a metric space S, in this case, the sphere $\mathbb{S}^2$, with metric $d(\cdot, \cdot)$, representing the great circle distance on the sphere. We define the function $T : \Omega \times \mathbb{S}^2 \to \mathbb{R}$, where Ω is the sample space, representing all possible outcomes of the random function. The fixed points $x_1, \ldots, x_n \in \mathbb{S}^2$ are specific locations on the sphere, and the vector $(T(x_1), \ldots, T(x_n))$ follows a multivariate Gaussian distribution. The function T is isotropic if the mean and covariance matrix of $(T(x_1), \ldots, T(x_n))$ and $(T(gx_1), \ldots, T(gx_n))$ are identical for any rotation g matrix on the sphere.

2.3.1.2. *Spherical Harmonics and Orthonormal Basis.* To construct a Gaussian random field concretely, we first need to introduce an orthogonal or orthonormal basis on the space of square-integrable (L^2) functions on the sphere, circle, or SO(3). Let $\langle f, g\rangle$ be an inner product of the functional space, and the norm is defined as $\|f\| = \sqrt{\langle f, f\rangle}$. An orthonormal basis of the space of functions on the sphere, circle, or SO(3) is a set of functions $\{f_k\}$ such that $\langle f_k, f_{k'}\rangle = \delta_{kk'}$.

Spherical harmonics play a crucial role in this context. They are defined by $\{Y_m^\ell : \ell = 0, \ldots, \infty, m = -\ell, \ldots, \ell\}$, where ℓ is the degree of the harmonic, representing the number of oscillations on the sphere's surface, and m is the order of the harmonic.

Here are some examples of the orthonormal basis in the sphere, circle, and SO(3):

- The Fourier basis $\{e^{ik\theta} : k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(\mathbb{S}^1)$, the space of square-integrable functions on the circle, with respect to the inner product $\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\bar{g}(\theta)d\theta$.
- The spherical harmonics $\{Y_m^\ell : \ell \in \mathbb{N}_0, m = -\ell, \ldots, \ell\}$ form an orthonormal basis of $L^2(\mathbb{S}^2)$, the space of square-integrable functions on the sphere.
- The Wigner-D functions $\{D^\ell_{mm'} : \ell \in \mathbb{N}_0, m, m' = -\ell, \ldots, \ell\}$ form an orthonormal basis of $L^2(\mathrm{SO}(3))$, the space of square-integrable functions on the rotation group SO(3).

2.3.1.3. *Temperature Map and Isotropic Gaussian Random Field.* The temperature map of the Cosmic Microwave Background is described as a real-valued function on the sphere, $f : \mathbb{S} \to \mathbb{R}$, where $\mathbb{S}$ denotes the space of the sphere. This temperature map can be decomposed using spherical harmonics $\{Y_\ell^m\}$.

A square-integrable function on the sphere, $f(x) \in \mathbb{C}$ for $x \in \mathbb{S}$, can be decomposed by spherical harmonics basis functions as:

$$f(x) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} a_\ell^m Y_\ell^m(x)$$

Here, the coefficients $\{a_\ell^m\}$ are the spherical harmonic coefficients of the function f, and the transformation from f to $\{a_\ell^m\}$ is called the spherical harmonics transform. If the coefficients satisfy $a_\ell^m = (-1)^m a_\ell^{-m*}$, then f is a real-valued function suitable for modeling the temperature map.

We describe the construction of an isotropic Gaussian Random Field on the sphere, following the comprehensive mathematical details in [**LS15**, **LTT18**]. Let C_ℓ be a sequence of positive real numbers with a finite sum $\sum_{\ell=0}^{\infty}(2\ell+1)C_\ell$. Suppose a_ℓ^m are complex Gaussian random variables

satisfying:

$$E(a_m^\ell) = 0 \quad \text{and} \quad E(a_m^\ell a_{m'}^{\ell'*}) = C_\ell \delta_{\ell\ell'} \delta_{mm'}.$$

If the coefficients a_ℓ^m satisfy the condition $a_\ell^m = (-1)^m a_\ell^{-m*}$ for a real-valued function, then $f = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_\ell^m Y_\ell^m$ is an isotropic Gaussian Random Field on the sphere. The covariance between any two points x and y is:

$$\mathrm{E}(f(x)f(y)) = \sum_{\ell=0}^{\infty} C_\ell \frac{2\ell+1}{4\pi} P_\ell(\langle x_1, x_2 \rangle)$$

where P_ℓ are Legendre polynomials, and $\langle x_1, x_2 \rangle = \cos(\theta)$, with θ being the great circle distance between x and y.

2.3.1.4. *E-mode and B-mode Polarization in the Cosmic Microwave Background.* E-mode and B-mode polarization maps can be represented as two real-valued functions on the sphere. These functions can be decomposed using Spin-weighted spherical harmonics, which are a generalization of the standard spherical harmonics to include a spin-weight parameter s. The spin-weight can take integer values, and in the context of the CMB, we are particularly interested in $s = 2$ and $s = -2$.

For a given spin weight s, the spin-weighted spherical harmonics are defined as:

$${}_sY_\ell^m(\theta, \phi) = (-1)^s \sqrt{\frac{(2\ell+1)(\ell-s)!}{4\pi(\ell+s)!}} e^{im\phi} P_\ell^s(\cos\theta),$$

where P_ℓ^s are the associated Legendre polynomials, and s, ℓ, m are the spin-weight, degree, and order of the harmonic, respectively.

For a given map X, the coefficients $a_\ell^{m,X}$ are given by:

$$a_\ell^{m,X} = \int_{\mathbb{S}} X(\theta, \phi) {}_sY_\ell^{m*}(\theta, \phi) \sin\theta \, d\theta d\phi,$$

where X represents the temperature map T or polarization maps E-mode E, and B-mode B, and the integral is over the sphere.

The correlations between temperature T, E-mode, and B-mode are quantified through the power spectra, defined as:

$$C_\ell^{XX'} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a_\ell^{m,X} a_\ell^{m,X'*},$$

where $X, X' \in \{T, E, B\}$, and the asterisk denotes complex conjugation.

In the context of the CMB, the lensing effect from light aberration due to peculiar velocity can have a significant impact on the observed power spectra. The Peculiar velocity of the observer in the universe can cause distortions in the observed CMB, leading to the observed anisotropies in the temperature and polarization maps.

The mathematical framework of E-mode and B-mode polarization, along with the defined power spectra, offers a profound understanding of the CMB. Analyzing these components and their correlations with temperature, considering the lensing effect from light aberration due to peculiar velocity, allows scientists to explore the universe's dynamics and structure. This analysis can both lead to a more comprehensive understanding of cosmic phenomena and contribute to the ongoing refinement of cosmological models.

2.3.1.5. *Toy Model of CMB on the Circle.* In the section on Isotropic Gaussian Random Fields, we introduced the concept of constructing Gaussian random fields on various spaces, including the sphere and the circle. While the spherical model is essential for a realistic representation of the CMB, it can be mathematically intricate.

To provide a more tractable model that aligns with the principles outlined in the Isotropic Gaussian Random Field section, we develop a toy model of the CMB defined on the circle instead of the sphere. This model leverages the Fourier basis, previously mentioned, which is more straightforward to manipulate. Let us consider a sequence of real positive numbers C_k such that $C_k = C_{-k}$ and $\sum_k C_k < \infty$ where k is an integer index. We can define a sequence of Gaussian random variables a_k that satisfies:

$$E[a_k] = 0 \quad \text{and} \quad E[a_k a_{k'}^*] = C_k \delta_{kk'},$$

where E denotes the expectation. Then, we can define a function $f(\theta)$ on the circle as:

$$f(\theta) = \sum_k a_k e^{ik\theta},$$

which represents an isotropic Gaussian random field on the circle, analogous to the construction on the sphere. To ensure that f is a real-valued function, we impose the condition:

$$a_k = a_{-k}^*.$$

This toy model on the circle shares essential characteristics with the spherical model described in the Isotropic Gaussian Random Field section. The use of the Fourier basis simplifies the mathematical treatment, making it an excellent tool for understanding the underlying mathematics. The insights gained from this toy model can be applied to the more complex spherical model, providing a bridge between the intuitive understanding and the rigorous mathematical framework required for the analysis of the CMB. This approach allows for a more comprehensive exploration of the lensing effect from light aberration due to peculiar velocity, as well as other cosmological phenomena.

2.3.2. CMB as Isotropic Gaussian Random Field. We model the primordial CMB, the CMB without observation noise or distortion, as an isotropic Gaussian random field (GRF) on the sphere $\mathbb{S}^2$. A Gaussian random field f is defined such that for any integer k and any points $x_1, \ldots, x_k \in \mathbb{S}^2$, the vector $(f(x_1), \ldots, f(x_k))$ follows a multivariate Gaussian distribution. The field f is isotropic if the covariance between any two points x_1 and x_2 on the sphere is given by

$$\mathrm{Cov}(f(x_1), f(x_2)) = \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} C_\ell P_\ell(\cos(d(x_1, x_2))),$$

where $C_\ell \geq 0$ is the power spectrum of the field, defined as

$$C_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \mathrm{E}[a_{\ell m} a_{\ell m}^*],$$

P_ℓ is the Legendre polynomial of degree ℓ, and $d(x_1, x_2)$ represents the great circle distance between x_1 and x_2 on the sphere. The sequence C_ℓ, also known as the power spectrum of the CMB, is dependent on the physical cosmology model defined by physicists. Analyzing the CMB temperature map often involves estimating the power spectra C_ℓ, which in turn affects the cosmological parameters. Given the measured primordial CMB spherical harmonic coefficients $a_{\ell m}$, the maximum likelihood

estimators of C_ℓ are

$$\hat{C}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a_{\ell m} a^*_{\ell m}.$$

In some tasks, we assume fiducial values for the power spectrum $C_\ell$and estimate other unknowns in the model.

The statistical model of the CMB is a Gaussian Random Field on the sphere, expressed as a function with spherical harmonics coefficients $(a_{\ell m})$ being a Gaussian random vector. In the case of an isotropic Gaussian random field, the covariance between coefficients can be expressed as

$$C_{\ell\ell' mm'} := \mathrm{Cov}(a_{\ell m}, a_{\ell' m'}) = \mathrm{E}[a_{\ell m} a^*_{\ell' m'}] = \delta_{\ell\ell'}\delta_{mm'} C_\ell, \tag{2.1}$$

in other words, $(a_{\ell m}) \sim N(\mathbf{a}_0, \mathbf{C})$, where $\mathbf{a}_0$ is a mean vector (often assumed to be zero) and $\mathbf{C}$ is a diagonal matrix with entries C_ℓ. The spherical harmonic coefficients $a_{\ell m}$ are independent and uncorrelated, with the variance of each $a_{\ell m}$ depending only on the index ℓ.

The observed CMB temperature map incorporates additive noise, characterized as an isotropic Gaussian random field with coefficients $n_{\ell m} \sim N(0, N_\ell)$. We assume the noise to be isotropic, a standard assumption for noise in observed CMB data. The observation can be expressed as

$$\mathbf{d} = \mathbf{a} + \mathbf{n}, \tag{2.2}$$

where $\mathbf{d}$, $\mathbf{a}$, and $\mathbf{n}$ represent the coefficients of the observed CMB temperature map, the primordial CMB temperature map, and the noise, respectively. The observation process may include a Gaussian beam that attenuates high-frequency components of the CMB, resulting in a non-constant noise variance across frequency. However, in this work, we assume the noise variance is constant across ℓ, i.e. N_ℓ is constant. This assumption is crucial in this dissertation, as it facilitates the derivation of specific estimators and their theoretical properties. We assume that the isotropic Gaussian noise is invariant under peculiar velocity transformations. This simplification allows for a more tractable mathematical treatment and leads to simpler expressions for the estimators. The necessity of this assumption for ensuring invariance under peculiar velocity transformations will be discussed in subsequent chapters. Furthermore, portions of the CMB on the sphere may be unobserved or masked due to foreground contamination, primarily from the Milky Way. While

this masking can introduce uncertainty and bias in spectrum estimation, this dissertation does not address such cases. The assumption of an isotropic Gaussian noise model remains central to the methodology and theoretical framework presented in this work.

2.4. Peculiar Velocity Effects on the CMB

2.4.1. Introduction to Peculiar Velocity and Its Relevance to CMB.

2.4.1.1. *Definition and Significance.* Peculiar velocity refers to the velocity of an object relative to a rest frame, which in the context of cosmology, is often the CMB rest frame. This concept is vital in understanding the observed anisotropies in the CMB, a relic radiation from the Big Bang that fills the universe.

The study of peculiar velocity is essential for understanding the large-scale structure and dynamics of the universe, including the motion of galaxies. The Lorentz Boost, a combined effect of the Relativistic Doppler Effect and Relativistic Light Aberration, plays a crucial role in the observed characteristics of the CMB.

2.4.1.2. *Relativistic Doppler Effect and Light Aberration.* The Relativistic Doppler Effect refers to the change in frequency or wavelength of a wave in relation to an observer moving relative to the wave's source. In the context of the CMB, this effect causes a shift in the observed temperature.

Relativistic Light Aberration refers to the apparent change in the direction of light emitted from a moving source as observed in a different inertial frame of reference. This effect alters the apparent position of celestial objects and also influences how we observe the CMB.

These relativistic effects, collectively known as the Lorentz Boost, have significant implications for the observed properties of the CMB, particularly in relation to its modeling as an isotropic Gaussian Random Field.

2.4.1.3. *Impact on CMB Dipole.* The dipole of the CMB refers to the largest-scale variation in its temperature, and it is measured to be much larger than the other multipoles, which are smaller-scale variations. Cosmologists assume that this dipole is solely attributed to the Lorentz Boost due to our peculiar velocity with respect to the CMB rest frame.

By combining the measured temperatures of WMAP dipole [**HWH$^+$09**] with the COBE monopole [**Lin96**, **MFS$^+$99**], the direction of this peculiar velocity can be inferred to be $l = 263.99° \pm 0.14°$, $b = 48.26°$ in galactic coordinates, and its modulus to be $\beta \equiv |\boldsymbol{\beta}| = (1.231 \pm 0.003) \times 10^{-3}$.

2.4.1.4. *Illustration of Peculiar Velocity Effect.* Figure 2.3 illustrates the effect of peculiar velocity on the isotropic Gaussian Random Field. The images depict the Cosmic Microwave Background as observed under varying magnitudes of simulated peculiar velocity. The leftmost image represents the CMB without peculiar velocity, while the subsequent images show the effect of increasing peculiar velocity. This effect appears as a contraction toward one point and an expansion in the opposite direction on the sphere, visually illustrating the combined influence of the Lorentz Boost.

2.4.2. Measurement of Peculiar Velocity.

2.4.2.1. *Introduction to Measurement Techniques.* Measuring peculiar velocity is a nuanced task that involves the integration of various cosmological observations and statistical methods. This section presents an overview of the primary techniques employed in the measurement of peculiar velocity, with a particular emphasis on leveraging CMB observations and statistical methodologies.

2.4.2.2. *Basic Measurement Using Dipole Observations.* In cosmology, the term "dipole" refers to a specific pattern of temperature variation in the CMB. Early observations, such as those from the Cosmic Background Explorer (COBE) and Wilkinson Microwave Anisotropy Probe (WMAP), have allowed scientists to measure this dipole pattern with great accuracy.

The basic method of measuring peculiar velocity assumes that the observed dipole is solely due to our motion relative to the CMB rest frame. By assuming that the observed CMB dipole is solely due to our peculiar velocity, we can estimate its direction and magnitude. This approach is justified by the precise measurement of the CMB dipole and the underlying assumptions of standard cosmological models.

2.4.2.3. *Advanced Measurement Using High-Resolution Data.* With the advent of more high-resolution observations, such as those from the Planck satellite, new methods have been developed to measure peculiar velocity. These methods utilize the high-frequency power spectrum data of the CMB, which provides detailed information about temperature variations on small scales.

One such method is the Quadratic Estimator, which leverages the high-frequency power spectrum data to create a more accurate estimation of peculiar velocity. This method has been described in various papers, including [**CvL02**, **KK11**, **PYSS10**].

2.4.2.4. *Introduction of Maximum Likelihood Estimator.* In addition to these methods, this dissertation will introduce a maximum likelihood estimator, a statistical approach that aims to find the values of the parameters that maximize the likelihood function. This method offers a robust and efficient way to estimate peculiar velocity, aligning with the statistical focus of the research.

2.4.2.5. *Conclusion.* The measurement of peculiar velocity has evolved from basic methods relying on dipole observations to more advanced techniques utilizing high-resolution data. By analyzing the temperature variations in the CMB and employing statistical techniques like the Quadratic Estimator and maximum likelihood estimation, we can obtain a comprehensive understanding of peculiar velocity. These methods not only contribute to our broader understanding of the universe but also offer exciting opportunities for statistical innovation and application.

2.4.3. Challenges and Contribution.

2.4.3.1. *Challenges in Measuring Peculiar Velocity.* The measurement of peculiar velocity, although advanced with the development of new techniques and high-resolution data, still presents several challenges:

- Assumptions and Limitations: The basic method of measuring peculiar velocity relies on the assumption that the observed dipole is solely due to our motion relative to the CMB rest frame. This assumption may not always hold, leading to potential inaccuracies.
- Utilization of High-Frequency Data: While high-frequency data from observations like Planck provide valuable insights, utilizing these data effectively can be complex. The Quadratic Estimators, although powerful, may not attain optimal efficiency in terms of root-mean-square error when the peculiar velocity is not small.
- Statistical Complexity: The application of advanced statistical methods requires careful consideration of the underlying statistical properties of the data. Ensuring that these methods are both accurate and computationally efficient is a challenging task.

2.4.3.2. *Contribution of This Dissertation.* This dissertation makes a significant contribution to the field of peculiar velocity measurement by introducing a Maximum Likelihood estimator. The Maximum Likelihood estimator addresses a limitation of Quadratic Estimators by achieving optimal efficiency in terms of Root-Mean-Square Error, even for non-small peculiar velocities. By leveraging the valuable high-frequency data more effectively, this method offers a more robust and precise estimation of peculiar velocity.

The introduction of the Maximum Likelihood estimator not only addresses some of the existing challenges but also opens new avenues for statistical innovation in cosmology. It represents a blend of statistical rigor and physical understanding, tailored to provide insights into the challenging phenomenon of peculiar velocity.

2.5. Mathematics of Gaussian random field

We detail the mathematics of the Gaussian Random Field on the circle and sphere, which model the primordial CMB. The Gaussian random field on the circle serves as a simplified toy model of CMB, reducing the complexity of the sphere model while still retaining essential features of the peculiar velocity problem on CMB in 2D version.

2.5.1. Functions on Circle. We begin by describing the space of functions on the circle $\mathbb{S}^1$. We define the inner product on functions on the circle $\mathbb{S}^1$. Let $f : \mathbb{S}^1 \to \mathbb{C}$ and $g : \mathbb{S}^1 \to \mathbb{C}$ be two complex-valued functions on the sphere $\mathbb{S}^1$. The inner product of f and g is defined by the integral:

$$\langle f, g\rangle := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\overline{g(\theta)}\, d\theta,$$

where $\overline{g(\theta)}$ is the complex conjugate of $g(\theta)$. If f and g are real-valued functions, the complex conjugation can be ignored. Note that $\sqrt{\langle f, f\rangle}$ is real for any complex-valued function f. The L^2 norm of a function f is given by

$$\|f\|_2 := \sqrt{\langle f, f\rangle}. \tag{2.3}$$

We define the L^2 space of functions on the circle $\mathbb{S}^1$ to be the collection of all functions on the circle with finite L^2 norm, i.e., $L^2_{\mathbb{C}}(\mathbb{S}^1) = \{f : \langle f, f\rangle < \infty\}$. By the polarization identity, the inner product $\langle f, g\rangle$ is finite for any $f, g \in L^2_{\mathbb{C}}(\mathbb{S}^1)$.

We consider the Fourier basis of $L^2_{\mathbb{C}}(\mathbb{S}^1)$ to be the set of functions:

$$\text{Fourier}(N) = \{e_k : k = -N, \dots, -1, 0, 1, \dots, N\}, \tag{2.4}$$

$$e_k(\theta) = \exp(ik\theta) \quad \text{for } \theta \in [-\pi, \pi). \tag{2.5}$$

Note that the set of Fourier basis Fourier(N) is orthonormal in $L^2_{\mathbb{C}}(\mathbb{S}^1)$, i.e.,

$$\langle e_k, e_{k'} \rangle = \delta_{kk'} = \begin{cases} 1 & \text{if } k = k' \\ 0 & \text{if } k \neq k' \end{cases}.$$

Here, N represents the highest frequency considered in the Fourier basis.

Let $f \in L^2_{\mathbb{C}}(\mathbb{S}^1)$. f admits a Fourier decomposition

$$f = \sum_{k=-\infty}^{\infty} \langle f, e_k \rangle e_k.$$

We call the complex sequence $f_k = \langle f, e_k \rangle$ the Fourier coefficients of f.

Suppose $f \in L^2_{\mathbb{R}}(\mathbb{S}^1)$ is a real-valued function on the sphere $\mathbb{S}^1$. The Fourier coefficients f_k of f satisfy $f_k = \overline{f}_{-k}$ for all k. Conversely, if a function f has Fourier coefficients f_k satisfying $f_k = \overline{f}_{-k}$, then f is real-valued.

2.5.2. Fourier series. Denote $\mathbb{S}^1$ as the unit circle. A point on the unit circle $\mathbb{S}^1$ can be represented as $e^{i\theta} \in \mathbb{C}$ for $\theta \in [-\pi, \pi]$. Let $L^2(\mathbb{S}^1)$ be the space of square-integrable complex-valued or real-valued functions on the circle $\mathbb{S}^1$. The norm of $f \in L^2(\mathbb{S}^1)$ is $\|f\|_2 = (\int_{\mathbb{S}^1} f(x) f^*(x)\, dx)^{1/2}$. For any $f, g \in L^2(\mathbb{S}^1)$, the inner product between f and g is defined to be $\langle f, g \rangle := \int_{\mathbb{S}^1} f(x) g^*(x)\, dx$ where $g^*(x)$ is the complex conjugate of $g(x)$. By the parallelogram law, this inner product is well defined, given that the norms of f and g are finite.

Denote the Fourier basis function $e_k : \mathbb{S}^1 \to \mathbb{C}$, which is defined by $e_k(x) := \frac{1}{\sqrt{2\pi}} e^{ikx}$, for $x \in [-\pi, \pi)$ and for integer k. The complex-valued $L^2(\mathbb{S}^1)$ has the orthonormal basis $\{e_k : k = \dots, -1, 0, 1, \dots\}$. For complex-valued $f \in L^2(\mathbb{S}^1)$, the Fourier decomposition is

$$f(x) = \sum_{k \in \mathbb{Z}} f_k e_k(x) \quad \text{for } x \in \mathbb{S}^1,$$

where

$$f_k = \int_{-\pi}^{\pi} f(x) e_k^*(x)\, dx,$$

is the (normalized) Fourier coefficients of f. The infinite sum is defined in the sense that

$$\int_{-\pi}^{\pi} \left| f(x) - \sum_{k=-K}^{K} f_k e_k(x) \right|^2 dx \to 0 \quad \text{as } K \to \infty.$$

For real-valued $f \in L^2(\mathbb{S}^1)$, the Fourier coefficients of f must satisfy

$$f_{-k} = f_k^* \quad \text{for } k \in \mathbb{Z}.$$

We define a set of real orthonormal basis of $L^2(\mathbb{S}^1)$:

$$\{\frac{1}{\sqrt{\pi}} \sin(kx) : k = -1, -2, \dots\} \cup \{\frac{1}{\sqrt{2\pi}}\} \cup \{\frac{1}{\sqrt{\pi}} \cos(kx) : k = 1, 2, \dots\}$$

and the real Fourier decomposition is

$$f(\theta) = \sum_{k<0} a_k \frac{1}{\sqrt{\pi}} \sin(k\theta) + a_0 \frac{1}{\sqrt{2\pi}} + \sum_{k>0} a_k \frac{1}{\sqrt{\pi}} \cos(k\theta)$$

for $a_k \in \mathbb{R}$ and the relationship between f_k and a_k is

$$\begin{aligned} f_k &= a_k \frac{1}{\sqrt{2}} + a_{-k} \frac{1}{\sqrt{2}} i \quad \text{for } k < 0 \\ f_0 &= a_0 \\ f_k &= a_k \frac{1}{\sqrt{2}} - a_{-k} \frac{1}{\sqrt{2}} i \quad \text{for } k > 0. \end{aligned}$$

The transformation between (f_k) and (a_k) is orthogonal/unitary, i.e. $\sum_k f_k f_k^* = \sum_k a_k^2$.

We denote the space of real Fourier coefficients of real-valued function under the above real orthonormal basis by $l^2(\mathbb{S}^1)$. The inner product of $l^2(\mathbb{S}^1)$ is defined to be

$$\langle (a_k), (b_k) \rangle = \sum_{k \in \mathbb{Z}} a_k b_k \tag{2.6}$$

for any $(a_k), (b_k) \in l^2(\mathbb{S}^1)$. The space of $l^2(\mathbb{S}^1)$ with this inner product is isometry to the space of real-valued $L^2(\mathbb{S}^1)$. As an element in $l^2(\mathbb{S}^1)$ can be concretely expressed as a sequence, or an

infinite-dimensional vector, we prefer work inside $l^2(\mathbb{S}^1)$ instead of $L^2(\mathbb{S}^1)$. Moreover, in practice, we restrict ourselves to a finite-dimensional vector space $\{(a_k \text{ for } k = -N, \ldots, N) : (a_k) \in l^2(\mathbb{S}^1)\}$ where N is the highest frequency that we can measure or compute numerically.

2.5.3. Functions on Sphere. We now extend our discussion to the space of functions on the sphere $\mathbb{S}^2$.

Let f and g be complex-valued functions on $\mathbb{S}^2$. Similar to the case of the circle, we can define the inner product, L^2 norm, and Fourier basis for functions on the sphere $\mathbb{S}^2$. Similar to the case of the circle, we define the inner product and L^2 norm for functions on the sphere $\mathbb{S}^2$. The definitions are analogous to those on the circle, but with suitable modifications to accommodate the spherical geometry. Let f and g be complex-valued function on the sphere $\mathbb{S}^2$. The inner product is

$$\langle f, g\rangle = \int_{\mathbb{S}^2} f\bar{g}dS = \int_{-\pi}^{\pi}\int_{0}^{\pi} f(\theta,\phi)\bar{g}(\theta,\phi)\sin\theta d\theta d\phi$$

where the bar $\bar{g}$ denotes the complex conjugate of the function g and the coordinate (θ, ϕ) is a spherical coordinate. For real-valued functions, the inner product is the same definition, but complex conjugation has no effect on the real-valued functions.

The Fourier basis of $L^2_{\mathbb{C}}(\mathbb{S}^2)$ consists of spherical harmonics, which are eigenfunctions of the Laplace-Beltrami operator on the sphere. The details of the spherical harmonics and their properties will be discussed in the following sections.

2.5.4. Spherical Harmonic. The temperature map of the CMB is modeled as a real-valued function on the unit sphere. To analyze CMB, we apply spherical harmonic decomposition.

The space of all complex-valued functions on sphere $\mathbb{S}$ with inner product $\langle\cdot,\cdot\rangle$ is denoted by $L^2(\mathbb{S})$. The spherical harmonics $\{Y_\ell^m : \ell = 0, 1, 2, \ldots, m = -\ell, -\ell+1, \ldots, \ell-1, \ell\}$ forms a set of orthonormal basis of $L^2(\mathbb{S})$. For any $f \in L^2(\mathbb{S})$, f can be decomposed as

$$f = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\langle f, Y_{\ell m}\rangle Y_\ell^m.$$

The space of functions on SO(3) can be generated by Wigner D function [**KR08**]:

$$f(\phi,\theta,\psi) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\sum_{m'=-\ell}^{\ell} f^\ell_{mm'}D^\ell_{mm'}(\phi,\theta,\psi).$$

The relationship between spin-weighted spherical harmonics and Wigner D function is

$$D^{\ell}_{sm}(\phi,\theta,\psi) = \sqrt{\frac{4\pi}{2\ell+1}} e^{is\psi} {}_{-s}Y^{\ell}_{m}(\theta,\phi)$$

Denote by $L^2(\mathbb{S}^2)$ the square-integrable complex-valued functions on the sphere. We use the spherical coordinate system $(\theta,\phi) \mapsto (\sin\theta\cos\phi, \sin\theta, \sin\phi, \cos\theta) \in \mathbb{S}^2$ for $\theta \in [0,\pi]$ and $\phi \in [0, 2\pi)$.

For $f, g \in L^2(\mathbb{S}^2)$, define inner product

$$\langle f, g\rangle_{\mathbb{S}^2} := \int f(\theta,\phi) g^*(\theta,\phi) \sin\theta d\theta d\phi.$$

For $\ell = 0, 1, \ldots$ and $m = -\ell, \ldots, \ell$, the spherical harmonics basis functions $\{Y^m_\ell\} \subset L^2(\mathbb{S}^2)$ are defined as:

$$Y^m_\ell(\theta,\phi) := \frac{(-1)^m}{\sqrt{2\pi}} N_{\ell m} P^m_\ell(\cos\theta) e^{im\phi}$$

where $N_{\ell m} = \sqrt{\frac{(2\ell+1)}{2}\frac{(\ell-m)!}{(\ell+m)!}}$ and P^m_ℓ is the associated Legendre polynomial.

Spherical Harmonics form a set of orthonormal basis $\{Y^m_\ell\}$ on $L^2(\mathbb{S}^2)$ with respect to the inner product $\langle , \rangle_{\mathbb{S}^2}$:

$$\langle Y^m_\ell, Y^{m'}_{\ell'}\rangle_{\mathbb{S}^2} = \delta_{\ell\ell'}\delta_{mm'}.$$

Note that $\{N_{\ell m} P^m_\ell\}$ forms an orthonormal basis for $L^2([-1,1])$, the space of square-integrable functions on the interval $[-1,1]$, with respect to the standard inner product $\langle f, g\rangle = \int_{-1}^{1} f(x) g(x)^* dx$.

For every function $f \in L^2(\mathbb{S}^2)$, we can find unique spherical harmonic coefficients $[a_{\ell m}]$ such that

$$f = \sum_{\ell,m} a_{\ell m} Y^m_\ell.$$

It can be shown that

$$Y^{-m}_\ell = (-1)^m Y^{m*}_\ell.$$

Hence, f is real-valued if and only if spherical harmonic coefficients satisfies

$$a_{\ell,-m} = (-1)^m a^*_{\ell m}.$$

We can define a set of real spherical harmonic basis:

$$\tilde{Y}_\ell^m = \begin{cases} Y_\ell^m & \text{if } m = 0 \\ \frac{(-1)^m}{\sqrt{2}} Y_\ell^m + \frac{1}{\sqrt{2}} Y_\ell^{-m} & \text{if } m > 0 \\ \frac{(-1)^m i}{\sqrt{2}} Y_\ell^m + \frac{-i}{\sqrt{2}} Y_\ell^{-m} & \text{if } m < 0 \end{cases}$$

which is orthonormal with respect to the inner product

$$\langle f, g \rangle := \int_{\mathbb{S}^2} f g d\Omega. \tag{2.7}$$

For any real-valued $f \in L^2(\mathbf{S}^2)$, square-integrable function on $\mathbb{S}^2$, there exists a unique sequence $(a_{\ell m} : \ell = 0, \ldots, m = -\ell, \ldots, \ell)$ representing f

$$f(\mathbf{x}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} \tilde{Y}_\ell^m \quad \text{for all } \mathbf{x} \in \mathbb{S}^2.$$

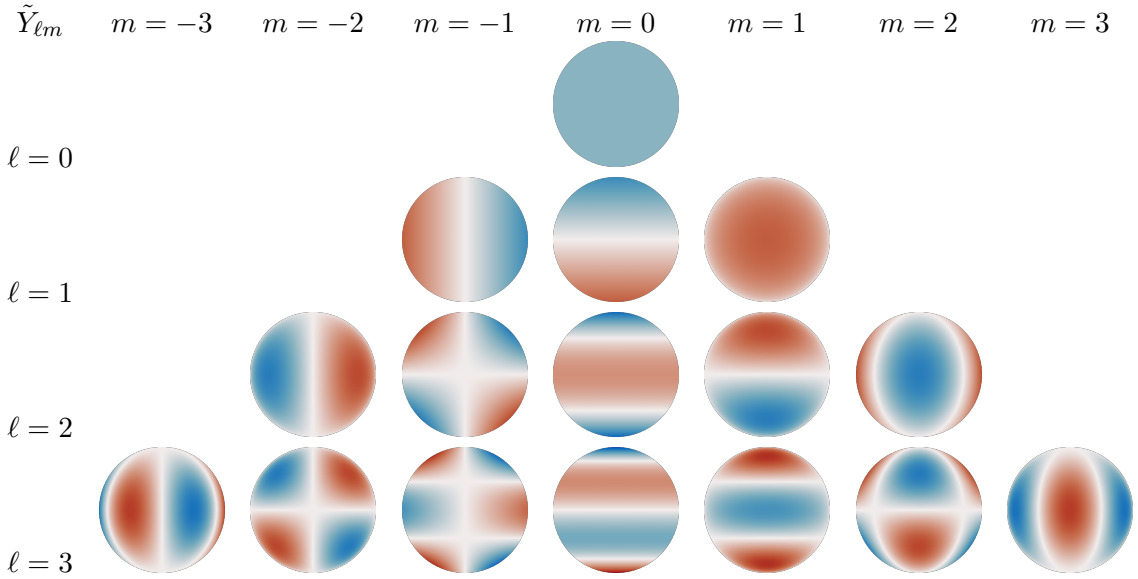


TABLE 2.1. Illustrations of the real spherical harmonic basis.

2.5.5. Gaussian Random Field. We focus on the Gaussian random field defined on the unit circle $\mathbb{S}^1$ and unit sphere $\mathbb{S}^2$.

For a real-valued random field f on $\mathbb{S}^d$ with $E[f(\mathbf{x})] < \infty$ for all $\mathbf{x} \in \mathbb{S}^d$ where $d = 1$ or 2, the covariance function $C(\mathbf{x}, \mathbf{y}) = \mathrm{Cov}(f(\mathbf{x}), f(\mathbf{y}))$ must be positive definite on $(\mathbf{x}, \mathbf{y}) \in \mathbb{S}^d \times \mathbb{S}^d$, i.e.,

$$\mathrm{Var}\left(\sum_{j=1}^{n} \alpha_j f(x_j)\right) = \sum_{j,k=1}^{n} \alpha_j \alpha_k C(x_j, x_k) \geq 0,$$

for all finite n, all $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{S}^d$, and all real $\alpha_1, \ldots, \alpha_n$.

Let $\{e_k : k \in \mathbb{Z}\}$ be the Fourier basis of the unit circle $\mathbb{S}^1$. Let $(C_k : k = 0, 1, 2, \ldots)$ be a sequence of positive numbers such that $\sum_k C_k < \infty$. Suppose we have $f_k \sim N(0, C_k)$ independently. Define a real-valued random function $f : \mathbb{S}^d \to \mathbb{R}$ that $f(x) = \sum_{k \in \mathbb{Z}} f_k e_k(x)$ for $x \in [-\pi, \pi)$. For any $x, y \in [-\pi, \pi)$, the covariance function $\mathrm{Cov}(f(x), f(y))$ is

$$\begin{aligned}
E[f(x)f(y)] &= \sum_{k<0} C_k \frac{1}{\pi} \sin(kx)\sin(ky) + C_0 \frac{1}{2\pi} + \sum_{k>0} C_k \frac{1}{\pi} \cos(kx)\cos(ky) \\
&= \sum_{k<0} C_k \frac{1}{\pi} (\cos(k(x-y)) - \cos(k(x+y))) + C_0 \frac{1}{2\pi} + \sum_{k>0} C_k \frac{1}{\pi} (\cos(k(x-y)) + \cos(k(x+y))) \\
&= C_0 \frac{1}{2\pi} + \sum_{k>0} C_k \frac{1}{\pi} 2\cos(k|x-y|)
\end{aligned}$$

which is positive definitely and only depends on the distance between x and y. f is an isotropic Gaussian random field on the unit circle.

On the other hand, if f is an isotropic Gaussian random field on the sphere with covariance function $\mathrm{Cov}(f(x), f(y)) = C(|x-y|)$, we can apply Fourier transform to $C(|x-y|)$ to obtain the sequence $(C_k : k = 0, 1, 2, \ldots)$ such that $C(|x-y|) = \frac{1}{2\pi} C_0 + \sum_{k>0} C_k \frac{1}{\pi} 2\cos(k|x-y|)$.

2.5.6. 2-dimensional toy model of Gaussian Random Field. In the 2-dimensional toy model of CMB, the statistical model of the Fourier coefficients is given by:

$$\mathbf{f}^{\mathrm{obs}} = \mathbf{f}^{\mathrm{ave}} + \mathbf{f}^{\mathrm{pri}} + \boldsymbol{\epsilon},$$

where $\mathbf{f}^{\mathrm{obs}}$ is the vector of Fourier coefficients of the observed CMB, $\mathbf{f}^{\mathrm{ave}}$ is the vector of Fourier coefficients of the (known or unknown) positive average of primordial CMB, $\mathbf{f}^{\mathrm{pri}}$ is a mean-zero random vector modeling the Fourier coefficients of the fluctuation of primordial CMB, and $\boldsymbol{\epsilon}$ is the observation noise from the instrument.

The term $\mathbf{f}^{\text{ave}}$ is employed to model the average temperature of the CMB, specifically capturing a constant value of 2.71 Kelvin. This constant serves two purposes in the model. First, it represents a non-random value, reflecting the fact that this temperature is a fixed characteristic of the CMB. Second, it is uniform across the sphere, indicating that this temperature is consistent in all directions. Based on these considerations, we assume the form of the CMB temperature map to be:

$$\mathbf{f}^{\text{ave}} = (f_0^{\text{ave}} > 0; f_k^{\text{ave}} = 0 \text{ for } k \neq 0),$$

where f_0^{ave} is assumed to be known and represents this constant temperature. f_0^{ave} may be a fiducial value, which refers to a specific value considered as a fixed or reference point in the analysis. This fiducial value may may be assumed based on the CMB observation, either through mathematical-physics-theoretical or observational-empirical methods [**Fix09**].

The vector $\mathbf{f}^{\text{pri}}$ is used to model the Fourier coefficients of the fluctuation in primordial CMB. $\mathbf{f}^{\text{pri}}$ is a vector of Fourier coefficients of a mean-zero isotropic Gaussian random field. Let f_k^{pri} denote the frequency k component of $\mathbf{f}^{\text{pri}}$ for $k \in \mathbb{Z}$. $f_k^{\text{pri}} \sim N(0, C_k)$ independently for all $k \in \mathbb{Z}$, where $C_k > 0$ is the variance of f_k. The sequence $(C_k : k = 0, 1, \dots)$ is determined by the cosmological model and are considered to be fiducial values, given the set of most probable cosmological models and parameters. In reality, it may be an unknown sequence that requires estimation based on CMB observations and cosmological models. If a cosmological model and corresponding cosmological parameters are assumed, the fiducial values of $(C_k : k = 0, 1, \dots)$ can also be obtained through mathematical-physical theory [**LC11**].

The random vector $\boldsymbol{\epsilon}$ is used to model the observation noise of the instrument. It is again an isotropic Gaussian random field. The Fourier coefficients $\epsilon_k \sim N(0, C_{\text{noise}})$ independently, where values of $C_{\text{noise}} > 0$ are assumed based on the instrument's configuration. The isotropic Gaussian random field implies that the statistical properties of the observation noise are invariant under rotations in this context.

FIGURE 2.2. Illustrations of the Cosmic Microwave Background as an Isotropic Gaussian Random Field. The top image employs an orthographic projection, rendering the heat map on a 3D sphere. In contrast, the bottom image utilizes a Mollweide projection, an area-preserving transformation that projects the sphere onto a flat 2D plane. It is essential to note the reversed perspectives in these projections: the orthographic view represents an "outside looking in" perspective, while the Mollweide view adopts an "inside looking out" perspective aligned with galactic coordinates. The orthographic projection shows only half of the sphere, while the Mollweide projection shows the entire sphere.

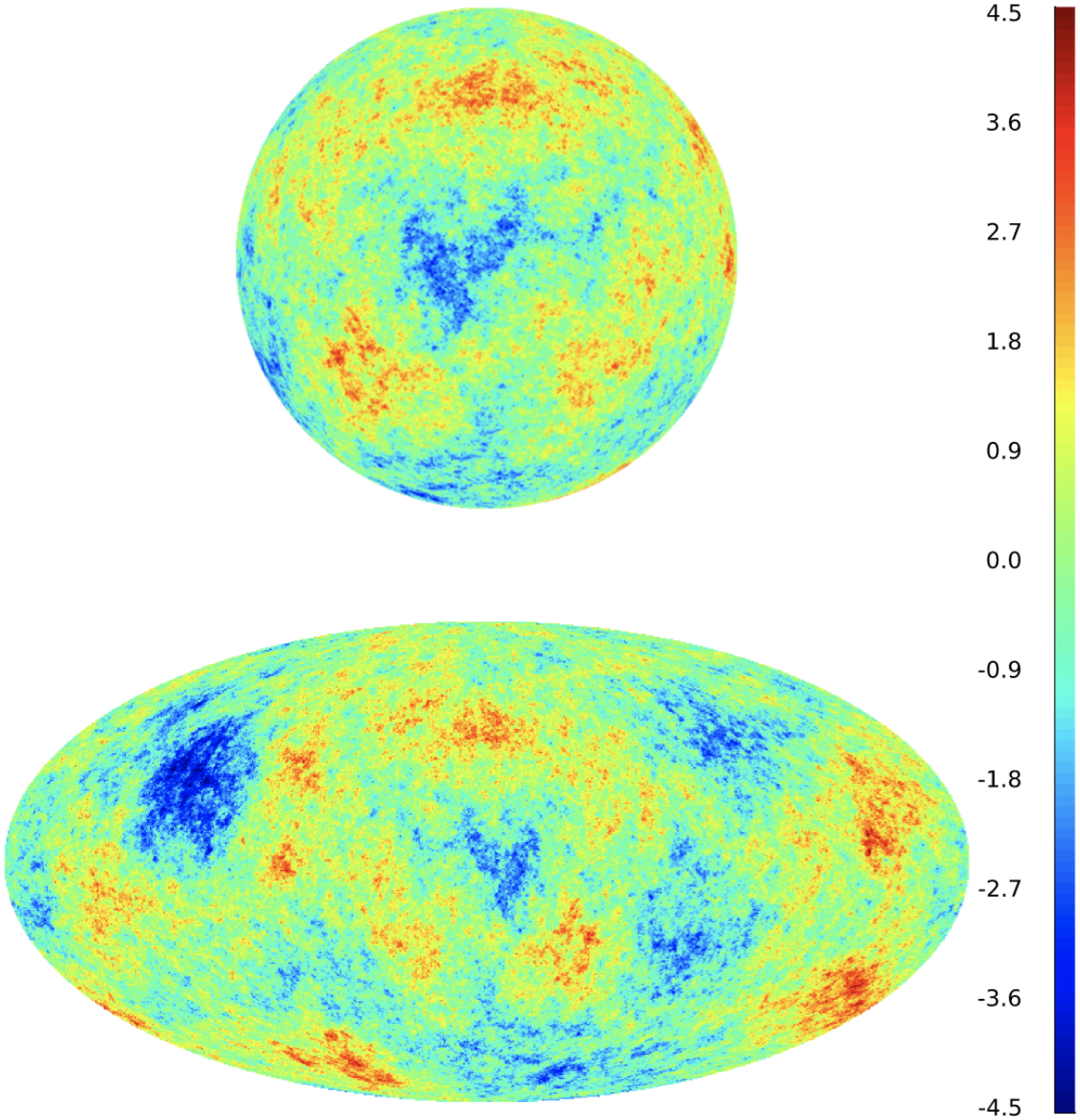
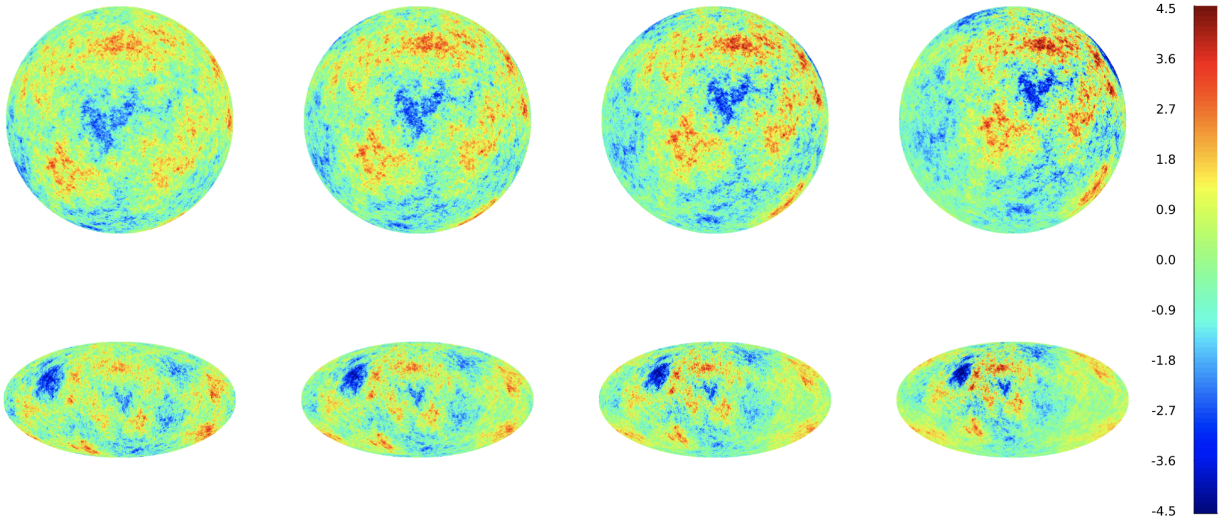


FIGURE 2.3. Illustrations on the Cosmic Microwave Background observed under simulated peculiar velocity effect. From left to right, the peculiar velocity has magnitudes $[0.0, 0.2, 0.4, 0.6]$ in units of lightspeed.The direction of the peculiar velocity is $(x = 1.0, y = 1.0, z = 1.0)$. The leftmost one represents the CMB without any peculiar velocity. From left to right, we can see there is a shrinking toward one spot while enlarging in the opposite on the sphere.



CHAPTER 3

Lorentz Boosts

3.1. Introduction

In this chapter, we explore the mathematical framework that describes the peculiar velocity effect on the CMB, focusing on aspects that are particularly relevant to statistical analysis and modeling.

3.2. Peculiar Velocity Relative to the CMB

The measurement of the observational CMB is influenced by the relativistic Doppler effect and relativistic light aberration, both consequences of the peculiar motion of the Milky Way relative to the CMB rest frame. In the context of the CMB, the term "dipole" refers to the largest anisotropy in the first spherical harmonic ($\ell = 1$), where the amplitude fluctuation is encoded in one cosine function. The observed dipole in the CMB, with an amplitude of around 3.3621 ± 0.0010 mK, is much larger than signals at nearby frequencies, a result specifically attributed to the relativistic Doppler effect. Moreover, relativistic light aberration causes a contraction of the spatial area in the direction of the peculiar velocity and an expansion in the opposite direction. This phenomenon, known as anisotropy, introduces variations in the CMB's temperature across different directions in the sky, appearing slightly warmer in the direction of movement and cooler in the opposite direction. In statistical terms, anisotropy refers to the property of being directionally dependent, as opposed to isotropy, which implies uniformity in all directions.

In this chapter, we will provide a detailed mathematical description of the relativistic Doppler effect and relativistic light aberration. Together, these effects are referred to as the "Lorentz Boost," named for their mathematical equivalence to a (restricted) Lorentz Boost in the Lorentz group. This chapter lays the foundation for understanding how these relativistic effects contribute to our observations of the universe through the lens of the CMB.

3.2.1. Relativistic Light Aberration. Relativistic Light Aberration is a phenomenon that describes the apparent change in the direction of light due to the relative motion between the observer and the source. In this section, we will explore the mathematical framework that describes this effect, focusing on the concepts of peculiar velocity [**ACM+11**] and rapidity [**DC14**].

- **Peculiar Velocity**: Let $\boldsymbol{\beta}$ denote the peculiar velocity, which is expressed in units of the speed of light. It is constrained by $\|\boldsymbol{\beta}\| < 1$, as any relative motion is limited by the speed of light. The magnitude of the peculiar velocity is given by $\beta = \|\boldsymbol{\beta}\|$, where $0 \leq \beta < 1$.
- **Rapidity**: The rapidity $\boldsymbol{\eta}$ is a hyperbolic angle defined as

$$\boldsymbol{\eta} := \frac{\operatorname{arctanh}(\beta)}{\beta}\boldsymbol{\beta}.$$

  Here, $\eta = \|\boldsymbol{\eta}\|$ ranges from 0 to ∞, and it is related to β by $\beta = \tanh\eta$. The Lorentz factor γ, essential in special relativity, is defined as $\gamma = 1/\sqrt{1-\beta^2} = \cosh\eta$.

For the purpose of simplifying the mathematics of peculiar velocity effects, we can imagine aligning our coordinate system such that the z-axis is along the peculiar velocity, if we know its direction. In this setup, $\boldsymbol{\beta}/\|\boldsymbol{\beta}\| = (0,0,1)$. We, the observers, are then moving along the z-axis at a speed β. Let $\mathbf{p}$ be the observed position of a point on the sphere or circle, affected by light aberration as we are moving at speed β. The great circle distance between $\mathbf{p}$ and $\boldsymbol{\beta}/\|\boldsymbol{\beta}\|$, represented as θ as in spherical coordinate system, can be determined using:

$$\cos\theta = \langle \mathbf{p}, \frac{1}{\|\boldsymbol{\beta}\|}\boldsymbol{\beta}\rangle = \langle \mathbf{p}, (0,0,1)\rangle.$$

Since we assume the peculiar velocity is along z-axis, the point $\mathbf{p}$ can be written as $\mathbf{p} = (x,y,z) = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta)$ in 3-dimensional space with spherical coordinate (θ,φ). Let $\mathbf{p}' = (x',y',z') = (\cos\varphi'\sin\theta', \sin\varphi'\sin\theta', \cos\theta')$ be the intrinsic position of the same point, without the effect of light aberration. In other words, θ represents the observation angle when in motion, while θ' represents the intrinsic angle when stationary. The relativistic light aberration leaves φ unchanged, while transforming between θ and θ'. The transformation from $\mathbf{p}'$ to the observed point $\mathbf{p}$ is described by the following equations:

$$\cos\theta = \frac{\cos\theta' + \beta}{1 + \beta\cos\theta'} \quad \text{and} \quad \cos\theta' = \frac{\cos\theta - \beta}{1 - \beta\cos\theta}.$$

These equations capture the essence of the relativistic light aberration effect, mapping the intrinsic position to the observed position. Expressing the transformation in terms of rapidity, we find that the cosine and sine of θ and θ' are related by

$$\begin{aligned}
\cos\theta &= \frac{\cosh\eta\cos\theta' + \sinh\eta}{\cosh\eta + \sinh\eta\cos\theta'} \\
\sin\theta &= \frac{\sin\theta'}{\cosh\eta + \sinh\eta\cos\theta'} \\
\cos\theta' &= \frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta} \\
\sin\theta' &= \frac{\sin\theta}{\cosh\eta - \sinh\eta\cos\theta}.
\end{aligned}$$

In spherical coordinate, the transformation can be written as

$$(\cos\varphi\sin\theta', \sin\varphi\sin\theta', \cos\theta') = \left(\frac{\cos\varphi\sin\theta}{\cosh\eta - \sinh\cos\theta}, \frac{\sin\varphi\sin\theta}{\cosh\eta - \sinh\cos\theta}, \frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\cos\theta}\right)$$

The transformation can also be written as a diffeomorphism $\Phi_\eta : \mathbb{S}^2 \to \mathbb{S}^2$:

$$(\theta', \varphi') = \Phi_\eta(\theta, \varphi) = \left(\arctan2\left(\frac{\sin\theta}{\cosh\eta - \sinh\cos\theta}, \frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\cos\theta}\right), \varphi\right).$$

The function arctan2 is the two-argument arctangent, returning the angle in $[0, 2\pi)$ so that $\theta = \arctan2(\sin\theta, \cos\theta)$ and taking into account the quadrant of $(\sin\theta, \cos\theta)$.

Now we explore the toy model of the circle $\mathbb{S}^1$, represented by coordinates $(\cos\theta, \sin\theta)$ with $\theta \in [-\pi, \pi]$. In this model, we consider a two-dimensional peculiar velocity. To simplified the equations of the transformation, we assume the peculiar velocity is along the x-axis. Nevertheless, in the latter sections, we will state the transformation in arbitary direction.

Given the rapidity η, which is the norm of the rapidity, along the x-axis, we can define a diffeomorphism $\Phi_\eta : \mathbb{S}^1 \to \mathbb{S}^1$ that describes the transformation between the observed angle θ and the intrinsic angle θ'. This transformation is expressed as:

$$\begin{aligned}
\theta' = \Phi_\eta(\theta) &= \arctan\left(\frac{\sin\theta}{\cosh\eta - \sinh\eta\cos\theta}, \frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta}\right), \\
\theta = \Phi_\eta^{-1}(\theta') = \Phi_{-\eta}(\theta') &= \arctan\left(\frac{\sin\theta'}{\cosh\eta + \sinh\eta\cos\theta'}, \frac{\cosh\eta\cos\theta' + \sinh\eta}{\cosh\eta + \sinh\eta\cos\theta'}\right).
\end{aligned}$$

The derivative of this transformation with respect to θ quantifies into how small changes in θ affect θ':

$$\frac{d\theta'}{d\theta} = \frac{d}{d\theta}\Phi_\eta(\theta) = \frac{1}{\cosh\eta - \sinh\eta\cos\theta} = \frac{1}{\gamma(1-\beta\cos\theta)} \quad \text{where } \gamma = \frac{1}{\sqrt{1-\beta^2}}.$$

It also serves as a tool to help obtain the maximum likelihood estimator in later sections. Furthermore, we present a useful identity that facilitates the manipulation of formulas in the transformation:

$$\frac{1}{\gamma(1-\beta\cos\theta)} = \gamma(1+\beta\cos\theta').$$

This identity, connecting the observed and intrinsic angles, is a useful tool for understanding relativistic light aberration in this simplified model. The toy model serves as a stepping stone towards comprehending more advanced scenarios and offers insights into the fundamental geometric and physical principles at play.

3.2.2. Relativistic Doppler effect. This section introduces an attempt to describe the relativistic Doppler effect in mathematical terms, with a particular focus on its intrinsic relationship with relativistic light aberration. Unlike traditional approaches in physics literature, our formulation closely ties the effect to the phenomenon of relativistic light aberration. Specifically, we consider the relativistic light aberration on spheres and circles as a diffeomorphism and define the relativistic Doppler effect as the Jacobian of this diffeomorphism. Further mathematical details will be discussed in subsequent sections.

Let (θ, φ) denote the coordinates in a moving frame, and (θ', φ') denote those in a stationary frame. The transformation of a function $f(\theta, \varphi)$ under the Doppler effect is expressed as:

$$f(\theta,\varphi) \to A(\theta,\varphi)^{\frac{1}{2}} f(\theta,\varphi),$$

where $A(\theta, \varphi)$ is termed the Doppler effect multiplier.

For the sphere $\mathbb{S}^2$, let's assume that the peculiar velocity is $(0, 0, \beta)$ along the z-axis, with rapidity $\eta = \operatorname{atanh}(\beta)$. The Doppler effect multiplier $A_\eta(\theta, \varphi)$ is then defined as:

$$A_\eta(\theta,\varphi) = \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right)^2.$$

For the circle $\mathbb{S}^1$, assume that the peculiar velocity is $(\beta, 0)$ along the x-axis, with the same rapidity $\eta = \text{atanh}(\beta)$. The multiplier $A_\eta(\theta)$ is:

$$A_\eta(\theta) = \frac{1}{\cosh\eta - \sinh\eta\cos\theta}.$$

The Doppler effect multiplier serves as the square root of the ratio between the surface area elements in different coordinate systems. For instance, the surface elements for the sphere $\mathbb{S}^2$ in (θ, φ) and (θ', φ') are connected by:

$$\sin\theta' d\theta' d\varphi' = \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right)^2 \sin\theta d\theta d\varphi$$

This ratio between the surface area elements acts as Doppler effect multiplier for the sphere.

For the circle $\mathbb{S}^1$, the line elements in (θ) and (θ') are related by:

$$d\theta' = \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right) d\theta.$$

3.2.3. Lorentz Boost. Given an arbitrary diffeomorphism Φ and its corresponding Jacobian A, the Lorentz Boost operator, denoted $\mathbf{B}$, is defined as acting on a function f in the space $L^2_{\mathbb{C}}(\mathbb{S}^2)$ as follows:

$$(\mathbf{B}f)(\mathbf{x}) = A(\mathbf{x})^{\frac{1}{2}} f(\Phi(\mathbf{x})),$$

where $\mathbf{x} \in \mathbb{S}^2$.

To illustrate the concept, let's look at its applications to CMB observations:

- Spherical CMB: For a Lorentz Boost along the z-axis with rapidity η, the diffeomorphism Φ_η and its Jacobian A_η are defined as:

$$\Phi_\eta(\theta, \varphi) = \left(\arccos\left(\frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta}\right), \varphi\right)$$
$$A_\eta(\theta, \varphi) = \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right)^2$$

  where (θ, φ) is the spherical coordinate. The observed CMB f^{obs} in terms of the primordial CMB f^{pri} is:

$$f^{\text{obs}}(\theta, \varphi) = \frac{1}{\cosh\eta - \sinh\eta\cos\theta} f^{\text{pri}}\left(\arccos\left(\frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta}\right), \varphi\right)$$

- Circular CMB: For a Lorentz Boost along the x-axis with rapidity η, the diffeomorphism Φ_η and its Jacobian A_η are:

$$\Phi_\eta(\theta) = \arccos\left(\frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta}\right)$$
$$A_\eta(\theta) = \frac{1}{\cosh\eta - \sinh\eta\cos\theta}$$

  where θ is the polar coordinate. The observed CMB f^{obs} in terms of the primordial CMB f^{pri} is:

$$f^{\text{obs}}(\theta) = \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right)^{\frac{1}{2}} f^{\text{pri}}(\arccos\left(\frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta}\right)).$$

The Lorentz Boost operator can be generalized to include a power r:

$$(\mathbf{B}_\eta^{(r)} f)(\mathbf{x}) = A_\eta^{\frac{r}{2}}(\mathbf{x}) f(\Phi_\eta(\mathbf{x})).$$

When $r = 1$, the operator is unitary, which is the central focus of this dissertation. For $r = 0$, the transformation reduces to a mere diffeomorphism.

For a Lorentz Boost in an arbitrary direction $\boldsymbol{\eta} = (\eta_x, \eta_y, \eta_z)$ and $\eta = \|\boldsymbol{\eta}\|$, a rotation R is used to align $\boldsymbol{\eta}$ with the z-axis, so that $R\boldsymbol{\eta} = (0, 0, \eta)$. The Lorentz Boost is then applied to the rotated function $\mathbf{R}^{-1}f$:

$$(\mathbf{R}^{-1}f)(x) = f(R^{-1}x),$$

The function is subsequently rotated back to its original orientation, which yields $\mathbf{B}_{\boldsymbol{\eta}} := \mathbf{R}\mathbf{B}_\eta\mathbf{R}^{-1}$.

3.3. Lorentz Boost of Toy CMB model on Circle

In this section, we extend our discussion to focus on the details of the toy model of the CMB in a circle, as well as rapidity in two-dimensional space. The Lorentz Boost incorporates two primary effects: Relativistic Doppler Effect and Relativistic Light Aberration. Although these effects were initially formulated in a three-dimensional space with x, y, and z axes, they can be restricted to the x and y axes. We consider an Isotropic Gaussian Random Field defined on a circle. First, we present the formulas for the special cases where the peculiar velocity is aligned with the x-axis and

y-axis. We then extend this to derive the general formula for any direction using mathematical methods from the special theory of relativity.

3.3.1. Lorentz Boost on circle to x and y directions. We are going to explicitly write down the Lorentz Boost to x-axis and y-axis respectively. Let $\boldsymbol{\beta} \in \mathbb{R}^2$ represent the peculiar velocity relative to the CMB rest frame, normalized by the speed of light. According to Einstein's theory of relativity, $\|\boldsymbol{\beta}\|_2 < 1$. The rapidity vector $\boldsymbol{\eta}$ is then defined as:

$$\boldsymbol{\eta} = \frac{\operatorname{arctanh}(\|\boldsymbol{\beta}\|_2)}{\|\boldsymbol{\beta}\|_2}\boldsymbol{\beta}.$$

Consider a unit circle $\mathbb{S}^1$ defined as $\mathbb{S}^1 = \{(\cos\theta, \sin\theta) : \theta \in [-\pi, \pi)\}$. Let $f : \mathbb{S}^1 \to \mathbb{R}$ be a primordial CMB temperature map, and $\theta \in [-\pi, \pi]$ serves as the coordinate in motion. The Lorentz Boost acting on f in the x-axis direction (at radian 0) with rapidity $\eta = \|\boldsymbol{\eta}\|_2$ is given by:

$$(\mathbf{B}_\eta f)(\theta) = \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right)^{1/2} f\left(\arctan\left(\frac{\sin\theta}{\cosh\eta\cos\theta - \sinh\eta}\right)\right).$$

On the other hand, the Lorentz Boost acting on f in the y-axis direction (at radian $\frac{\pi}{2}$) is:

$$(\mathbf{B}_\eta f(\theta)) = \left(\frac{1}{\cosh\eta - \sinh\eta\sin\theta}\right)^{1/2} f\left(\arctan\left(\frac{\cosh\eta\sin\theta - \sinh\eta}{\cos\theta}\right)\right).$$

3.3.2. Lorentz Boost in general direction on circle. To derive the Lorentz Boost operator $L^2(\mathbb{S}^1)$ in general direction $\boldsymbol{\eta} \in \mathbb{R}^2$, we start with the Lorentz transformation of a light path to derive the relativistic light aberration. Assume we are in a moving frame relative to the CMB with rapidity $(\eta\cos\phi, \eta\sin\phi)$. The Lorentz transformation from an event (t, x, y) in the CMB rest frame to an event (t', x', y') in our moving frame is denoted by $\mathbf{B}_{\eta,\phi}(t, x, y) = (t', x', y')$, and is given by:

$$\mathbf{B}_{\eta,\phi} = \begin{pmatrix} \cosh\eta & -\sinh\eta\cos\phi & -\sinh\eta\sin\phi \\ -\sinh\eta\cos\phi & \cosh\eta\cos^2\phi + \sin^2\phi & (\cosh\eta - 1)\cos\phi\sin\phi \\ -\sinh\eta\sin\phi & (\cosh\eta - 1)\cos\phi\sin\phi & \cos^2\phi + \cosh\eta\sin^2\phi \end{pmatrix}.$$

Alternatively, the Lorentz transformation can be parameterized by the rapidity vector $\boldsymbol{\eta} = (\eta_x, \eta_y)$ as:

$$\mathbf{B}_{\boldsymbol{\eta}} = \begin{pmatrix} \cosh(\|\boldsymbol{\eta}\|_2) & -\frac{\eta_x}{\|\boldsymbol{\eta}\|_2}\sinh(\|\boldsymbol{\eta}\|_2) & -\frac{\eta_y}{\|\boldsymbol{\eta}\|_2}\sinh(\|\boldsymbol{\eta}\|_2) \\ -\frac{\eta_x}{\|\boldsymbol{\eta}\|_2}\sinh(\|\boldsymbol{\eta}\|_2) & \frac{\eta_x^2}{\|\boldsymbol{\eta}\|_2^2}\cosh(\|\boldsymbol{\eta}\|_2) + \frac{\eta_y^2}{\|\boldsymbol{\eta}\|_2^2} & \frac{\eta_x\eta_y}{\|\boldsymbol{\eta}\|_2^2}(\cosh(\|\boldsymbol{\eta}\|_2) - 1) \\ -\frac{\eta_y}{\|\boldsymbol{\eta}\|_2}\sinh(\|\boldsymbol{\eta}\|_2) & \frac{\eta_x\eta_y}{\|\boldsymbol{\eta}\|_2^2}(\cosh(\|\boldsymbol{\eta}\|_2) - 1) & \frac{\eta_x^2}{\|\boldsymbol{\eta}\|_2^2} + \frac{\eta_y^2}{\|\boldsymbol{\eta}\|_2^2}\cosh(\|\boldsymbol{\eta}\|_2) \end{pmatrix}.$$

Let a light path emanating from the direction $(\cos\theta, \sin\theta)$ be parameterized as $\mathbf{p}(t) = (t, t\cos\theta, t\sin\theta)$ for $t \in \mathbb{R}$. The apparent light path, $\mathbf{p}'(\tau) = (\tau, \tau\cos\theta', \tau\sin\theta')$, under the rapidity $\boldsymbol{\eta} = (\eta\cos\phi, \eta\sin\phi)$, can be expressed as:

$$\mathbf{p}'(\tau) = \mathbf{B}_{\eta,\phi}\mathbf{p}(t)$$

where $\mathbf{B}_{\eta,\phi}$ is

$$\mathbf{B}_{\eta,\phi} := \begin{pmatrix} \cosh\eta & -\sinh\eta\cos\phi & -\sinh\eta\sin\phi \\ -\sinh\eta\cos\phi & \cosh\eta\cos^2\phi + \sin^2\phi & (\cosh\eta - 1)\cos\phi\sin\phi \\ -\sinh\eta\sin\phi & (\cosh\eta - 1)\cos\phi\sin\phi & \cos^2\phi + \cosh\eta\sin^2\phi \end{pmatrix}.$$

In other words, we have the system of equations:

$$\begin{pmatrix} \tau \\ \tau\cos\theta' \\ \tau\sin\theta' \end{pmatrix} = \begin{pmatrix} \cosh\eta & -\sinh\eta\cos\phi & -\sinh\eta\sin\phi \\ -\sinh\eta\cos\phi & \cosh\eta\cos^2\phi + \sin^2\phi & (\cosh\eta - 1)\cos\phi\sin\phi \\ -\sinh\eta\sin\phi & (\cosh\eta - 1)\cos\phi\sin\phi & \cos^2\phi + \cosh\eta\sin^2\phi \end{pmatrix} \begin{pmatrix} t \\ t\cos\theta \\ t\sin\theta \end{pmatrix}.$$

The Lorentz factor $\gamma(\theta)$ is defined as:

$$\gamma(\theta) = \frac{dt}{d\tau} = \frac{1}{\frac{d\tau}{dt}} = \frac{1}{\cosh\eta - \sinh\eta\cos\phi\cos\theta - \sinh\eta\sin\phi\sin\theta}.$$

Indeed, the Jacobian of the transformation $\theta \mapsto \theta'$ equals the Lorentz factor $\gamma(\theta)$. Moreover, the apparent angle θ' can be determined by solving:

$$\begin{aligned} \tan\theta' &= \frac{\tau\sin\theta'}{\tau\cos\theta'} \\ &= \frac{-\sinh\eta\sin\phi + ((\cosh\eta - 1)\cos\phi\sin\phi)\cos\theta + (\cos^2\phi + \cosh\eta\sin^2\phi)\sin\theta}{-\sinh\eta\cos\phi + (\cosh\eta\cos^2\phi + \sin^2\phi)\cos\theta + ((\cosh\eta - 1)\cos\phi\sin\phi)\sin\theta}. \end{aligned}$$

Considering the Lorentz factor γ and appearance angle θ' as functions of θ, the Lorentz Boost operation can be represented as a linear transformation:

$$f(\theta) \mapsto \gamma(\theta)^{\frac{1}{2}} f(\theta'(\theta)).$$

Here, the multiplicative operation $f(\theta) \mapsto \sqrt{\gamma(\theta)} f(\theta)$ corresponds to the relativistic Doppler effect, while the transformation $f(\theta) \mapsto f(\theta'(\theta))$ represents the relativistic light aberration. The composite operation is denoted by $\mathbf{B}_{\boldsymbol{\eta}} : L^2(\mathbb{S}^1) \to L^2(\mathbb{S}^1)$ as:

$$(\mathbf{B}_{\boldsymbol{\eta}} f)(\theta) = \gamma(\theta)^{\frac{1}{2}} f(\theta'(\theta)).$$

3.3.3. Unitarity of the Lorentz Boost Operator. To investigate the unitarity of the Lorentz Boost operator, we examine the interplay between the Doppler effect and light aberration for functions defined on a circle. Let $\Phi_{\eta,\phi}(\theta) = \theta'(\theta)$ denote the pure light aberration, and let $\frac{d}{d\theta}\Phi_{\eta,\phi}(\theta)$ represent its Jacobian. It can be shown that this Jacobian is related to the Doppler effect multiplier $\gamma(\theta)$ as follows:

$$\frac{d}{d\theta}\Phi_{\eta,\phi}(\theta) = \gamma(\theta).$$

Consequently, the area elements in the respective coordinates θ and θ' are related by the Doppler effect multiplier:

$$d\theta' = \gamma(\theta) d\theta.$$

This implies that the Lorentz Boost operation is unitary when the Doppler effect multiplier is applied with an appropriate exponent:

$$\int (f(\theta'))^2 d\theta' = \int \left(\gamma(\theta)^{\frac{1}{2}} f(\theta'(\theta))\right)^2 d\theta.$$

The unitary property is achieved by setting the multiplier to be the square root of the derivative of the angle transformation function. Thus, the Lorentz Boost can be expressed as:

$$f(\theta) \mapsto \left|\frac{d}{d\theta}\Phi_{\eta,\phi}(\theta)\right|^{\frac{1}{2}} f(\Phi_{\eta,\phi}(\theta)).$$

3.4. Lorentz Boost via Flow Formalism

3.4.1. Introduction to Flow. In this section, we introduce the concept of a Flow, a class of diffeomorphisms generated by a vector field. This formalism allows us to express transformed random fields as solutions to ordinary differential equations (ODEs). Consequently, the transformation can be represented by the exponential of a linear operator. The group of diffeomorphisms is parameterized and has a representation as a linear operator acting on a function space. Utilizing the Flow formalism, we can explicitly determine the parametric form of the generator of the matrix exponential, or more generally, the Lie algebra of the Lie group.

Although an explicit expression for the Lorentz Boost operation exists, it is cumbersome for analytical manipulations and approximations. We demonstrate that the Lorentz Boost operation can be cast as a particular type of Flow. This reformulation facilitates easier manipulation and approximation.

We first present and prove several propositions related to the Flow. Subsequently, we show how the Lorentz Boost can be formulated within the Flow framework. This approach aids us in constructing the representation of the group of diffeomorphisms (in this case, the Lorentz Group) acting on function spaces defined on circles or spheres.

Let $\mathcal{M}$ be a manifold equipped with a vector field X. The tangent space at a point p in $\mathcal{M}$ is denoted by $T_p\mathcal{M}$. The symbol $X_p f$ represents the directional derivative of a scalar function f along X at the point p, and it belongs to $T_p\mathcal{M}$. We use Xf to denote the vector field of directional derivatives defined on the entire manifold $\mathcal{M}$.

Definition 3.4.1. A Flow $\Phi_t : \mathcal{M} \to \mathcal{M}$, generated by a vector field X on a manifold $\mathcal{M}$, is a solution of the system

$$\frac{d}{dt}\Phi_t(p) = X(\Phi_t(p)),$$

in the tangent space $T_{\Phi_t(p)}\mathcal{M}$ for all p in $\mathcal{M}$ and initial condition $\Phi_0(p) = p$.

Definition 3.4.2. We define the pushforward of a diffeomorphism. Let $\Phi : \mathcal{M} \to \mathcal{M}$ be a diffeomorphism on a manifold $\mathcal{M}$. The pushforward, denoted by $D\Phi(p) : T_p\mathcal{M} \to T_{\Phi(p)}\mathcal{M}$, is a

linear transformation satisfying

$$(D\Phi(p))\left.\frac{d}{dt}\right|_{t=0}\gamma(t) = \left.\frac{d}{dt}\right|_{t=0}\Phi(\gamma(t))$$

for all smooth curve $\gamma(s)$ on $\mathcal{M}$ with $\gamma(0) = p$.

Note that in the case of Euclidean space $\mathcal{M} = \mathbb{R}^n$, the pushforward of a diffeomorphism is indeed the Jacobian matrix of the diffeomorphism.

LEMMA 3.4.1. Let X be a vector field on a manifold $\mathcal{M}$, and let $\Phi_t(p)$ be a Flow generated by X. The pushforward $D\Phi_t$, satisfies

$$(D\Phi_t(p))X(p) = X(\Phi_t(p))$$

and thus

$$\frac{d}{dt}\Phi_t = (D\Phi_t)X, \quad \Phi_0 = \text{identity map.}$$

PROOF. For any $p \in \mathcal{M}$, let $\gamma(t) := \Phi_t(p)$. Firstly

$$\frac{d}{dt}\Phi_t(p) = (D\Phi_t(p))X(p)$$

by definition 3.4.2. Thus,

$$(D\Phi_t(p))X(p) = \frac{d}{dt}\Phi_t(p) = X(\Phi_t(p))$$

by the system in Definition 3.4.1. Therefore, the Flow equation from Definition 3.4.1 becomes

$$\frac{d}{dt}\Phi_t(p) = (D\Phi_t(p))X(p).$$

□

DEFINITION 3.4.3. Let X be a vector field on a manifold $\mathcal{M}$ with a metric tensor g. The divergence of X, denoted by $\operatorname{div}(X)$, is defined as

$$\operatorname{div}(X) := \frac{1}{\sqrt{|\det g|}}\partial_a\left(\sqrt{|\det g|}X^a\right),$$

where ∂_a denotes the partial derivative with respect to the coordinate x^a.

The definition of divergence for a vector field on a smooth manifold can be found in [**Lee12**].

LEMMA 3.4.2. Let Φ_t be a Flow generated by the vector field X. Then

$$\left.\frac{d}{ds}\right|_{s=0} \det D\Phi_s = \operatorname{div}(X),$$

and

$$\frac{d}{dt} \det D\Phi_t = \operatorname{div}(X) \det D\Phi_t.$$

In summary, the rate of change of the area element $\det D\Phi_t$ is given by the divergence of Φ_t. Those are key equations that we need in proving the main theorem, Theorem 3.4.2, in the Flow formulation of Lorentz Boosts. The term involving divergence in the differential equation in the proposition indeed comes from those equations.

3.4.2. Rotation on circle. To clarify the concept of Integral Flow, let's consider the example of rotation on a circle. We start with a function $f \in L^2(\mathbb{S}^1, \mathbb{R})$ and apply a rotation transformation: $f(\theta) \to f(\theta + \phi)$ for some angle ϕ. This rotation can be represented as an Integral Flow.

We define the diffeomorphism for rotation as $\Phi_\phi(\theta) := \arctan\left(\frac{\sin(\theta+\phi)}{\cos(\theta+\phi)}\right)$. Upon direct differentiation, we find that both $\frac{d}{d\theta}\Phi_\phi(\theta)$ and $\frac{d}{d\phi}\Phi_\phi(\theta)$ are equal to 1. Then the metric $g(\theta)$ in the case of circle is the ratio of $\frac{d}{d\phi}\Phi_\phi(\theta)$ to $\frac{d}{d\theta}\Phi_\phi(\theta)$:

$$g(\theta) = \frac{\frac{d}{d\phi}\Phi_\phi(\theta)}{\frac{d}{d\theta}\Phi_\phi(\theta)} = 1.$$

Given these properties, the corresponding Integral Flow operator $\mathbf{R}$ is defined as $(\mathbf{R}f)(\theta) = \frac{d}{d\theta}f(\theta)$.

Now, let's examine the action of $\mathbf{R}$ on trigonometric functions to understand its nature as a rotation operator:

$$(\mathbf{R}f)(\theta) = \frac{d}{d\theta}f(\theta),$$

$$\mathbf{R}\cos(k\theta) = k\sin(-k\theta),$$

$$\mathbf{R}1 = 0,$$

$$\mathbf{R}\sin(k\theta) = k\cos(-k\theta).$$

The real Fourier decomposition of $\mathbf{R}f$ is then given by:

$$\mathbf{R}f = \begin{pmatrix} & & & & & & \cdots \\ & & & & & 2 & \\ & & & & 1 & & \\ & & & 0 & & & \\ & & -1 & & & & \\ & -2 & & & & & \\ \cdots & & & & & & \end{pmatrix} \begin{pmatrix} \vdots \\ f_{-2} \\ f_{-1} \\ f_0 \\ f_1 \\ f_2 \\ \vdots \end{pmatrix}.$$

Using the definition of the matrix exponential, we can express $\exp(\phi\mathbf{R})$ as:

$$\exp(\phi\mathbf{R}) = \sum_{n=0}^{\infty} \frac{\phi^n \mathbf{R}^n}{n!},$$

which leads to its real Fourier representation:

$$\exp(\phi\mathbf{R}) = \begin{pmatrix} \cdots & & & & & & \cdots \\ & \cos(2\phi) & & & & \sin(2\phi) & \\ & & \cos(\phi) & & \sin(\phi) & & \\ & & & 1 & & & \\ & & -\sin(\phi) & & \cos(\phi) & & \\ & -\sin(2\phi) & & & & \cos(2\phi) & \\ \cdots & & & & & & \cdots \end{pmatrix}.$$

The matrix $\exp(\phi\mathbf{R})$ is a representation of the rotation group SO(2) acting on the vector space of Fourier coefficients of functions on the circle.

In summary, the rotation $f(\theta+\phi)$ can be expressed as $(\exp(\phi\mathbf{R})f)(\theta)$, serving as a straightforward example of an operator acting on functions on a circle. This also shows that the vector field is left-invariant, as $V(\theta) = 1$ for all θ.

3.4.3. Formulation of Light Aberration using Flow. Consider a smooth scalar function $f : \mathcal{M} \to \mathbb{R}$. We introduce a family of functions f_t parameterized by $t \in \mathbb{R}$ defined by $f_t(p) := f(\Phi_t(p))$ all $p \in \mathcal{M}$. Our objective is to demonstrate that f_t satisfies the following differential

equation:

$$\frac{d}{dt} f_t = X f_t, \tag{3.1}$$

where $X f_t$ is the directional derivative of f_t along the vector field X. The directional derivative of a function $g : \mathcal{M} \to \mathbb{R}$ along a vector field X generated by a Flow Φ_t is defined as

$$X g(p) = \lim_{\delta \to 0} \frac{g(\Phi_\delta(p)) - g(p)}{\delta}.$$

By direct calculation, we can confirm that f_t satisfies Equation (3.1). This proof relies on the properties of Flows as 1-parameter groups, specifically $\Phi_{t_1}(\Phi_{t_2}(p)) = \Phi_{t_1+t_2}(p)$. This property comes from the uniqueness of solutions to ordinary differential equations. The detailed steps are as follows:

$$\begin{aligned}
\frac{d}{dt} f_t(p) &= \lim_{\delta \to 0} \frac{f_{t+\delta}(p) - f_t(p)}{\delta} \\
&= \lim_{\delta \to 0} \frac{f(\Phi_{t+\delta}(p)) - f(\Phi_t(p))}{\delta} \\
&= \lim_{\delta \to 0} \frac{f(\Phi_t(\Phi_\delta(p))) - f(\Phi_t(p))}{\delta} \\
&= \lim_{\delta \to 0} \frac{f_t(\Phi_\delta(p)) - f_t(p)}{\delta} \\
&= X f_t(p).
\end{aligned}$$

Later in Theorem 3.4.2, we will define $f_t(p) := (\det D\Phi_t(p))^m f(\Phi_t(p))$ for any positive power m and show that it satisfies an extended version Equation 3.1 that involves an extra term with divergence of the vector field X.

To illustrate the concept of Flow, let's consider an example of rotation in a 2-dimensional plane. The diffeomorphism $\Phi_t(x, y)$ on $\mathbb{R}^2$ is given by

$$\Phi_t(x, y) := ((\cos t)x - (\sin t)y, (\sin t)x + (\cos t)y) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

This transformation represents a rotation around the origin by angle t. The Jacobian matrix $D\Phi_t(x, y)$ is

$$D\Phi_t(x, y) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

On the other hand, differentiating $\Phi_t(x, y)$ with respect to t gives

$$\frac{d}{dt}\Phi_t(x, y) = \begin{pmatrix} -\sin t & -\cos t \\ \cos t & -\sin t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Note that $\frac{d}{dt}\Phi_t(x, y)$ and $D\Phi_t(x, y)$ are different by a rotation matrix. This leads us to define the vector field $X(x, y)$ as a rotation:

$$X(x, y) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (-y, x).$$

We find that $\Phi_t(x, y)$ is a Flow generated by $X(x, y)$, satisfying the equation:

$$\frac{d}{dt}\Phi_t(x, y) = X(x, y)\Phi_t(x, y).$$

For a smooth function $f(x, y)$ in a 2D plane, we define $f_t(x, y) := (\det D\Phi_t(x, y))^r f(\Phi_t(x, y))$. We can directly verify that the determinant of $D\Phi_t(x, y)$ and the divergence of $X(x, y)$ are trivial:

$$\det(D\Phi_t(x, y)) = 1 \text{ and } \operatorname{div} X = 0.$$

So, the family of functions f_t can be explicitly written as:

$$f_t(x, y) = f((\cos t)x - (\sin t)y, (\sin t)x + (\cos t)y).$$

We can then confirm that Equation 3.1 holds through a series of calculations: First, we directly differentiate the function $f_t(x, y)$ w.r.t. t to obtain an expression of LHS

$$\begin{aligned}
\frac{d}{dt}f_t &= \frac{d[\Phi_t(x,y)]_x}{dt}(\frac{\partial}{\partial x}f)(\Phi_t(x,y)) + \frac{d[\Phi_t(x,y)]_y}{dt}(\frac{\partial}{\partial y}f)(\Phi_t(x,y)) \\
&= \frac{d}{dt}(x\cos t - y\sin t)(\frac{\partial}{\partial x}f)(\Phi_t(x,y)) + \frac{d}{dt}(x\sin t + y\cos t)(\frac{\partial}{\partial y}f)(\Phi_t(x,y)) \\
&= (-x\sin t - y\cos t)(\frac{\partial}{\partial x}f)(\Phi_t(x,y)) + (x\cos t - y\sin t)(\frac{\partial}{\partial y}f)(\Phi_t(x,y)) \\
&= [X(\Phi_t(x,y))]_x(\frac{\partial}{\partial x}f)(\Phi_t(x,y)) + [X(\Phi_t(x,y))]_y(\frac{\partial}{\partial y}f)(\Phi_t(x,y)) \\
&= (X(\Phi_t(x,y))) \cdot (\nabla f)(\Phi_t(x,y)) \\
&= (-[\Phi_t(x,y)]_y, [\Phi_t(x,y)]_x) \cdot (\nabla f)(\Phi_t(x,y)).
\end{aligned}$$

On the other hand, we can write down ∇f_t, the gradient of $f_t(x, y)$, in term of ∇f, the gradient of $f(x, y)$.

$$\begin{aligned}
(\nabla f_t)(x,y) &= (D\Phi_t(x,y))^T(\nabla f)(\Phi_t(x,y)) \\
&= \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} (\frac{\partial}{\partial x}f)(\Phi_t(x,y)) \\ (\frac{\partial}{\partial y}f)(\Phi_t(x,y)) \end{pmatrix} \\
\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} (\nabla f_t)(x,y) &= (\nabla f)(\Phi_t(x,y)).
\end{aligned}$$

We now substitute ∇f with the expression involving ∇f_t and obtain

$$\begin{aligned}
\frac{d}{dt}f_t &= (-[\Phi_t(x,y)]_y, [\Phi_t(x,y)]_x) \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} (\nabla f_t)(x,y) \\
&= (-(x\sin t + y\cos t), x\cos t - y\sin t) \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} (\nabla f_t)(x,y) \\
&= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} -\sin t & \cos t \\ -\cos t & -\sin t \end{pmatrix} \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} (\nabla f_t)(x,y) \\
&= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\nabla f_t)(x,y) \\
&= X(x,y) \cdot (\nabla f_t)(x,y) \\
&= Xf_t.
\end{aligned}$$

which matches the equation (3.1).

Now, let $\mathbf{B}$ be the linear operator defined as $\mathbf{B}f := Xf$. Importantly, $\mathbf{B}$ is independent of t. The solution to the differential equation $\frac{d}{dt}f_t = \mathbf{B}f_t$ can be expressed in terms of the exponential of the linear operator:

$$f_t = \exp(\mathbf{B})f$$

where the exponential of the linear operator is defined by the series expansion:

$$\exp(\mathbf{B}) := I + \mathbf{B} + \frac{1}{2!}\mathbf{B}^2 + \frac{1}{3!}\mathbf{B}^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{B}^k.$$

Expressing f_t in terms of the exponential of $\mathbf{B}$ is advantageous, particularly when $\mathbf{B}$ acts sparsely on the basis functions of the function space. This property is demonstrated in the case of Lorentz Boost, where we will show that it can be interpreted as a Flow generated by a vector field that sparsely acts on both Fourier and spherical harmonic bases.

3.4.4. Formulation of Lorentz Boost using Flow in $\mathbb{R}^n$. The Lorentz Boost encompasses both light aberration and the Doppler effect. Light aberration is characterized by the transformation $f(p) \to f(\Phi(p))$, where Φ is some diffeomorphism. Similarly, the Doppler effect is described by

$f(p) \to A(p)f(p)$, where $A(p)$ is a scalar function. The Lorentz Boost operator $\mathbf{B}$ acts on a function f as follows:

$$(\mathbf{B}f)(p) := A(p)^{\frac{r}{2}} f(\Phi(p)).$$

We aim to demonstrate that the family of functions $f_t = A_t(p)^{\frac{1}{2}} f(\Phi_t(p))$ that are governed by the Lorentz Boost operator satisfies the differential equation

$$\begin{aligned} \frac{\partial}{\partial t} f_t &= \frac{1}{2} r \left(\nabla \cdot X\right) f_t + \left(X \cdot \nabla\right) f_t \\ f_0 &= f \end{aligned} \tag{3.2}$$

where $\nabla \cdot X$ and $X \cdot \nabla$ are defined as

$$\nabla \cdot X = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} X_i$$

$$X \cdot \nabla = \sum_{i=1}^{n} X_i \frac{\partial}{\partial x_i}.$$

In this section, we particularly consider the case where the manifold is a Euclidean space $\mathcal{M} = \mathbb{R}^n$. In Euclidean spaces, we can write down specific formulas or equations, making it easier to perform direct calculations and prove theorems.

In the case of Euclidean space $\mathcal{M} = \mathbb{R}^n$, recall that for any diffeomorphism Φ the Jacobian matrix $D\Phi(x)$ at point x in $\mathbb{R}^n$ has an explicit matrix form:

$$D\Phi = \begin{pmatrix} \vdots & \cdots & \vdots \\ \frac{\partial \Phi}{\partial x_1} & \cdots & \frac{\partial \Phi}{\partial x_n} \\ \vdots & \cdots & \vdots \end{pmatrix}.$$

Indeed, Jacobian matrix $D\Phi(x)$ is the pushforward of Φ. Pushforward is the generalization of Jacobian matrix to manifold, and thus Lemma 3.4.1 applies.

THEOREM 3.4.1. Let $\Phi_t : \mathbb{R}^n \to \mathbb{R}^n$ be a Flow generated by the vector field X. Let $r \geq 0$. For any smooth function $f : \mathbb{R}^n \to \mathbb{R}$ (or $\mathbb{C}$) let

$$f_t(x) := (\det D\Phi_t(x))^{\frac{r}{2}} f(\Phi_t(x)).$$

Then f_t satisfies the differential equation (3.2).

PROOF. We start by invoking Lemma 3.4.1, which tells us that the Flow Φ_t follows the equation

$$\frac{d}{dt}\Phi_t(x) = (D\Phi_t)X(x)$$

where $D\Phi_t(x)$ represents the Jacobian matrix of the diffeomorphism Φ_t at a point $x \in \mathbb{R}^n$. From this equation, it follows that X can be expressed as

$$X = (D\Phi_t)^{-1}\frac{d}{dt}\Phi_t.$$

for the t as long as the equation is valid.

In this proof, our aim is to establish that f_t satisfies the differential equation (3.2). We note that $f_0 = f$ as Φ_0 is the identity map. By computing the derivative of $f_t(x) = (\det D\Phi_t(x))^{r/2} f(\Phi_t(x))$ with respect to t, we obtain:

$$\begin{aligned}\frac{d}{dt}f_t &= (r/2)(\det D\Phi_t)^{r/2}\mathrm{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right) f\circ\Phi_t + (\det D\Phi_t)^{r/2}\frac{d}{dt}\Phi_t\cdot((\nabla f)\circ\Phi_t)\\ &= (r/2)\mathrm{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right) f_t + (\det D\Phi_t)^{r/2}\frac{d}{dt}\Phi_t\cdot((\nabla f)\circ\Phi_t)\,.\end{aligned}$$

Similarly, by directly calculating the gradient of $f_t(x) = (\det D\Phi_t(x))^{r/2} f(\Phi_t(x))$, we find:

$$\begin{aligned}\nabla f_t &= (r/2)(\det D\Phi_t)^{r/2}\begin{pmatrix}\mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t\right)\\ \vdots\\ \mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t\right)\end{pmatrix} f\circ\Phi_t + (\det D\Phi_t)^{r/2}(D\Phi_t)^T((\nabla f)\circ\Phi_t)\\ &= (r/2)\begin{pmatrix}\mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t\right)\\ \vdots\\ \mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t\right)\end{pmatrix} f_t + (\det D\Phi_t)^{r/2}(D\Phi_t)^T((\nabla f)\circ\Phi_t)\end{aligned}$$

Rearranging the terms and applying the inverse of $(D\Phi_t)^T$, we have

$$(\det D\Phi_t)^r(\nabla f)\circ\Phi_t = ((D\Phi_t)^T)^{-1}\left(\nabla f_t - (r/2)\begin{pmatrix}\mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t\right)\\ \vdots\\ \mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t\right)\end{pmatrix} f_t\right).$$

Combining the above equations, we have

$$\begin{aligned}
\frac{d}{dt}f_t &= (r/2)\mathrm{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right)f_t + (\det D\Phi_t)^{r/2}\frac{d}{dt}\Phi_t\cdot((\nabla f)\circ\Phi_t)\\
&= (r/2)\mathrm{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right)f_t + \frac{d}{dt}\Phi_t\cdot\left(((D\Phi_t)^T)^{-1}(\nabla f_t - (r/2)\begin{pmatrix}\mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t\right)\\ \vdots \\ \mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t\right)\end{pmatrix}f_t)\right)\\
&= (r/2)\mathrm{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right)f_t + (D\Phi_t)^{-1}\frac{d}{dt}\Phi_t\cdot\left((\nabla f_t - (r/2)\begin{pmatrix}\mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t\right)\\ \vdots \\ \mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t\right)\end{pmatrix}f_t)\right)\\
&= (r/2)\left(\mathrm{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right) - (D\Phi_t)^{-1}\frac{d}{dt}\Phi_t\cdot\begin{pmatrix}\mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t\right)\\ \vdots \\ \mathrm{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t\right)\end{pmatrix}\right)f_t + (D\Phi_t)^{-1}\frac{d}{dt}\Phi_t\cdot\nabla f_t.
\end{aligned}$$

The main point now is to show that the first term in the expression for $\frac{d}{dt}f_t$ can be simplified to $r\nabla\cdot\left((D\Phi_t)^{-1}\frac{d}{dt}\Phi_t\right)f_t$. To do this, we use Lemma 3.4.3 which gives us:

$$\nabla\cdot(Av) = (\nabla\cdot A)v + \mathrm{tr}(ADv).$$

Setting $A = (D\Phi_t)^{-1}$ and $v = \frac{\partial}{\partial x_i}\Phi_t$, we deduce:

$$\begin{aligned}
(\nabla\cdot(D\Phi_t)^{-1})\frac{\partial}{\partial x_i}\Phi_t &= \nabla\cdot((D\Phi_t)^{-1}\frac{\partial}{\partial x_i}\Phi_t) - \mathrm{tr}((D\Phi_t)^{-1}D\frac{\partial}{\partial x_i}\Phi_t)\\
&= -\mathrm{tr}((D\Phi_t)^{-1}\frac{\partial}{\partial x_i}D\Phi_t)\\
(\nabla\cdot(D\Phi_t)^{-1})D\Phi_t &= -\begin{pmatrix}\mathrm{tr}((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t)\\ \vdots \\ \mathrm{tr}((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t)\end{pmatrix}^T.
\end{aligned}$$

Hence, the first term of the expression of $\frac{d}{dt}f_t$

$$\begin{aligned}
&\left(\operatorname{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right) - (D\Phi_t)^{-1}\frac{d}{dt}\Phi_t \cdot \begin{pmatrix} \operatorname{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_1}D\Phi_t\right) \\ \vdots \\ \operatorname{tr}\left((D\Phi_t)^{-1}\frac{\partial}{\partial x_n}D\Phi_t\right) \end{pmatrix}\right) \\
&= \left(\operatorname{tr}\left((D\Phi_t)^{-1}\frac{d}{dt}D\Phi_t\right) + (D\Phi_t)^{-1}\frac{d}{dt}\Phi_t \cdot ((\nabla\cdot(D\Phi_t)^{-1})D\Phi_t)^T\right) \\
&= \left(\operatorname{tr}\left((D\Phi_t)^{-1}D\frac{d}{dt}\Phi_t\right) + (\nabla\cdot(D\Phi_t)^{-1})\frac{d}{dt}\Phi_t\right) \\
&= \nabla\cdot\left((D\Phi_t)^{-1}\frac{d}{dt}\Phi_t\right).
\end{aligned}$$

Substituting this result back into the expression for $\frac{d}{dt}f_t$, we establish:

$$\frac{d}{dt}f_t = \frac{r}{2}\left(\nabla\cdot X\right)f_t + \left(X\cdot\nabla\right)f_t.$$

This concludes our proof. □

LEMMA 3.4.3. Let $\mathbf{A}$ be a matrix and $\mathbf{v}$ be a vector. Then the divergence of the vector $\mathbf{Av}$ is given by:

$$\nabla\cdot(\mathbf{Av}) = (\nabla\cdot\mathbf{A})\mathbf{v} + \operatorname{tr}(\mathbf{A}D\mathbf{v})$$

where $D\mathbf{v}$ is the Jacobian matrix of $\mathbf{v}$ with respect to the coordinates x_1, x_2, x_3, and $\nabla\cdot\mathbf{A}$ is the divergence of each column of $\mathbf{A}$, explicitly,

$$D\mathbf{v} = \begin{pmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} & \frac{\partial v_1}{\partial x_3} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} & \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_3}{\partial x_1} & \frac{\partial v_3}{\partial x_2} & \frac{\partial v_3}{\partial x_3} \end{pmatrix}$$

and

$$\nabla\cdot\mathbf{A} = \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \end{bmatrix}\mathbf{A} = \begin{pmatrix} \frac{\partial A_{11}}{\partial x_1} + \frac{\partial A_{21}}{\partial x_2} + \frac{\partial A_{31}}{\partial x_3} \\ \frac{\partial A_{12}}{\partial x_1} + \frac{\partial A_{22}}{\partial x_2} + \frac{\partial A_{32}}{\partial x_3} \\ \frac{\partial A_{13}}{\partial x_1} + \frac{\partial A_{23}}{\partial x_2} + \frac{\partial A_{33}}{\partial x_3} \end{pmatrix}^T.$$

PROOF. We begin by recalling the formula that links the divergence of a vector field $\mathbf{v}$ to its Jacobian matrix:

$$\nabla \cdot \mathbf{v} = \operatorname{tr}(D\mathbf{v}).$$

To prove the lemma, we first calculate $D(\mathbf{Av})$ by applying the product rule for differentiation:

$$D(\mathbf{Av}) = \left[(\frac{\partial}{\partial x_1}\mathbf{A})\mathbf{v} \quad (\frac{\partial}{\partial x_2}\mathbf{A})\mathbf{v} \quad (\frac{\partial}{\partial x_3}\mathbf{A})\mathbf{v} \right] + \mathbf{A}D(\mathbf{v})$$

Next, we take the trace of both terms. After simplifying, the trace of the first term is $(\nabla \cdot \mathbf{A})\mathbf{v}$.

This concludes the proof of Lemma 3.4.3. □

It's important to note that if $X = (D\Phi_t)^{-1}\frac{d}{dt}\Phi_t$ is independent of t, then the corresponding differential equation becomes linear. This is a crucial part of the proof for establishing Theorem 3.4.1. In this situation, we can define an operator $\mathbf{K} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ as follows:

$$\mathbf{K}f = \frac{r}{2}(\nabla \cdot X)\, f + (X \cdot \nabla)\, f.$$

With $\mathbf{K}$ thus defined, the solution to the differential equation is given by:

$$f_t = \exp(t\mathbf{K}) f_0$$

where the exponential of an operator is defined as the power series:

$$\exp(t\mathbf{K}) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbf{K}^k.$$

3.4.5. Formulation of Lorentz Boost using Flow in manifold. The theorem and proof we discussed earlier is limited to Euclidean spaces, denoted by $\mathbb{R}^n$. The proof employs straightforward multivariate calculus techniques. However, the result is not applicable to manifolds like the sphere that deviate from the Euclidean structure. In order to generalize this, one can introduce a term involving the divergence of the Flow vector field.

3.4.5.1. *Notation and Preliminaries.* To formalize the general theorem of the Lorentz Boost with Flow, let $\mathcal{M}$ be a manifold and X be a vector field on $\mathcal{M}$. The divergence of X on $\mathcal{M}$ is written as $\operatorname{div} X$. For any function f on $\mathcal{M}$, its directional derivative along X is Xf. In the realm of differential geometry, the concept of a "pullback" is very useful.

DEFINITION 3.4.4. Let Φ be a diffeomorphism on $\mathcal{M}$. Its pullback, denoted by Φ^*, acts on any function f on $\mathcal{M}$ as $(\Phi^* f)(x) = f(\Phi(x))$. Additionally, if $\omega = dx^1 \wedge \ldots \wedge dx^n$ is the volume form on $\mathcal{M}$, its pullback $\Phi^*(\omega)$ satisfies

$$\int_\Omega \Phi^*(\omega)\omega = \int_{\Phi(\Omega)} \omega, \quad \text{for } \Omega \subset \mathcal{M}.$$

For a manifold $\mathcal{M}$ equipped with the metric h, the pullback of h is $\Phi^* h$ and is defined as $\Phi^* h(X, Y) = h((D\Phi)X, (D\Phi)Y)$ for vector fields X and Y.

3.4.5.2. *Lorentz Boost Operator with Flow.* The concept of light aberration is thought of as a family of diffeomorphisms on the manifold, each generated by a Flow. The Doppler effect is understood as a ratio of the change in area under this family of diffeomorphisms. Given such a Flow, one can then define the Lorentz Boost operator corresponding to it.

DEFINITION 3.4.5. Suppose Φ_t is a Flow generated by a vector field X on a manifold $\mathcal{M}$. The Lorentz Boost operator $\mathbf{B}_t$ corresponding to the Flow Φ_t is defined as

$$(\mathbf{B}_t f)(p) := (A_t(p))^{r/2} f(\Phi_t(p)),$$

where r is a positive number and A_t is the ratio between the area element of the pullback of the Flow and the original area element, defined as

$$A_t := \frac{\Phi_t^*(\omega)}{\omega}.$$

Furthermore, A_t can be interpreted in different contexts. For example, if the Flow preserves area, such as in a rotation, then $A_t = 1$. If the Flow is conformal, A_t becomes the square of the conformal factor. A_t can also be understood as the Radon–Nikodym derivative, satisfying:

$$\int_\Omega A_t d\mu = \int_{\Phi_t(\Omega)} d\mu,$$

where μ is a measure on the manifold. In differential geometry terms, if $\mathcal{M}$ is equipped with a metric h, A_t is:

$$A_t = \sqrt{\frac{\det \Phi^* h}{\det h}}.$$

3.4.5.3. *Generalized Lorentz Boost Theorem.*

THEOREM 3.4.2. Consider a manifold $\mathcal{M}$ with a volume form $\omega = dx^1 \wedge \ldots \wedge dx^n$. Let $\Phi_t : \mathcal{M} \to \mathcal{M}$ be a Flow on $\mathcal{M}$ generated by a vector field X on $\mathcal{M}$. Assume $\Phi_0(x) = x$ and

$$\frac{d}{dt}\Phi_t(x) = X(\Phi_t(x)).$$

For m being a positive number and any $f \in L^2(\mathcal{M})$, define a family of functions f_t as

$$f_t := \left(\frac{\Phi_t^*(\omega)}{\omega}\right)^m \Phi_t^* f.$$

Then, for all t, f_t satisfies

$$\frac{\partial}{\partial t} f_t = m(\operatorname{div} X) f_t + X f_t,$$

with the initial condition $f_0 = f$.

The case of $m = 0$ was shown earlier through direct calculation. Theorem 3.4.2 extends this by introducing a multiplier power m, motivated by the Lorentz Boost operation that commonly involves multipliers of varying powers for different physical observables. In the case where the Lorentz Boost is an unitary operator, $m = \frac{1}{2}$.

PROOF. Fix a point $p \in \mathcal{M}$. We examine the action of the vector field X on f_t:

$$\begin{aligned}
X_p[f_t] &= X_p\left[\frac{\Phi_t^*(\omega)^m \Phi_t^* f^m}{\omega}\right] \\
&= \frac{d}{ds}\bigg|_{s=0} \Phi_s^*\left[\frac{\Phi_t^*(\omega)^m \Phi_t^* f}{\omega^m}\right] \\
&= \frac{d}{ds}\bigg|_{s=0}\left[\frac{\Phi_{s+t}^*(\omega)^m \Phi_{s+t}^* f}{\Phi_s^*(\omega)^m}\right] \\
&= \frac{\frac{d}{dt}[\Phi_t^*(\omega)^m \Phi_t^* f]}{\Phi_0^*(\omega)^m} - m\left[\frac{\Phi_t^*(\omega)^m \Phi_t^* f}{\Phi_0^*(\omega)^m}\right]\frac{\frac{d}{dt}\big|_{s=0}\Phi_s^*(\omega)}{\Phi_0^*(\omega)} \\
&= \frac{d}{dt} f_t(p) - m f_t(p)\frac{\operatorname{div} X|_p \omega}{\omega} \\
&= \frac{d}{dt} f_t(p) - m f_t(p) \operatorname{div} X|_p
\end{aligned}$$

Here we used the fact that $\frac{d}{ds}|_{s=0}\Phi_s^*\omega = L_X\omega = \operatorname{div}(X)\omega$. Thus, we obtain:

$$\frac{d}{dt} f_t(p) = m f_t(p) \operatorname{div} X|_p + X_p[f_t].$$

This confirms the differential equation in the theorem:

$$\frac{\partial}{\partial t} f_t = m(\operatorname{div} X) f_t + X f_t.$$

with the initial condition $f_0 = f$.

□

3.4.6. Lorentz Boost on sphere.

3.4.6.1. *General Formulation.* To study the Lorentz Boost effect on a unit sphere $\mathbb{S}^2$, we begin by setting up the required notations. Let $\mathbb{S}^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$. Points on $\mathbb{S}^2$ can be represented using spherical coordinates (θ, φ) as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\varphi \sin\theta \\ \sin\varphi \sin\theta \\ \cos\theta \end{pmatrix}.$$

We also need to describe the tangent vector field on the sphere. A tangent vector at a point $\mathbf{x} \in \mathbb{S}^2$ is orthogonal to $\mathbf{x}$, meaning the $\mathbf{v} \cdot \mathbf{x} = 0$.

Basis vectors $\mathbf{e}_\theta$ and $\mathbf{e}_\varphi$ in spherical coordinates are:

$$\mathbf{e}_\theta = \begin{pmatrix} \cos\varphi \cos\theta \\ \sin\varphi \cos\theta \\ -\sin\theta \end{pmatrix}, \quad \mathbf{e}_\varphi = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix}.$$

These basis vectors form an orthonormal frame at each point (θ, φ). Any tangent vector field X can be uniquely represented as:

$$X(\theta, \varphi) = X_1(\theta, \varphi)\mathbf{e}_\theta + X_2(\theta, \varphi)\mathbf{e}_\varphi.$$

Next, we consider divergence, gradient and derivative operations on the sphere. These are different from their counterparts in $\mathbb{R}^n$. The divergence of vector field X on $\mathbb{S}^2$ is:

$$\operatorname{div} X = \frac{1}{\sin\theta} \frac{\partial}{\partial\theta}(\sin\theta X_1) + \frac{1}{\sin\theta} \frac{\partial X_2}{\partial\varphi}.$$

The gradient on $\mathbb{S}^2$ of a scalar field f is:

$$\nabla_{\mathbb{S}^2} f = \mathbf{e}_\theta \frac{\partial f}{\partial \theta} + \mathbf{e}_\varphi \frac{1}{\sin\theta} \frac{\partial f}{\partial \varphi}$$

and the derivative of f along X is:

$$X[f] = X_1 \frac{\partial f}{\partial \theta} + X_2 \frac{1}{\sin\theta} \frac{\partial f}{\partial \varphi}.$$

It is essential to understand that the divergence operation on $\mathbb{S}^2$ serves as a measure of the rate of change of volume associated with the Flow generated by the tangent vector field X. Importantly, this notion of divergence is distinct from the Cartesian counterpart often used in $\mathbb{R}^2$ or $\mathbb{R}^3$. This distinction arises because we are working in spherical coordinates, a curvilinear system, where the volume element is not constant but depends on the specific coordinates (θ, φ). Therefore, the divergence operation cannot be simplified to $\sum_i \frac{\partial}{\partial x_i} X_i$ like it is in Cartesian coordinates.

For the Lorentz Boost, we consider a boost in the prescribed direction to the North pole or the z-axis. The rapidity η ranges over $\mathbb{R}$, and the corresponding speed is $\beta = \tanh\eta$ with range $(-1, 1)$. The Lorentz Boost effect is a combination of light aberration and the Doppler effect. The light aberration is given by a family of diffeomorphism $\Phi_\eta : \mathbb{S}^2 \to \mathbb{S}^2$. If we move at speed $\beta = \tanh(\eta)$ and observe a star at position x, then the actual position of the star in the sky when we are stationary is $\Phi_\eta(x)$. This map is generated by a left-invariant vector field X, such that:

$$\frac{d}{d\eta}\Phi_\eta(x) = X(\Phi_\eta(x)) \quad \text{and} \quad \frac{d}{d\eta}\Phi_\eta(x) = D\Phi_\eta X(x).$$

Here, $D\Phi_\eta$ is the pushforward of Φ_η. The explicit form of X that induces Φ_η under a Lorentz Boost can be obtained by solving the above differential equations. The solutions allow us to understand how the positions of celestial objects shift as we approach the speed of light.

The Jacobian matrix $D\Phi_\eta(\theta, \varphi)$ is given by:

$$D\Phi_\eta(\theta, \varphi) = \begin{pmatrix} \frac{1}{\cosh\eta - \sinh\eta\cos\theta} & 0 \\ 0 & 1 \end{pmatrix}.$$

Consider the vector field $X(\theta, \varphi)$, defined as:

$$X(\theta, \varphi) = \sin\theta \mathbf{e}_\theta + 0\mathbf{e}_\varphi.$$

This vector field generates the Flow Φ_η. It can be verified that $X(\theta, \varphi)$ satisfies the differential equation in Equation 3.4.1.

Now, let $f : \mathbb{S}^2 \to \mathbb{R}$ represent a smooth scalar field on the sphere. Suppose f is the field we aim to measure in the primordial sky. However, the Lorentz Boost along the z-axis with rapidity η alters the observed scalar field, denoted as f_η. This observed field is expressed in terms of f as:

$$f_\eta(x) = A_\eta(x)^{\frac{r}{2}} f(\Phi_\eta(x)),$$

where r is a non-negative integer, and A_η accounts for the Doppler effect. To find the Doppler effect multiplier A_η, we begin by expressing the metric tensor g of the sphere $\mathbb{S}^2$:

$$g = \begin{pmatrix} g_{\theta\theta} & g_{\theta\varphi} \\ g_{\varphi\theta} & g_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}$$

We can also write down the pullback $\Phi_\eta^* g$:

$$\Phi_\eta^* g = \begin{pmatrix} (\Phi_\eta^* g)_{\theta\theta} & (\Phi_\eta^* g)_{\theta\varphi} \\ (\Phi_\eta^* g)_{\varphi\theta} & (\Phi_\eta^* g)_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right)^2 & 0 \\ 0 & \left(\frac{\sin\theta}{\cosh\eta - \sinh\eta\cos\theta}\right)^2 \end{pmatrix}.$$

Hence, the Doppler effect multiplier A_η is given by:

$$A_\eta(\theta, \varphi) = \frac{\sqrt{\det \Phi\eta^* g}}{\sqrt{\det g}} = \left(\frac{1}{\cosh\eta - \sinh\eta\cos\theta}\right)^2.$$

We are now ready to determine the Lorentz Boosts along different axes. We first compute the divergence of $X = \sin\theta\mathbf{e}_\theta + 0\mathbf{e}_\varphi$ corresponding to Lorentz Boost toward the z-axis and directional derivative along X:

$$\begin{aligned} X &= \sin\theta\mathbf{e}_\theta \\ \operatorname{div} X &= \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin^2\theta) = 2\cos\theta \\ Xf &= \sin\theta\frac{\partial}{\partial\theta}f \end{aligned}$$

Upon inserting these values into the differential equation, we arrive at:

$$\frac{d}{d\eta} f_\eta(\theta, \varphi) = \frac{1}{2} r \cos\theta f_\eta(\theta, \varphi) + \sin\theta \frac{d}{d\theta} f_\eta(\theta, \varphi).$$

Likewise, we can find the Lorentz Boost in the direction of the x-axis as follows:

$$\begin{aligned} X &= \cos\theta\cos\varphi \mathbf{e}_\theta + \sin\varphi \mathbf{e}_\varphi \\ \operatorname{div} X &= \frac{1}{\sin\theta}\left(-\frac{\partial}{\partial\theta}(\sin\theta\cos\theta\cos\varphi) + \frac{\partial}{\partial\varphi}(\sin\varphi)\right) = 2\sin\theta\cos\varphi \\ Xf &= -\cos\theta\cos\varphi\frac{\partial}{\partial\theta} f + \frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial\varphi} f. \end{aligned}$$

Similarly, the Lorentz Boost in the y-axis direction can be determined as:

$$\begin{aligned} X &= \cos\theta\sin\varphi \mathbf{e}_\theta + \cos\varphi \mathbf{e}_\varphi \\ \operatorname{div} X &= \frac{1}{\sin\theta}\left(-\frac{\partial}{\partial\theta}(\sin\theta\cos\theta\sin\varphi) - \frac{\partial}{\partial\varphi}(\cos\varphi)\right) = 2\sin\theta\sin\varphi \\ Xf &= -\cos\theta\sin\varphi\frac{\partial}{\partial\theta} f - \frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial\varphi} f. \end{aligned}$$

For Lorentz Boosts along different axes, we can summarize the respective operators $Kf := \frac{1}{2}(\operatorname{div} X)f + Xf$ as:

$$\begin{aligned} K_x &= 2\sin\theta\cos\varphi - \cos\theta\cos\varphi\frac{\partial}{\partial\theta} + \frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial\varphi} \\ K_y &= 2\sin\theta\sin\varphi - \cos\theta\sin\varphi\frac{\partial}{\partial\theta} - \frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial\varphi} \\ K_z &= 2\cos\theta + \sin\theta\frac{\partial}{\partial\theta}. \end{aligned}$$

3.4.6.2. *Representation in Spherical Harmonics.* In the study of Lorentz transformations, it is crucial to understand how these operators act on functions defined on a sphere. Here, we explore this action in the context of spherical harmonics. The Lorentz Boost Lie algebra linear operators, denoted as $\mathbf{K} = (K_x, K_y, K_z)$, can be explicitly defined in terms of spherical harmonics. Consider f as a real-valued function defined on a sphere $\mathbb{S}^2$, which can be decomposed into spherical harmonics as follows:

$$f = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell m} Y_{\ell m}$$

Here, $Y_{\ell m}$ are the spherical harmonics, which form an orthonormal basis for square-integrable functions on S^2, and $f_{\ell m}$ are the corresponding expansion coefficients. These coefficients satisfy the condition $f_{\ell,-m} = (-1)^m (f_{\ell m})^*$ to ensure that f is real-valued.

The Lorentz Boost Lie algebra is spanned by the linear operators $\mathbf{K} = (K_x, K_y, K_z),$, which can be explicitly defined in the spherical harmonic basis. The action of K_x, K_y, and K_z on f can be expressed as:

$$
\begin{aligned}
(K_x f)_{\ell m} &= -\frac{1}{2}\sqrt{\frac{(\ell+m)\ell^2(\ell+m-1)}{4\ell^2-1}} f_{\ell-1,m-1} \\
&\quad + \frac{1}{2}\sqrt{\frac{(\ell-m)\ell^2(\ell-m-1)}{4\ell^2-1}} f_{\ell-1,m+1} \\
&\quad - \frac{1}{2}\sqrt{\frac{(\ell-m+1)(\ell+1)^2(\ell-m+2)}{4(\ell+1)^2-1}} f_{\ell+1,m-1} \\
&\quad + \frac{1}{2}\sqrt{\frac{(\ell+m+1)(\ell+1)^2(\ell+m+2)}{4(\ell+1)^2-1}} f_{\ell+1,m+1} \\
(K_y f)_{\ell m} &= \frac{i}{2}\sqrt{\frac{(\ell+m)\ell^2(\ell+m-1)}{4\ell^2-1}} f_{\ell-1,m-1} \\
&\quad + \frac{i}{2}\sqrt{\frac{(\ell-m)\ell^2(\ell-m-1)}{4\ell^2-1}} f_{\ell-1,m+1} \\
&\quad + \frac{i}{2}\sqrt{\frac{(\ell-m+1)(\ell+1)^2(\ell-m+2)}{4(\ell+1)^2-1}} f_{\ell+1,m-1} \\
&\quad + \frac{i}{2}\sqrt{\frac{(\ell+m+1)(\ell+1)^2(\ell+m+2)}{4(\ell+1)^2-1}} f_{\ell+1,m+1} \\
(K_z f)_{\ell m} &= \sqrt{\frac{(\ell+m)(\ell-m)\ell^2}{4\ell^2-1}} f_{\ell-1,m} \\
&\quad - \sqrt{\frac{(\ell+m+1)(\ell-m+1)(\ell+1)^2}{4(\ell+1)^2-1}} f_{\ell+1,m}
\end{aligned}
\tag{3.3}
$$

where i is the imaginary unit.

To understand how the Lorentz Boost operators act on f, consider the following decomposition:

$$f = \sum_{\ell} \sum_{m} f_{\ell m} Y_{\ell m}$$

Then, $Kf = \sum_\ell \sum_m f_{\ell m} K Y_{\ell m}$. Our objective is to express Kf in terms of the original basis $Y_{\ell m}$. First, we note:

$$Kf = \sum_\ell \sum_m (Kf)_{\ell m} Y_{\ell m}$$

Further calculations yield:

$$\begin{aligned} Kf &= \sum_\ell \sum_{m'} f_{\ell' m'} K Y_{\ell' m'} \\ &= \sum_{\ell'} \sum_{m'} f_{\ell' m'} \sum_\ell \sum_m (K Y_{\ell' m'})_{\ell m} Y_{\ell m} \\ &= \sum_\ell \sum_m \left(\sum_{\ell'} \sum_{m'} f_{\ell' m'} (K Y_{\ell' m'})_{\ell m} \right) Y_{\ell m} \end{aligned}$$

Finally, as $\{Y_{\ell m}\}$ forms a basis, we obtain:

$$(Kf)_{\ell m} = \sum_{\ell'} \sum_{m'} f_{\ell' m'} (K Y_{\ell' m'})_{\ell m}.$$

To derive the explicit form of (K_x, K_y, K_z) in the spherical harmonic basis, we can perform direct calculation. We start with K_z first since it is the simplest one. To derive the explicit form of K_z in the spherical harmonic basis, we start with Equation (24) from [**DC14**]:

$$K_z Y_{\ell m} = B_{\ell+1,m} Y_{\ell+1,m} - B_{\ell,m} Y_{\ell-1,m}$$

where $B_{\ell m} = \sqrt{\frac{(\ell^2 - m^2)\ell^2}{4\ell^2 - 1}}$. Then we can write

$$(K_z Y_{\ell' m'})_{\ell m} = \begin{cases} B_{\ell,m} & \text{if } (\ell, m) = (\ell' + 1, m') \\ -B_{\ell+1,m} & \text{if } (\ell, m) = (\ell' - 1, m') \\ 0 & \text{otherwise.} \end{cases}$$

By substituting this into our general expression for Kf, we obtain:

$$\begin{aligned}
(K_z f)_{\ell m} &= \sum_{\ell'} \sum_{m'} f_{\ell' m'} (K_z Y_{\ell' m'})_{\ell m} \\
&= \sum_{\ell'} f_{\ell' m} (K_z Y_{\ell' m})_{\ell m} \\
&= f_{\ell-1,m} (K_z Y_{\ell-1,m})_{\ell m} + f_{\ell+1,m} (K_z Y_{\ell+1,m})_{\ell m} \\
&= f_{\ell-1,m} B_{\ell,m} - f_{\ell+1,m} B_{\ell+1,m} \\
&= \sqrt{\frac{(\ell^2 - m^2)\ell^2}{4\ell^2 - 1}} f_{\ell-1,m} \\
&\quad - \sqrt{\frac{((\ell+1)^2 - m^2)(\ell+1)^2}{4(\ell+1)^2 - 1}} f_{\ell+1,m}
\end{aligned}$$

For K_x and K_y, we introduce the operators $K_+ = K_x + iK_y$ and $K_- = K_x - iK_y$ where i is the imaginary unit. Utilizing Equation (28) in [**DC14**] , we find:

$$\begin{aligned}
K_+ Y_{\ell m} &= -B_{\ell m} \sqrt{\frac{\ell - m - 1}{\ell + m}} Y_{\ell-1,m+1} - B_{\ell+1,m} \sqrt{\frac{\ell + m + 2}{\ell - m + 1}} Y_{\ell+1,m+1} \\
K_- Y_{\ell m} &= B_{\ell m} \sqrt{\frac{\ell + m - 1}{\ell - m}} Y_{\ell-1,m-1} + B_{\ell+1,m} \sqrt{\frac{\ell - m + 2}{\ell + m + 1}} Y_{\ell+1,m-1}
\end{aligned}$$

Since $K_x = \frac{1}{2}(K_+ + K_-)$ and $K_y = -\frac{i}{2}(K_+ - K_-)$, these equations allow us to express K_x and K_y as follows:

$$\begin{aligned}
K_x Y_{\ell m} =& \frac{1}{2} \left(-B_{\ell m} \sqrt{\frac{\ell - m - 1}{\ell + m}} Y_{\ell-1,m+1} - B_{\ell+1,m} \sqrt{\frac{\ell + m + 2}{\ell - m + 1}} Y_{\ell+1,m+1} \right) \\
&+ \frac{1}{2} \left(B_{\ell m} \sqrt{\frac{\ell + m - 1}{\ell - m}} Y_{\ell-1,m-1} + B_{\ell+1,m} \sqrt{\frac{\ell - m + 2}{\ell + m + 1}} Y_{\ell+1,m-1} \right) \\
K_y Y_{\ell m} =& -\frac{i}{2} \left(-B_{\ell m} \sqrt{\frac{\ell - m - 1}{\ell + m}} Y_{\ell-1,m+1} - B_{\ell+1,m} \sqrt{\frac{\ell + m + 2}{\ell - m + 1}} Y_{\ell+1,m+1} \right) \\
&+ \frac{i}{2} \left(B_{\ell m} \sqrt{\frac{\ell + m - 1}{\ell - m}} Y_{\ell-1,m-1} + B_{\ell+1,m} \sqrt{\frac{\ell - m + 2}{\ell + m + 1}} Y_{\ell+1,m-1} \right)
\end{aligned}$$

where $B_{\ell m} := \sqrt{\frac{(\ell^2 - m^2)\ell^2}{4\ell^2 - 1}}$.

By substituting these into our general expression for Kf, we obtain:

$$(K_x Y_{\ell' m'})_{\ell m} = \begin{cases} \frac{1}{2} B_{\ell+1,m+1} \sqrt{\frac{(\ell+1)+m}{(\ell+1)-(m+1)}} & \text{if } (\ell, m) = (\ell' - 1, m' - 1) \\ -\frac{1}{2} B_{\ell+1,m-1} \sqrt{\frac{(\ell+1)-(m-1)-1}{(\ell+1)+(m-1)}} & \text{if } (\ell, m) = (\ell' - 1, m' + 1) \\ \frac{1}{2} B_{\ell,m+1} \sqrt{\frac{(\ell-1)-(m+1)+2}{(\ell-1)+(m+1)+1}} & \text{if } (\ell, m) = (\ell' + 1, m' - 1) \\ -\frac{1}{2} B_{\ell,m-1} \sqrt{\frac{(\ell-1)+(m-1)+2}{(\ell-1)-(m-1)+1}} & \text{if } (\ell, m) = (\ell' + 1, m' + 1) \\ 0 & \text{otherwise.} \end{cases}$$

and

$$(K_y Y_{\ell' m'})_{\ell m} = \begin{cases} \frac{i}{2} B_{\ell+1,m+1} \sqrt{\frac{(\ell+1)+(m+1)-1}{(\ell+1)-(m+1)}} & \text{if } (\ell, m) = (\ell' - 1, m' - 1) \\ \frac{i}{2} B_{\ell+1,m-1} \sqrt{\frac{(\ell+1)-(m-1)-1}{(\ell+1)+(m-1)}} & \text{if } (\ell, m) = (\ell' - 1, m' + 1) \\ \frac{i}{2} B_{\ell,m+1} \sqrt{\frac{(\ell-1)-(m+1)+2}{(\ell-1)+(m+1)+1}} & \text{if } (\ell, m) = (\ell' + 1, m' - 1) \\ \frac{i}{2} B_{\ell,m-1} \sqrt{\frac{(\ell-1)+(m-1)+2}{(\ell-1)-(m-1)+1}} & \text{if } (\ell, m) = (\ell' + 1, m' + 1) \\ 0 & \text{otherwise.} \end{cases}$$

After simplifications, we get

$$(K_x Y_{\ell' m'})_{\ell m} = \begin{cases} \frac{1}{2} \sqrt{\frac{(l+m+2)(\ell+1)^2(\ell+m+1)}{4(\ell+1)^2-1}} & \text{if } (\ell, m) = (\ell' - 1, m' - 1) \\ -\frac{1}{2} \sqrt{\frac{(\ell-m+2)(\ell+1)^2(\ell-m+1)}{4(\ell+1)^2-1}} & \text{if } (\ell, m) = (\ell' - 1, m' + 1) \\ \frac{1}{2} \sqrt{\frac{(\ell-m-1)\ell^2(\ell-m)}{4\ell^2-1}} & \text{if } (\ell, m) = (\ell' + 1, m' - 1) \\ -\frac{1}{2} \sqrt{\frac{(\ell+m-1)\ell^2(\ell+m)}{4\ell^2-1}} & \text{if } (\ell, m) = (\ell' + 1, m' + 1) \\ 0 & \text{otherwise.} \end{cases}$$

and

$$(K_y Y_{\ell' m'})_{\ell m} = \begin{cases} \frac{i}{2} \sqrt{\frac{(\ell+m+2)(\ell+1)^2(\ell+m+1)}{4(\ell+1)^2-1}} & \text{if } (\ell, m) = (\ell' - 1, m' - 1) \\ \frac{i}{2} \sqrt{\frac{(\ell-m+2)(\ell+1)^2(\ell-m+1)}{4(\ell+1)^2-1}} & \text{if } (\ell, m) = (\ell' - 1, m' + 1) \\ \frac{i}{2} \sqrt{\frac{(\ell-m-1)\ell^2(\ell-m)}{4\ell^2-1}} & \text{if } (\ell, m) = (\ell' + 1, m' - 1) \\ \frac{i}{2} \sqrt{\frac{(\ell+m-1)\ell^2(\ell+m)}{4\ell^2-1}} & \text{if } (\ell, m) = (\ell' + 1, m' + 1) \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$
\begin{aligned}
(K_x f)_{\ell m} =& \sum_{\ell'} \sum_{m'} f_{\ell' m'} (K_x Y_{\ell' m'})_{\ell m} \\
=& f_{\ell-1,m-1} (K_x Y_{\ell-1,m-1})_{\ell m} + f_{\ell-1,m+1} (K_x Y_{\ell-1,m+1})_{\ell m} \\
&+ f_{\ell+1,m-1} (K_x Y_{\ell+1,m-1})_{\ell m} + f_{\ell+1,m+1} (K_x Y_{\ell+1,m+1})_{\ell m} \\
=& -\frac{1}{2} \sqrt{\frac{(\ell+m-1)\ell^2(\ell+m)}{4\ell^2-1}} f_{\ell-1,m-1} + \frac{1}{2} \sqrt{\frac{(\ell-m-1)\ell^2(\ell-m)}{4\ell^2-1}} f_{\ell-1,m+1} \\
&- \frac{1}{2} \sqrt{\frac{(\ell-m+2)(\ell+1)^2(\ell-m+1)}{4(\ell+1)^2-1}} f_{\ell+1,m-1} + \frac{1}{2} \sqrt{\frac{(\ell+m+2)(\ell+1)^2(\ell+m+1)}{4(\ell+1)^2-1}} f_{\ell+1,m+1}
\end{aligned}
$$

and

$$
\begin{aligned}
(K_y f)_{\ell m} =& \sum_{\ell'} \sum_{m'} f_{\ell' m'} (K_y Y_{\ell' m'})_{\ell m} \\
=& f_{\ell-1,m-1} (K_y Y_{\ell-1,m-1})_{\ell m} + f_{\ell-1,m+1} (K_y Y_{\ell-1,m+1})_{\ell m} \\
&+ f_{\ell+1,m-1} (K_y Y_{\ell+1,m-1})_{\ell m} + f_{\ell+1,m+1} (K_y Y_{\ell+1,m+1})_{\ell m} \\
=& \frac{i}{2} \sqrt{\frac{(\ell+m-1)\ell^2(\ell+m)}{4\ell^2-1}} f_{\ell-1,m-1} + \frac{i}{2} \sqrt{\frac{(\ell-m-1)\ell^2(\ell-m)}{4\ell^2-1}} f_{\ell-1,m+1} \\
&+ \frac{i}{2} \sqrt{\frac{(\ell-m+2)(\ell+1)^2(\ell-m+1)}{4(\ell+1)^2-1}} f_{\ell+1,m-1} + \frac{i}{2} \sqrt{\frac{(\ell+m+2)(\ell+1)^2(\ell+m+1)}{4(\ell+1)^2-1}} f_{\ell+1,m+1}.
\end{aligned}
$$

We use $Y_{0,0}$ as example. Recall that $Y_{00} = \frac{1}{2}\sqrt{\frac{1}{\pi}}$. Then

$$
\begin{aligned}
K_x Y_{0,0} &= \frac{1}{2} \sqrt{\frac{1}{\pi}} \cos\varphi \sin\theta \\
K_y Y_{0,0} &= \frac{1}{2} \sqrt{\frac{1}{\pi}} \sin\varphi \sin\theta \\
K_z Y_{0,0} &= \frac{1}{2} \sqrt{\frac{1}{\pi}} \cos\theta
\end{aligned}
$$

Recall that

$$Y_{1,-1}(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{3}{2\pi}} \cdot e^{-i\varphi} \cdot \sin\theta$$
$$Y_{1,0}(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{3}{\pi}} \cdot \cos\theta$$
$$Y_{1,1}(\theta,\varphi) = -\frac{1}{2}\sqrt{\frac{3}{2\pi}} \cdot e^{i\varphi} \cdot \sin\theta$$

Then we can rewrite $K_xY_{0,0}, K_yY_{0,0}, K_zY_{0,0}$ as

$$K_xY_{0,0} = -\sqrt{\frac{2\pi}{3}}Y_{1,1} + \sqrt{\frac{2\pi}{3}}Y_{1,-1}$$
$$K_yY_{0,0} = -i\sqrt{\frac{2\pi}{3}}Y_{1,1} + i\sqrt{\frac{2\pi}{3}}Y_{1,-1}$$
$$K_zY_{0,0} = \sqrt{\frac{1}{3\pi}}Y_{1,0}.$$

We use $Y_{0,0}$ as example. To clarify the action of the Lorentz Boost operators K_x, K_y, and K_z on the spherical harmonic $Y_{0,0}$, let's first remember that $Y_{0,0} = \frac{1}{2}\sqrt{\frac{1}{\pi}}$. Applying these operators to $Y_{0,0}$ gives:

$$K_xY_{0,0} = \frac{1}{2}\sqrt{\frac{1}{\pi}}\cos\varphi\sin\theta,$$
$$K_yY_{0,0} = \frac{1}{2}\sqrt{\frac{1}{\pi}}\sin\varphi\sin\theta,$$
$$K_zY_{0,0} = \frac{1}{2}\sqrt{\frac{1}{\pi}}\cos\theta.$$

For reference, the expressions for $Y_{1,-1}, Y_{1,0}$, and $Y_{1,1}$ are:

$$Y_{1,-1}(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{3}{2\pi}}e^{-i\varphi}\sin\theta,$$
$$Y_{1,0}(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta,$$
$$Y_{1,1}(\theta,\varphi) = -\frac{1}{2}\sqrt{\frac{3}{2\pi}}e^{i\varphi}\sin\theta.$$

We can then rewrite $K_x Y_{0,0}, K_y Y_{0,0}$, and $K_z Y_{0,0}$ as:

$$\begin{aligned}
K_x Y_{0,0} &= -\sqrt{\frac{2\pi}{3}} Y_{1,1} + \sqrt{\frac{2\pi}{3}} Y_{1,-1}, \\
K_y Y_{0,0} &= -i\sqrt{\frac{2\pi}{3}} Y_{1,1} + i\sqrt{\frac{2\pi}{3}} Y_{1,-1}, \\
K_z Y_{0,0} &= \sqrt{\frac{1}{3\pi}} Y_{1,0}.
\end{aligned}$$

Recall that

$$\begin{aligned}
Y_{0,0} &= \frac{1}{2}\sqrt{\frac{1}{\pi}}, \\
Y_{1,-1} &= \frac{1}{2}\sqrt{\frac{3}{2\pi}} e^{-i\varphi} \sin\theta, \\
Y_{1,0} &= \frac{1}{2}\sqrt{\frac{3}{\pi}} \cos\theta, \\
Y_{1,1} &= -\frac{1}{2}\sqrt{\frac{3}{2\pi}} e^{i\varphi} \sin\theta, \\
Y_{2,-2} &= \frac{1}{4}\sqrt{\frac{15}{\pi}} e^{-2i\varphi} \sin^2\theta, \\
Y_{2,-1} &= \frac{1}{2}\sqrt{\frac{15}{2\pi}} e^{-i\varphi} \sin\theta\cos\theta, \\
Y_{2,0} &= \frac{1}{4}\sqrt{\frac{5}{\pi}} (3\cos^2\theta - 1), \\
Y_{2,1} &= -\frac{1}{2}\sqrt{\frac{15}{2\pi}} e^{i\varphi} \sin\theta\cos\theta, \\
Y_{2,2} &= \frac{1}{4}\sqrt{\frac{15}{\pi}} e^{2i\varphi} \sin^2\theta.
\end{aligned}$$

For $\ell = 1$ and $m = -1$, we can rewrite $K_x Y_{1,-1}, K_y Y_{1,-1}, K_z Y_{1,-1}$ as:

$$\begin{aligned}
K_x Y_{1,-1} &= \frac{f_{0,0}}{\sqrt{6}} - \frac{2f_{2,-2}}{\sqrt{5}} + \sqrt{\frac{2}{15}} f_{2,0} \\
K_y Y_{1,-1} &= \frac{i f_{0,0}}{\sqrt{6}} + \frac{2i f_{2,-2}}{\sqrt{5}} + i\sqrt{\frac{2}{15}} f_{2,0} \\
K_z Y_{1,-1} &= -\frac{2f_{2,-1}}{\sqrt{5}}
\end{aligned}$$

For $\ell = 1$ and $m = 0$, we can rewrite $K_x Y_{1,0}, K_y Y_{1,0}, K_z Y_{1,0}$ as:

$$K_x Y_{1,0} = -\sqrt{\frac{2}{5}} f_{2,-1} + \sqrt{\frac{2}{5}} f_{2,1}$$
$$K_y Y_{1,0} = i\sqrt{\frac{2}{5}} f_{2,-1} + i\sqrt{\frac{2}{5}} f_{2,1}$$
$$K_z Y_{1,0} = \frac{f_{0,0}}{\sqrt{3}} - \frac{4 f_{2,0}}{\sqrt{15}}$$

For $\ell = 1$ and $m = 1$, we can rewrite $K_x Y_{1,1}, K_y Y_{1,1}, K_z Y_{1,1}$ as:

$$K_x Y_{1,1} = -\frac{f_{0,0}}{\sqrt{6}} - \sqrt{\frac{2}{15}} f_{2,0} + \frac{2 f_{2,2}}{\sqrt{5}}$$
$$K_y Y_{1,1} = \frac{i f_{0,0}}{\sqrt{6}} + i\sqrt{\frac{2}{15}} f_{2,0} + \frac{2i f_{2,2}}{\sqrt{5}}$$
$$K_z Y_{1,1} = -\frac{2 f_{2,1}}{\sqrt{5}}$$

Here, $f_{\ell,m}$ are the coefficients in the expansion of the function f in terms of spherical harmonics $Y_{\ell m}$.

3.4.7. Lorentz Boost on circle. The objective of this section is to illustrate the differential equation governing the Lorentz Boost on a circle. While the general proof can be challenging, examining it in the one-dimensional circle case offers a more straightforward but still insightful scenario. The problem only requires basic calculus, making it suitable for a quick verification.

3.4.7.1. *Lorentz Boost along x-axis.* Consider the Lorentz Boost applied to a function $f(\theta)$ along the x-axis ($\theta = 0$) with a rapidity η. Let this transformed function be $f_\eta(\theta)$. Specifically, we define the transformation as follows:

$$f_\eta(\theta) = J_\eta^r(\theta) f(\Phi_\eta(\theta)),$$
$$\Phi_\eta(\theta) = \arctan\left(\frac{\sin\theta}{\cosh\eta\cos\theta - \sinh\eta}\right),$$
$$J_\eta(\theta) = \frac{d}{d\theta}\Phi_\eta(\theta) = \frac{1}{\cosh\eta - \sinh\eta\cos\theta}.$$

We assert that $f_\eta(\theta)$ satisfies the following differential equation

$$\frac{d}{d\eta} f_\eta(\theta) = r\cos\theta f_\eta(\theta) + \sin\theta \frac{d}{d\theta} f_\eta(\theta).$$

It is worth noting that $\operatorname{div} X = \cos\theta$ and $Xf = \sin\theta \frac{d}{d\theta} f$.

To establish the above claim, we perform direct calculations on $\Phi_\eta(\theta)$. These calculations yield:

$$\begin{aligned}
\frac{d}{d\eta}\Phi_\eta(\theta) &= \frac{\sin\theta}{\cosh\eta - \sinh\eta\cos\theta} = \sin(\Phi_\eta(\theta)) \\
\frac{d}{d\theta}\Phi_\eta(\theta) &= \frac{1}{\cosh\eta - \sinh\eta\cos\theta} \\
\frac{d}{d\eta}\frac{d}{d\theta}\Phi_\eta(\theta) &= \frac{1}{\cosh\eta - \sinh\eta\cos\theta}\frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta} \\
&= \left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)\cos(\Phi_\eta(\theta)) \\
\frac{d}{d\theta}\frac{d}{d\theta}\Phi_\eta(\theta) &= -\frac{\sin\theta\sinh\eta}{(\cosh\eta - \cos\theta\sinh\eta)^2} = -\sin(\Phi_\eta(\theta))\sinh\eta\left(\frac{d}{d\theta}\Phi_\eta(\theta)\right).
\end{aligned}$$

Furthermore, by applying the chain rule to $f_\eta(\theta)$ with respect to η and θ, we get

$$\begin{aligned}
\frac{d}{d\eta} f_\eta(\theta) &= rJ_\eta^{r-1}(\theta)\frac{d}{d\eta}J_\eta(\theta)f(\Phi_\eta(\theta)) + J_\eta^r(\theta)f'(\Phi_\eta(\theta))\frac{d}{d\eta}\Phi_\eta(\theta) \\
&= r\left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)^{r-1}\left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)\cos(\Phi_\eta(\theta))f(\Phi_\eta(\theta)) \\
&\quad + \left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)^r f'(\Phi_\eta(\theta))\sin(\Phi_\eta(\theta)) \\
&= r\left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)^r\cos(\Phi_\eta(\theta))f(\Phi_\eta(\theta)) + \left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)^r f'(\Phi_\eta(\theta))\sin(\Phi_\eta(\theta)) \\
&= r\cos(\Phi_\eta(\theta))f_\eta(\theta) + \left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)^r f'(\Phi_\eta(\theta))\sin(\Phi_\eta(\theta)) \\
&= r\cos(\Phi_\eta(\theta))f_\eta(\theta) + \left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)^{r+1} f'(\Phi_\eta(\theta))\sin\theta \\
\frac{d}{d\theta} f_\eta(\theta) &= rJ_\eta^{r-1}(\theta)\frac{d}{d\theta}J_\eta(\theta)f(\Phi_\eta(\theta)) + J_\eta^r(\theta)f'(\Phi_\eta(\theta))\frac{d}{d\theta}\Phi_\eta(\theta) \\
&= rJ_\eta^r(\theta)(-\sin(\Phi_\eta(\theta))\sinh\eta)f(\Phi_\eta(\theta)) + J_\eta^r(\theta)f'(\Phi_\eta(\theta))\frac{d}{d\theta}\Phi_\eta(\theta) \\
&= -r\sin(\Phi_\eta(\theta))\sinh\eta f_\eta(\theta) + \left(\frac{d}{d\theta}\Phi_\eta(\theta)\right)^{r+1} f'(\Phi_\eta(\theta))
\end{aligned}$$

Combining these expressions allows us to eliminate f', resulting in an equation solely in terms of f_η:

$$\frac{d}{d\eta} f_\eta(\theta) = r(\cos(\Phi_\eta(\theta)) + \sin(\Phi_\eta(\theta)) \sinh\eta \sin\theta) f_\eta(\theta) + \frac{d}{d\theta} f_\eta(\theta) \sin\theta.$$

The first term in RHS can be simplified:

$$\begin{aligned}
&\cos(\Phi_\eta(\theta)) + \sin(\Phi_\eta(\theta)) \sinh\eta \sin\theta \\
=&\frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta} + \frac{\sin\theta}{\cosh\eta - \sinh\eta\cos\theta} \sinh\eta\sin\theta \\
=&\frac{\cosh\eta\cos\theta - \sinh\eta}{\cosh\eta - \sinh\eta\cos\theta} + \frac{\sin^2\theta\sinh\eta}{\cosh\eta - \sinh\eta\cos\theta} \\
=&\frac{\cosh\eta\cos\theta - \sinh\eta + \sinh\eta - \sinh\eta\cos^2\theta}{\cosh\eta - \sinh\eta\cos\theta} \\
=&\cos\theta.
\end{aligned}$$

This confirms the differential equation for $f_\eta(\theta)$:

$$\frac{d}{d\eta} f_\eta(\theta) = r\cos\theta f_\eta(\theta) + \sin\theta \frac{d}{d\theta} f_\eta(\theta).$$

Let t belong to $\mathbb{R}$ in a neighborhood of 0. Consider a continuously varying family of diffeomorphisms $\Phi_t : \mathbb{S}^1 \to \mathbb{S}^1$ that satisfy the differential equation for a given function $g : \mathbb{S}^1 \to \mathbb{R}$:

$$\begin{aligned}
\frac{d}{dt}\Phi_t(x) &= g(x)\frac{d}{dx}\Phi_t(x), \\
\Phi_0(x) &= x.
\end{aligned}$$

In this case, we form a pair (Φ_t, g). Alternatively, starting with a function $g : \mathbb{S}^1 \to \mathbb{R}$, we can construct $\Phi_t : \mathbb{S}^1 \to \mathbb{S}^1$ as:

$$\Phi_t(x) = \exp\left(tg(x)\frac{d}{dx}\right) x.$$

For any function f that is first-order differentiable and belongs to $L^2(\mathbb{S}^1, \mathbb{R})$, and for any $r \in \mathbb{R}$, we define $f_t : \mathbb{S}^1 \to \mathbb{R}$ as follows:

$$f_t(x) := \left(\frac{d}{dx}\Phi_t(x)\right)^r f(\Phi_t(x)).$$

Upon differentiation, $f_t(x)$ satisfies the differential equation:

$$\begin{aligned}\frac{d}{dt}f_t(x) &= r\frac{d}{dx}\left(\frac{\frac{d}{dt}\Phi_t(x)}{\frac{d}{dx}\Phi_t(x)}\right)f_t(x) + \left(\frac{\frac{d}{dt}\Phi_t(x)}{\frac{d}{dx}\Phi_t(x)}\right)\frac{d}{dx}f_t(x) \\ &= r\left(\frac{d}{dx}g(x)\right)f_t(x) + g(x)\frac{d}{dx}f_t(x)\end{aligned}$$

This allows us to express $f_t(x)$ as an exponential function of a linear operator acting on f:

$$f_t(x) = \exp\left(t\left(r\frac{d}{dx}g(x) + g(x)\frac{d}{dx}\right)\right)f(x).$$

To verify this, we perform further differentiation of $f_t(x)$ with respect to t and x:

$$\frac{d}{dt}f_t(x) = r\left(\frac{d}{dx}\Phi_t(x)\right)^{r-1}\left(\frac{d^2}{dtdx}\Phi_t(x)\right)f(\Phi_t(x)) + \left(\frac{d}{dx}\Phi_t(x)\right)^{r}f'(\Phi_t(x))\frac{d}{dt}\Phi_t(x)$$

$$\frac{d}{dx}f_t(x) = r\left(\frac{d}{dx}\Phi_t(x)\right)^{r-1}\left(\frac{d^2}{dx^2}\Phi_t(x)\right)f(\Phi_t(x)) + \left(\frac{d}{dx}\Phi_t(x)\right)^{r}f'(\Phi_t(x))\frac{d}{dx}\Phi_t(x)$$

We then combine both equations to eliminate f':

$$\begin{aligned}\frac{d}{dt}f_t(x) &= r\left(\frac{d}{dx}\Phi_t(x)\right)^{r-1}\left(\frac{d^2}{dtdx}\Phi_t(x)\right)f(\Phi_t(x)) + \\ &\qquad \frac{1}{\frac{d}{dx}\Phi_t(x)}\left(\frac{d}{dx}f_t(x) - r\left(\frac{d}{dx}\Phi_t(x)\right)^{r-1}\left(\frac{d^2}{dx^2}\Phi_t(x)\right)f(\Phi_t(x))\right)\frac{d}{dt}\Phi_t(x) \\ &= r\left(\frac{\frac{d^2}{dxdt}\Phi_t(x)}{\frac{d}{dx}\Phi_t(x)} - \frac{\frac{d^2}{dx^2}\Phi_t(x)\frac{d}{dt}\Phi_t(x)}{\left(\frac{d}{dx}\Phi_t(x)\right)^2}\right)f_t(x) + \left(\frac{\frac{d}{dt}\Phi_t(x)}{\frac{d}{dx}\Phi_t(x)}\right)\frac{d}{dx}f_t(x) \\ &= r\frac{d}{dx}\left(\frac{\frac{d}{dt}\Phi_t(x)}{\frac{d}{dx}\Phi_t(x)}\right)f_t(x) + \left(\frac{\frac{d}{dt}\Phi_t(x)}{\frac{d}{dx}\Phi_t(x)}\right)\frac{d}{dx}f_t(x) \\ &= r\left(\frac{d}{dx}g(x)\right)f_t(x) + g(x)\frac{d}{dx}f_t(x) \quad \text{where } g(x) = \frac{\frac{d}{dt}\Phi_t(x)}{\frac{d}{dx}\Phi_t(x)}.\end{aligned}$$

Let $K_x : L^2(\mathbb{S}^1) \to L^2(\mathbb{S}^1)$ be the Lorentz Boost operator defined as:

$$(K_x f)(\theta) = r\cos\theta f(\theta) + \sin\theta\frac{d}{d\theta}f(\theta).$$

The solution of the differential equation

$$\frac{\partial}{\partial t} f_t = K_x f_t$$

$$f_0 = f$$

can be represented as an exponential function of a linear operator:

$$f_\eta = \mathbf{B}_{\eta,0} f = \exp\left(\eta K_x\right) f.$$

Furthermore, we show that K_x has a sparse representation in the real Fourier basis $(e_k : k \in \ldots, -2, -1, 0, 1, 2, \ldots)$

$$\begin{aligned}
&(e_k : k \in \ldots, -2, -1, 0, 1, 2, \ldots) \\
&:= \left\{ \frac{1}{\sqrt{\pi}} \sin(k\theta) : k = -1, -2, \ldots \right\} \bigcup \left\{ \frac{1}{\sqrt{2\pi}} \right\} \bigcup \left\{ \frac{1}{\sqrt{\pi}} \cos(k\theta) : k = 1, 2, \ldots \right\}.
\end{aligned}$$

By direct calculation,

$$\begin{aligned}
K_x \cos(k\theta) &= -\frac{1}{2}(k-r)\cos((k-1)\theta) + \frac{1}{2}(k+r)\cos((k+1)\theta) \\
K_x 1 &= r\cos\theta \\
K_x \sin(k\theta) &= -\frac{1}{2}(k-r)\sin((k-1)\theta) + \frac{1}{2}(k+r)\sin((k+1)\theta),
\end{aligned}$$

the real Fourier decomposition of $K_x f$ turns out to be:

$$K_x f = \begin{pmatrix}
-\frac{-2-r}{2} & & & & \\
0 & -\frac{-1-r}{2} & & & \\
\frac{-2+r}{2} & 0 & 0 & & \\
 & 0 & 0 & -\frac{1-r}{\sqrt{2}} & \\
 & & \frac{r}{\sqrt{2}} & 0 & -\frac{2-r}{2} \\
 & & & \frac{1+r}{2} & 0 \\
 & & & & \frac{2+r}{2}
\end{pmatrix} \begin{pmatrix} \vdots \\ a_{-2} \\ a_{-1} \\ a_0 \\ a_1 \\ a_2 \\ \vdots \end{pmatrix}.$$

We derive the Lorentz Boost in the direction of the x-axis with a rapidity of η, expressed in the complex Fourier representation. We begin by defining $J_\eta(\theta) = \frac{d}{d\theta}\Phi_\eta(\theta)$. Next, let $f_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta}$.

Using the Fourier representations $\cos\theta = \frac{\sqrt{2\pi}}{2}(f_{-1}(\theta) + f_1(\theta))$ and $\sin(\theta) = \frac{\sqrt{2\pi}}{2}i(f_{-1}(\theta) - f_1(\theta))$, the derivative becomes:

$$\frac{d}{d\eta}(J_\eta^r f(\Phi_\eta(\theta))) = rJ_\eta^{r-1}J_\eta \cos(\Phi_\eta(\theta))f(\Phi_\eta(\theta)) + J_\eta^r f'(\Phi_\eta(\theta))\sin(\Phi_\eta(\theta)).$$

Let $f_n(\theta) = \frac{1}{\sqrt{2\pi}}e^{in\theta}$. Since $\cos\theta = \frac{\sqrt{2\pi}}{2}f_{-1}(\theta) + \frac{\sqrt{2\pi}}{2}f_1(\theta)$ and $\sin(\theta) = \frac{\sqrt{2\pi}}{2}if_{-1}(\theta) - \frac{\sqrt{2\pi}}{2}if_1(\theta)$, we have

$$\begin{aligned}
\frac{d}{d\eta}(J_\eta^r f_n(\Phi_\eta(\theta))) &= rJ_\eta^r\left(\frac{\sqrt{2\pi}}{2}f_{-1}(\Phi_\eta(\theta)) + \frac{\sqrt{2\pi}}{2}f_1(\Phi_\eta(\theta))\right)f_n(\Phi_\eta(\theta)) \\
&\quad + J_\eta^r f_n'(\Phi_\eta(\theta))\left(\frac{\sqrt{2\pi}}{2}if_{-1}(\Phi_\eta(\theta)) - \frac{\sqrt{2\pi}}{2}if_1(\Phi_\eta(\theta))\right) \\
&= rJ_\eta^r\left(\frac{\sqrt{2\pi}}{2}f_{-1}(\Phi_\eta(\theta)) + \frac{\sqrt{2\pi}}{2}f_1(\Phi_\eta(\theta))\right)f_n(\Phi_\eta(\theta)) \\
&\quad + J_\eta^r inf_n(\Phi_\eta(\theta))\left(\frac{\sqrt{2\pi}}{2}if_{-1}(\Phi_\eta(\theta)) - \frac{\sqrt{2\pi}}{2}if_1(\Phi_\eta(\theta))\right) \\
&= rJ_\eta^r\left(\frac{1}{2}f_{n-1}(\Phi_\eta(\theta)) + \frac{1}{2}f_{n+1}(\Phi_\eta(\theta))\right) \\
&\quad + nJ_\eta^r\left(-\frac{1}{2}f_{n-1}(\Phi_\eta(\theta)) + \frac{1}{2}f_{n+1}(\Phi_\eta(\theta))\right) \\
&= -\frac{n-r}{2}J_\eta^r f_{n-1}(\Phi_\eta(\theta)) + \frac{n+r}{2}J_\eta^r f_{n+1}(\Phi_\eta(\theta)).
\end{aligned}$$

Having derived this equation, we can now proceed to find the complex Fourier decomposition of $\mathbf{K}_x f$, for a general function $f = \sum_n a_n f_n$. The Fourier coefficients of $\mathbf{K}_x f$ can be represented as:

$$\mathbf{K}_x f = \begin{pmatrix}
-\frac{-2-r}{2} & & & & \\
0 & -\frac{-1-r}{2} & & & \\
\frac{-2+r}{2} & 0 & -\frac{0-r}{2} & & \\
 & \frac{-1+r}{2} & 0 & -\frac{1-r}{2} & \\
 & & \frac{0+r}{2} & 0 & -\frac{2-r}{2} \\
 & & & \frac{1+r}{2} & 0 \\
 & & & & \frac{2+r}{2}
\end{pmatrix}\begin{pmatrix}
\vdots \\ a_{-2} \\ a_{-1} \\ a_0 \\ a_1 \\ a_2 \\ \vdots
\end{pmatrix}.$$

3.4.7.2. *Lorentz Boost along y-axis.* Similarly, we can derive the Lorentz Boost for functions in $L^2(\mathbb{S}^2)$ with rapidity η, directed along the positive y-axis ($\theta = \pi/2$). We start by defining the angle $\Phi_\eta(\theta)$ that captures light aberration:

$$\begin{aligned}
\Phi_\eta(\theta) &:= \arctan\left(\frac{\cosh\eta\sin\theta - \sinh\eta}{\cos\theta}\right) \\
\frac{d}{d\theta}\Phi_\eta(\theta) &= \frac{1}{\cosh\eta - \sinh\eta\sin\theta} \\
\frac{d}{d\eta}\Phi_\eta(\theta) &= -\cos(\Phi_\eta(\theta)) = -\frac{\cos\theta}{\cosh\eta - \sinh\eta\sin\theta}.
\end{aligned}$$

We also introduce the function $V(\theta) = -\cos(\theta)$ and the Lorentz Boost operator in the y-direction, $\mathbf{K}_y f(\theta)$, defined as follows:

$$(\mathbf{K}_y f)(\theta) = r\sin\theta f(\theta) - \cos\theta\frac{d}{d\theta}f(\theta).$$

Using these definitions, we can examine the action of $\mathbf{K}_y$ on various functions:

$$\begin{aligned}
\mathbf{K}_y\cos(k\theta) &= -\frac{1}{2}(k+r)\sin((-k-1)\theta) - \frac{1}{2}(k-r)\sin((-k+1)\theta) \\
\mathbf{K}_y 1 &= -r\sin(-\theta) \\
\mathbf{K}_y\sin(k\theta) &= -\frac{1}{2}(k+r)\cos((-k-1)\theta) - \frac{1}{2}(k-r)\cos((-k+1)\theta).
\end{aligned}$$

Finally, the real Fourier decomposition of $\mathbf{K}_y f(\theta)$ can be expressed as:

$$\begin{pmatrix}
 & & & & -\frac{2+r}{2} \\
 & & & -\frac{1+r}{2} & 0 \\
 & & -\frac{r}{\sqrt{2}} & 0 & -\frac{2-r}{2} \\
 & -\frac{-1+r}{\sqrt{2}} & 0 & 0 & \\
-\frac{-2+r}{2} & 0 & 0 & & \\
0 & -\frac{-1-r}{2} & & & \\
-\frac{-2-r}{2} & & & &
\end{pmatrix}
\begin{pmatrix} \vdots \\ a_{-2} \\ a_{-1} \\ a_0 \\ a_1 \\ a_2 \\ \vdots \end{pmatrix}.$$

3.4.7.3. *Lorentz Boost in any direction.* We aim to derive the Rotation operator with angle ϕ and the Lorentz Boost operator along the x-axis with rapidity η. Having these two operators, we

can construct the Lorentz Boost in the $(\cos\phi, \sin\phi)$-direction with rapidity η. The process involves rotating function f from the direction $(\cos\phi, \sin\phi)$ to the x-axis $(\cos 0, \sin 0)$, applying a Lorentz Boost along the x-axis with rapidity η to f, and then rotating f back to its original direction:

$$f \mapsto \exp(\phi\mathbf{R})f \mapsto \exp(\eta\mathbf{K}_x)\exp(\phi\mathbf{R})f \mapsto \exp(-\phi\mathbf{R})\exp(\eta\mathbf{K}_x)\exp(\phi\mathbf{R})f.$$

This can be concisely written as:

$$\mathbf{B}_{\eta,\phi}f = \exp(-\phi\mathbf{R})\exp(\eta\mathbf{K}_x)\exp(\phi\mathbf{R})f.$$

Since $\exp(\phi\mathbf{R})$ and $\exp(-\phi\mathbf{R})$ are inverses, we can simplify:

$$\exp(-\phi\mathbf{R})\exp(\eta\mathbf{K}_x)\exp(\phi\mathbf{R}) = \exp(\eta\exp(-\phi\mathbf{R})\mathbf{K}\exp(\phi\mathbf{R})).$$

By performing direct calculations, we find:

$$\begin{aligned}\exp(-\phi\mathbf{R})\mathbf{K}_x\exp(\phi\mathbf{R}) &= \begin{pmatrix} & -\frac{-1-r}{2}\cos\phi & & \frac{1+r}{2}\sin\phi & \\ \frac{-2+r}{2}\cos\phi & & -\frac{r}{\sqrt{2}}\sin\phi & & \frac{-2+r}{2}\sin\phi \\ & -\frac{-1+r}{\sqrt{2}}\sin\phi & & -\frac{1-r}{\sqrt{2}}\cos\phi & \\ -\frac{-2+r}{2}\sin\phi & & \frac{r}{\sqrt{2}}\cos\phi & & -\frac{2-r}{2}\cos\phi \\ & -\frac{-2-r}{2}\sin\phi & & \frac{1+r}{2}\cos\phi & \end{pmatrix} \\ &= \cos\phi\mathbf{K}_x + \sin\phi\mathbf{K}_y.\end{aligned}$$

For a deeper understanding, let's consider the Lie algebraic representation:

$$\begin{aligned}\exp(-\phi\mathbf{R})\exp(\eta\mathbf{K}_x)\exp(\phi\mathbf{R}) &= \exp(\eta\exp(-\phi\mathbf{R})\mathbf{K}_x\exp(\phi\mathbf{R})) \\ &= \exp(\mathrm{Ad}(\exp(-\phi\mathbf{R}))(\eta\mathbf{K}_x)) \\ &= \exp(\exp(\mathrm{ad}(-\phi\mathbf{R}))(\eta\mathbf{K}_x)).\end{aligned}$$

It can be verified that the commutators are:

$$[\mathbf{K}_x, \mathbf{K}_y] = \mathbf{R}$$

$$[\mathbf{R}, \mathbf{K}_x] = -\mathbf{K}_y$$

$$[\mathbf{R}, \mathbf{K}_y] = \mathbf{K}_x.$$

Therefore, we have:

$$(\mathrm{ad}(\mathbf{R}))^k \mathbf{K}_x = \begin{cases} (-1)^{\frac{k}{2}} \mathbf{K}_x & \text{if } k = 0, 2, 4, \dots \\ (-1)^{\frac{k-1}{2}} \mathbf{K}_y & \text{if } k = 1, 3, 5, \dots \end{cases}$$

Using this, we get:

$$\begin{aligned} \exp(\mathrm{ad}(-\phi\mathbf{R}))(\eta\mathbf{K}_x) &= \sum_{k=0}^{\infty} \frac{1}{k!} (\mathrm{ad}(-\phi\mathbf{R}))^k (\eta\mathbf{K}_x) \\ &= \sum_{k=0,2,4\dots} \frac{1}{k!} (\mathrm{ad}(-\phi\mathbf{R}))^k (\eta\mathbf{K}_x) + \sum_{k=1,3,5\dots}^{\infty} \frac{1}{k!} (\mathrm{ad}(-\phi\mathbf{R}))^k (\eta\mathbf{K}_x) \\ &= \eta \sum_{k=0,2,4\dots} \frac{(-\phi)^k (-1)^{\frac{k}{2}}}{k!} \mathbf{K}_x + \eta \sum_{k=1,3,5\dots}^{\infty} \frac{(-\phi)^k (-1)^{\frac{k-1}{2}}}{k!} \mathbf{K}_y \\ &= \eta \cos\phi \mathbf{K}_x + \eta \sin(-\phi) \mathbf{K}_y \\ &= \eta(\cos\phi \mathbf{K}_x - \sin\phi \mathbf{K}_y). \end{aligned}$$

Thus, the rotate-boost-rotate sequence can be expressed as:

$$\exp(-\phi\mathbf{R}) \exp(\eta\mathbf{K}_x) \exp(\phi\mathbf{R}) = \exp(\eta_x \mathbf{K}_x + \eta_y \mathbf{K}_y).$$

In this case, the Lorentz Boost with rapidity vector $\boldsymbol{\eta} = (\eta\cos\phi, \eta\sin\phi) = (\eta_x, \eta_y)$ becomes:

$$\mathbf{B}_{\boldsymbol{\eta}} f = \exp(\eta_x \mathbf{K}_x + \eta_y \mathbf{K}_y) f.$$

We can further define the operator $\mathbf{K}_\phi := \cos\phi\mathbf{K}_x + \sin\phi\mathbf{K}_y$ such that $\mathbf{B}_{\eta,\phi} = \exp(\eta\mathbf{K}_\phi)$, for

$$\begin{aligned}
&\eta_x\mathbf{K}_x f + \eta_y\mathbf{K}_y f \\
=&\eta_x\left(r\cos\theta f(\theta) + \sin\theta\frac{d}{d\theta}f(\theta)\right) + \eta_y\left(r\sin\theta f(\theta) - \cos\theta\frac{d}{d\theta}f(\theta)\right) \\
=&\eta\cos\phi\left(r\cos\theta f(\theta) + \sin\theta\frac{d}{d\theta}f(\theta)\right) + \eta\sin\phi\left(r\sin\theta f(\theta) - \cos\theta\frac{d}{d\theta}f(\theta)\right) \\
=&\eta\left(r\cos(\theta-\phi) f(\theta) + \sin(\theta-\phi)\frac{d}{d\theta}f(\theta)\right) \\
=&\eta\mathbf{K}_\phi f.
\end{aligned}$$

For any function f in $L^2(\mathbb{S}^1)$, the Lorentz Boosted function $f_{\eta,\phi} := \mathbf{B}_{\eta,\phi}f$ satisfies the following differential equation:

$$\frac{d}{d\eta}f_{\eta,\phi}(\theta) = \eta r\cos(\theta-\phi)f_{\eta,\phi}(\theta) + \eta\sin(\theta-\phi)\frac{d}{d\theta}f_{\eta,\phi}(\theta),$$

with the initial condition $f_{0,\phi}(\theta) = f(\theta)$.

3.5. Dipole and Lorentz Beat

In this section, we investigate the effects of the Lorentz Boost on anisotropy, focusing on two particular outcomes: the low-frequency dipole effect and a high-frequency phenomenon resembling acoustic beats, which we term the "Lorentz Beat." These effects suggest that high-frequency components of the CMB could contain signatures of the Lorentz Boost. Accordingly, high-frequency analysis is essential for accurately estimating the Lorentz Boost.

Let $(\mathbf{e}_k : k \in \mathbb{Z})$ be a real Fourier basis given by

$$(\mathbf{e}_k : k \in \mathbb{Z}) := \left\{\frac{1}{\sqrt{\pi}}\sin(k\theta) : k < 0\right\} \cup \left\{\frac{1}{\sqrt{2\pi}} : k = 0\right\} \cup \left\{\frac{1}{\sqrt{\pi}}\cos(k\theta) : k > 0\right\}.$$

First, let's define what we mean by the term *dipole* in this context. For any L^2 function f on the circle $\mathbb{S}^1$, its dipole is the component corresponding to the Fourier frequencies $|k| = 1$. Specifically, if f has the Fourier series $f = \sum_{k\in\mathbb{Z}} a_k\mathbf{e}_k$, then its dipole consists of the terms $a_{-1}\mathbf{e}_{-1} + a1\mathbf{e}_1$.

Next, consider the Lorentz Boost operator $\mathbf{B}_\eta : L^2(\mathbb{S}^1, \mathbb{R}) \to L^2(\mathbb{S}^1, \mathbb{R})$, defined as

$$f\eta(\theta) := \mathbf{B}_\eta f = J^r \eta f(\Phi_\eta(\theta)).$$

From earlier discussions, we know that

$$\mathbf{B}_\eta = \exp\left(\eta_x \mathbf{K}_x + \eta_y \mathbf{K} y\right),$$

and in a first-order approximation,

$$\mathbf{B}_\eta \approx \mathbf{I} + \eta_x \mathbf{K}_x + \eta_y \mathbf{K}_y,$$

where $\mathbf{I}$ is the identity operator. Our aim is to examine the forms of $\mathbf{K}_x f$ and $\mathbf{K}_y f$ for $f \in L^2(\mathbb{S}^1, \mathbb{R})$ and argue that they resemble a dipole.

To do this, we introduce the concept of *Lorentz Beat* and define the normalized Lorentz Beat functions as follows:

$$\begin{aligned}
\mathbf{b}_x^{k,r}(\theta) &:= \frac{1}{\sqrt{\pi}} \frac{r\cos(\theta) + |k|\sin(\theta)}{\sqrt{k^2 + r^2}} = \frac{r}{\sqrt{k^2 + r^2}} \mathbf{e}_1 + \frac{-|k|}{\sqrt{k^2 + r^2}} \mathbf{e}_{-1} \\
\mathbf{b}_y^{k,r}(\theta) &:= \frac{1}{\sqrt{\pi}} \frac{-|k|\cos(\theta) + r\sin(\theta)}{\sqrt{k^2 + r^2}} = \frac{-|k|}{\sqrt{k^2 + r^2}} \mathbf{e}_1 + \frac{-r}{\sqrt{k^2 + r^2}} \mathbf{e}_{-1}
\end{aligned}$$

It is notable that the Lorentz Beat functions are dipole functions and form an orthonormal pair. As $k \to \infty$, these normalized functions approach the limit

$$\begin{aligned}
\mathbf{b}_x^{k,r}(\theta) &\to -\mathbf{e}_{-1}(\theta) \quad \text{as } k \to \infty \\
\mathbf{b}_y^{k,r}(\theta) &\to -\mathbf{e}_1(\theta) \quad \text{as } k \to \infty.
\end{aligned}$$

The introduction of Lorentz Beat is motivated by the phenomenon of acoustic beats, in which the superposition of two sound waves with similar frequencies results in an interference pattern. For instance, two waves with frequencies k_1 and k_2 will produce a beat given by

$$\cos(k_1\theta) + \cos(k_2\theta) = 2\cos\left(\frac{k_1 + k_2}{2}\theta\right)\cos\left(\frac{k_1 - k_2}{2}\theta\right).$$

The resulted wave has frequency $\left|\frac{k_1+k_2}{2}\right|$, which is the average of k_1 and k_2. However, has an envelop of frequency $\left|\frac{k_1-k_2}{2}\right|$. As we assume k_1 and k_2 are similar, $\left|\frac{k_1-k_2}{2}\right|$ is small. As a result, the resulting sound wave appears to be like a low frequency sound, or beat. The audible beat frequency is $|k_1-k_2|$. We can then relate this concept back to the Lorentz Boost and its effects. For example, for $k=0$, we have the following: We have, for $k=0$

$$(\mathbf{K}_x\mathbf{e}_k)(\theta) = \frac{r}{\sqrt{2}}\frac{1}{\sqrt{\pi}}\cos(\theta) = \frac{r}{\sqrt{2}}\mathbf{e}_1$$
$$(\mathbf{K}_y\mathbf{e}_k)(\theta) = \frac{r}{\sqrt{2}}\frac{1}{\sqrt{\pi}}\sin(\theta) = -\frac{r}{\sqrt{2}}\mathbf{e}_{-1}$$

and for $k \geq 1$

$$\begin{aligned}(\mathbf{K}_x\mathbf{e}_k)(\theta) &= \frac{1}{\sqrt{\pi}}\left[-\frac{1}{2}(k-r)\cos((k-1)\theta) + \frac{1}{2}(k+r)\cos((k+1)\theta)\right] \\ &= \frac{1}{\sqrt{\pi}}\left[r\cos(\theta)\cos(k\theta) - k\sin(\theta)\sin(k\theta)\right] \\ (\mathbf{K}_y\mathbf{e}_k)(\theta) &= \frac{1}{\sqrt{\pi}}\left[-\frac{1}{2}(k+r)\sin((-k-1)\theta) - \frac{1}{2}(k-r)\sin((-k+1)\theta)\right] \\ &= \frac{1}{\sqrt{\pi}}\left[r\sin(\theta)\cos(k\theta) + k\cos(\theta)\sin(k\theta)\right]\end{aligned}$$

and for $k \leq -1$

$$\begin{aligned}(\mathbf{K}_x\mathbf{e}_k)(\theta) &= \frac{1}{\sqrt{\pi}}\left[-\frac{1}{2}(k-r)\sin((k-1)\theta) + \frac{1}{2}(k+r)\sin((k+1)\theta)\right] \\ &= \frac{1}{\sqrt{\pi}}\left[r\cos(\theta)\sin(k\theta) + k\sin(\theta)\cos(k\theta)\right] \\ (\mathbf{K}_y\mathbf{e}_k)(\theta) &= \frac{1}{\sqrt{\pi}}\left[\frac{1}{2}(k-r)\cos((-k-1)\theta) + \frac{1}{2}(k+r)\cos((-k+1)\theta)\right] \\ &= \frac{1}{\sqrt{\pi}}\left[r\sin(\theta)\sin(k\theta) - k\cos(\theta)\cos(k\theta)\right].\end{aligned}$$

For $k \geq 1$ and $k \leq -1$, detailed calculations show that the normalized operators $\mathbf{K}_x\mathbf{e}_k/N_k$ and $\mathbf{K}_y\mathbf{e}_k/N_k$ approximately match the Lorentz Beats $\mathbf{b}_x^{k,r}(\theta)$ and $\mathbf{b}_y^{k,r}(\theta)$, respectively, where $N_k := \sqrt{\frac{1}{2}(k^2+r^2)}$.

We are ready to look at a phenomenon that we call the Lorentz Beat phenomenon. Given a function $f = \sum_k a_k\mathbf{e}_k$, we examine the first-order approximation of the Lorentz Boost with rapidity

$\boldsymbol{\eta} = (\eta_x, \eta_y)$ around the origin $\mathbf{0} = (0, 0)$:

$$\mathbf{B}_{\boldsymbol{\eta}} f \approx f + \eta_x \mathbf{K}_x f + \eta_y \mathbf{K}_y f$$

Here, we regard the Lorentz Boost as a small perturbation of f characterized by $\eta_x \mathbf{K}_x f + \eta_y \mathbf{K}_y f$. We aim to show that this perturbation resembles a dipole function. To make this argument, we approximate $\mathbf{K}_x \mathbf{e}_k$ and $\mathbf{K}_y \mathbf{e}_k$ using Lorentz Beat functions $\mathbf{b}_x^{k,r}$ and $\mathbf{b}_y^{k,r}$:

$$\begin{aligned}
&\eta_x \mathbf{K}_x f + \eta_y \mathbf{K}_y f \\
=&\eta_x \mathbf{K}_x \sum_k a_k \mathbf{e}_k + \eta_y \mathbf{K}_y \sum_k a_k \mathbf{e}_k \\
=&\eta_x \sum_k a_k \mathbf{K}_x \mathbf{e}_k + \eta_y \sum_k a_k \mathbf{K}_y \mathbf{e}_k \\
\approx&\eta_x \sum_k a_k N_k \left(\frac{r}{\sqrt{k^2 + r^2}} \mathbf{e}_1 + \frac{-|k|}{\sqrt{k^2 + r^2}} \mathbf{e}_{-1} \right) + \eta_y \sum_k a_k N_k \left(\frac{-|k|}{\sqrt{k^2 + r^2}} \mathbf{e}_1 + \frac{-r}{\sqrt{k^2 + r^2}} \mathbf{e}_{-1} \right) \\
=&\left(\sum_k \left(\eta_x \frac{r}{\sqrt{2}} + \eta_y \frac{-|k|}{\sqrt{2}} \right) a_k \right) \mathbf{e}_1 + \left(\sum_k \left(\eta_x \frac{-|k|}{\sqrt{2}} + \eta_y \frac{-r}{\sqrt{2}} \right) a_k \right) \mathbf{e}_{-1}
\end{aligned}$$

Observe the $O(k)$ multiplier on the coefficient a_k. For an L^2 function with Fourier coefficients $(a_k : k \in \mathbb{Z})$, it is necessary to assume $\sum_k k^2 a_k^2 < \infty$ for the Fourier series to be absolutely convergent, i.e. $\sum_k a_k^2 < \infty$. We need to further assume $\sum_k k^2 a_k^2 < \infty$.

$$\begin{aligned}
&= \sum_k a_k \mathbf{e}_k + \eta_x \mathbf{K}_x \sum_k a_k \mathbf{e}_k + \eta_y \mathbf{K}_y \sum_k a_k \mathbf{e}_k \\
&= \sum_k a_k \mathbf{e}_k + \eta_x \sum_k a_k \mathbf{K}_x \mathbf{e}_k + \eta_y \sum_k a_k \mathbf{K}_y \mathbf{e}_k \\
&\approx \sum_k a_k \mathbf{e}_k + \eta_x \sum_k a_k \left(\frac{r}{\sqrt{k^2 + r^2}} \mathbf{e}_1 + \frac{-|k|}{\sqrt{k^2 + r^2}} \mathbf{e}_{-1} \right) + \eta_y \sum_k a_k \left(\frac{-|k|}{\sqrt{k^2 + r^2}} \mathbf{e}_1 + \frac{-r}{\sqrt{k^2 + r^2}} \mathbf{e}_{-1} \right) \\
&= a_0 \mathbf{e}_0 + \left(a_1 + \sum_k \left(\eta_x \frac{r}{\sqrt{2}} + \eta_y \frac{-|k|}{\sqrt{2}} \right) a_k \right) \mathbf{e}_1 + \left(a_{-1} + \sum_k \left(\eta_x \frac{-|k|}{\sqrt{2}} + \eta_y \frac{-r}{\sqrt{2}} \right) a_k \right) \mathbf{e}_{-1} + \sum_{|k|>1} a_k \mathbf{e}_k
\end{aligned}$$

We define the Lorentz Boost Dipole to be

$$a_1 \to a_1 + \sum_k \left(\eta_x \frac{r}{\sqrt{2}} + \eta_y \frac{-|k|}{\sqrt{2}} \right) a_k$$
$$a_{-1} \to a_{-1} + \sum_k \left(\eta_x \frac{-|k|}{\sqrt{2}} + \eta_y \frac{-r}{\sqrt{2}} \right) a_k$$

$$a_1 \to a_1 + \eta_x \sum_k \frac{r}{\sqrt{2}} a_k + \eta_y \sum_k \frac{-|k|}{\sqrt{2}} a_k$$
$$a_{-1} \to a_{-1} + \eta_x \sum_k \frac{-|k|}{\sqrt{2}} a_k + \eta_y \sum_k \frac{-r}{\sqrt{2}} a_k$$

Now let's examine the alternative formulation:

$$\mathbf{B}_{\boldsymbol{\eta}} f \approx f + \eta \mathbf{K}_\phi f = f + \eta r \cos(\theta - \phi) f + \eta \sin(\theta - \phi) \frac{d}{d\theta} f$$

Interestingly, f and $\frac{d}{d\theta} f$ are orthogonal in $L^2(\mathbb{S}^1, \mathbb{R})$:

$$\int_{-\pi}^{\pi} f \frac{d}{d\theta} f d\theta = \int_{-\pi}^{\pi} \frac{1}{2} \frac{d}{d\theta} f^2 d\theta = 0.$$

We now turn our attention to the two distinct components $r \cos(\theta - \phi) f$ and $\sin(\theta - \phi) \frac{d}{d\theta} f$. The term $\frac{d}{d\theta} f$ is especially prominent for high-frequency signals, represented as $\sum_k k a_k \mathbf{e}_k$. We would like to plot the decomposition of $(\mathbf{B}_{\boldsymbol{\eta}} f)$ in Figure 3.3 and 3.4:

$$\mathbf{B}_{\boldsymbol{\eta}} f \approx f + \eta r \cos(\theta - \phi) f + \eta \sin(\theta - \phi) \frac{d}{d\theta} f. \tag{3.4}$$

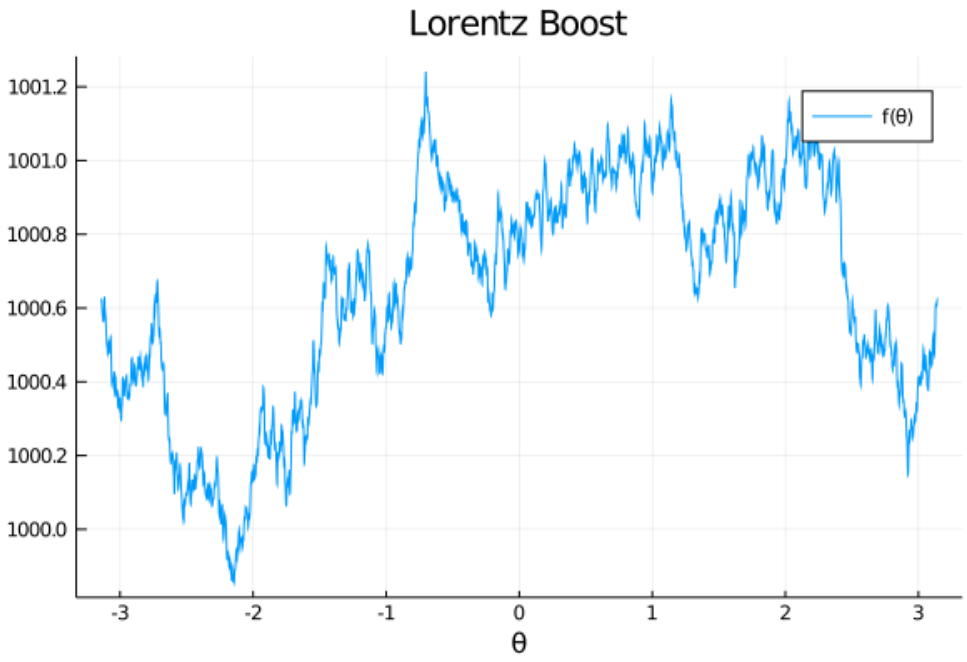


FIGURE 3.1. The initial state of the function f on $\mathbb{S}$ before the application of the Lorentz Boost.

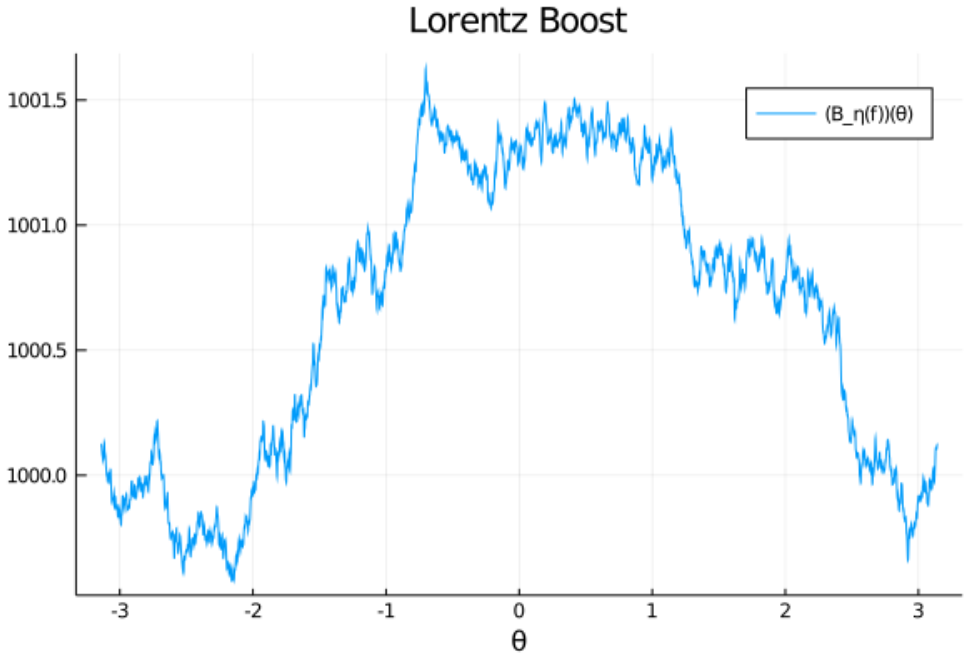


FIGURE 3.2. The function f on $\mathbb{S}$ following the application of the Lorentz Boost, illustrating the induced alterations

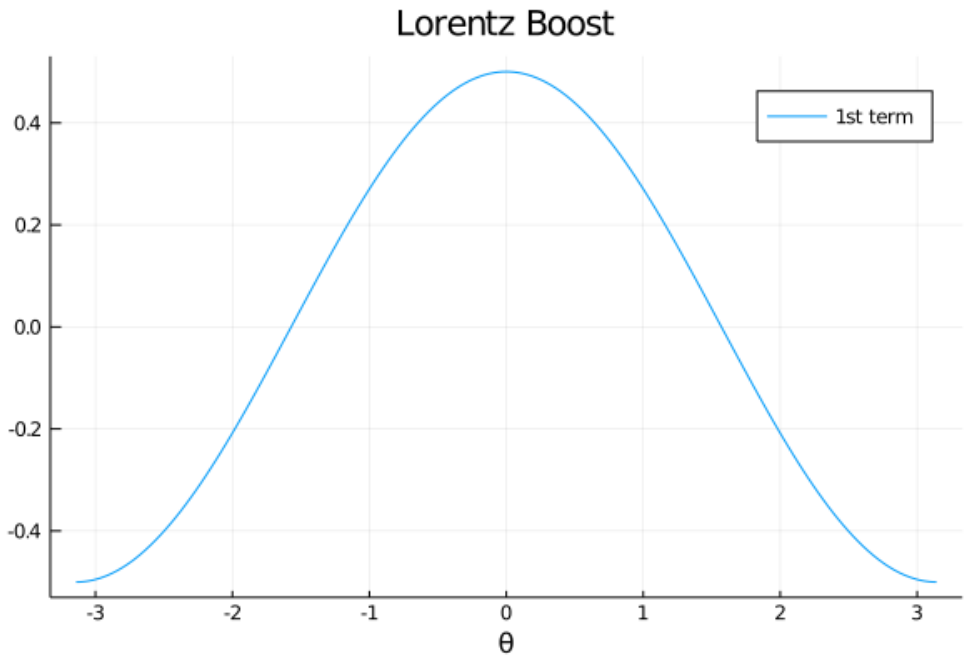


FIGURE 3.3. Illustration of the first term $\eta r \cos(\theta - \phi) f$ from the Lorentz Boost approximation given in Equation 3.4, highlighting its dipole-like structure.

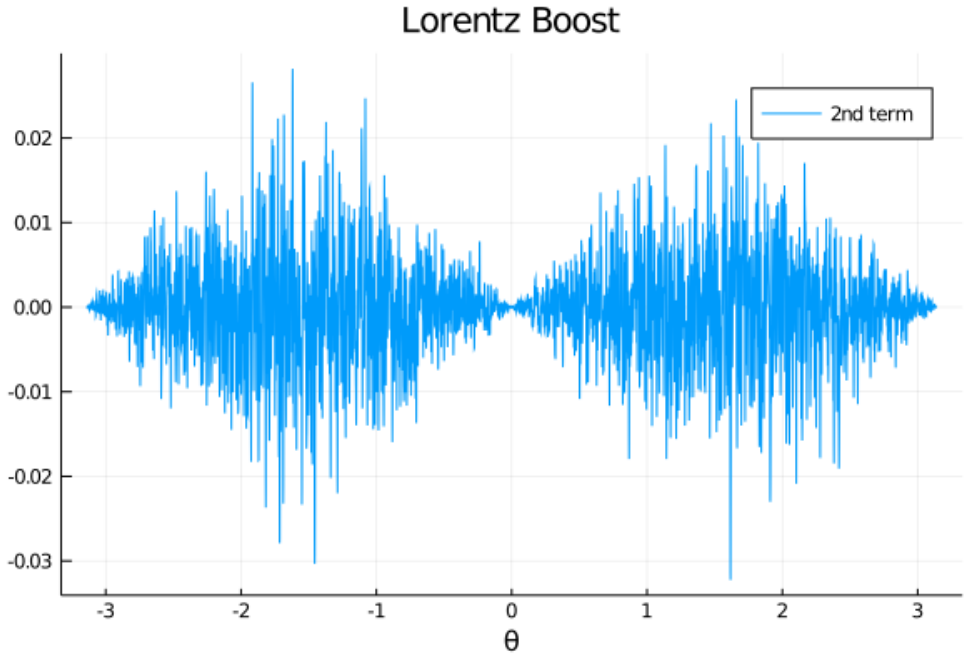


FIGURE 3.4. Illustration of the second term $\eta \sin(\theta - \phi) \frac{d}{d\theta} f$ from the Lorentz Boost approximation in Equation 3.4, emphasizing the high-frequency components akin to acoustic beats.

CHAPTER 4

Quadratic Estimator

4.1. Introduction

This chapter focuses on Quadratic Estimators, particularly as they apply to the CMB in the field of physical cosmology.

Initially, we review the established methodology for deriving the Quadratic Estimator of peculiar velocity as a least-squares minimization problem, as presented in works such as [**ACM$^+$11**]. This approach can also be interpreted as a Method of Moments, where the estimator is obtained by matching the first moment of the distribution.

Later in the chapter, we explore Quadratic Estimators from a novel perspective that integrates mathematical statistics, differential geometry, and numerical optimization. This approach enhances our understanding of Quadratic Estimators. Specifically, we frame the Quadratic Estimator as a one-step Newton method solution to a maximum likelihood problem. To facilitate this, we derive explicit formulas for the likelihood functions related to the estimation of peculiar velocity in the CMB. In Chapter 3, we have already discussed the general operation of Lorentz Boost on functions defined on a manifold, given an arbitrary smooth vector field. The peculiar velocity effect on the CMB serves as a special case where the manifold is a sphere and the vector field is defined as in Equation 3.4.6.1. Our formulation yields explicit formulas for both the gradient and the Hessian of the likelihood function with respect to the parameters of interest.

4.2. Quadratic Estimator of a Least Square Problem

We begin by reviewing the conventional approach to Quadratic Estimators. Let $\mathbf{x}$ be a random variable with zero mean and a given covariance matrix:

$$\mathrm{E}[\mathbf{x}\mathbf{x}^T] = \mathbf{C} + \lambda\mathbf{P},$$

where $\mathbf{C}$ and $\mathbf{P}$ are known square matrices, and λ is the scalar parameter of interest. Given observed data $\mathbf{x}$, our goal is to estimate λ. The estimation problem can be framed as a least-squares problem that aims to minimize the χ^2 functional:

$$\chi^2(\lambda) := \sum_{ij} \frac{(x_i x_j - (C_{ij} + \lambda P_{ij}))^2}{\sigma_{ij}^2},$$

Here, λ is the unknown parameter of interest, and $x_i x_j$ represents the sample correlation of the observed data $\mathbf{x}$, C_{ij} and P_{ij} are the (i,j)-th entries of matrices $\mathbf{C}$ and $\mathbf{P}$ respectively, and σ_{ij}^2 is the assumed variance of the random quantity $x_i x_j$. If σ_{ij}^2 is assumed to be 1, the minimizer of $\chi^2(\lambda)$ essentially becomes a Method of Moments solution. In this case, we find λ such that $C_{ij} + \lambda P_{ij}$ matches $x_i x_j$. For general values of σ_{ij}^2, the minimizer of $\chi^2(\lambda)$ solves a weighted least-squares problem.

4.3. Quadratic Estimator as Fisher Scoring Algorithm

Quadratic Estimators are used to estimate parameters in the covariance matrix of the data. Suppose the data is represented by $\mathbf{x}$ with a model $P(\mathbf{x}|\lambda)$; the parameter of interest is λ. We may construct a Quadratic Estimator for λ that is required to have the form

$$\hat{\lambda} = \mathbf{x}^T \mathbf{A} \mathbf{x} + b \tag{4.1}$$

where $\mathbf{A}$ is a symmetric matrix and b is a scalar. Consider the case that $\mathbf{x}$ has multivariate Gaussain distribution $N(\boldsymbol{\mu}, \mathbf{V})$ for a mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{V}$. For any symmetric matrix $\mathbf{A}_1$ and $\mathbf{A}_2$,

$$\mathrm{E}[\mathbf{x}^T \mathbf{A} \mathbf{x}] = \mathrm{tr}(\mathbf{A}\mathbf{V}) + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}$$

$$\mathrm{Cov}(\mathbf{x}^T \mathbf{A}_1 \mathbf{x}, \mathbf{x}^T \mathbf{A}_2 \mathbf{x}) = 2\mathrm{tr}(\mathbf{A}_1 \mathbf{V} \mathbf{A}_2 \mathbf{V}) + 4 \boldsymbol{\mu}^T \mathbf{A}_1 \mathbf{V} \mathbf{A}_2 \boldsymbol{\mu}.$$

More generally, when the parameter of interest is a vector $\boldsymbol{\lambda}$, the Quadratic Estimator of each component λ_i is a linear combination of $\sum_j (\mathbf{x}^T \mathbf{A}_j \mathbf{x} + b_j)$ for some matrices $\mathbf{A}_j$ and scalars b_j.

We follow the formulation of Quadratic Estimator in [**HL09**]. Suppose $\boldsymbol{\mu} = 0$ and $\mathbf{V}$ has a linear form in the parameters λ_k

$$\mathbf{V}(\boldsymbol{\lambda}) = \mathbf{C} + \sum_k \lambda_k \mathbf{P}_k \tag{4.2}$$

where $\boldsymbol{\lambda} = (\lambda_k)_k$ are the parameters of interests and $\mathbf{C}$ and $\mathbf{P}_k$ are known matrices. For example, in [**Teg97**], the CMB temperature data $\mathbf{x}$ is modeled as having a multivariate Gaussian distribution with covariance matrix $\mathrm{E}[\mathbf{x}\mathbf{x}^T] = \mathbf{V}(\mathbf{c}) = \mathbf{N} + \sum_{\ell \geq 0} C_\ell \mathbf{P}^\ell$ where $\mathbf{c} = (C_\ell)_\ell$ is the vector of spectral coefficients and $\mathbf{P}^\ell_{ij} = \frac{2\ell+1}{4\pi} P_\ell(\hat{\mathbf{r}}_i \cdot \hat{\mathbf{r}}_j)$ and $\hat{\mathbf{r}}_i$ is a unit vector pointing in the direction of pixel i and $\mathbf{N}$ is a known diagonal matrix that models the measurement noise. Here $\mathbf{c}$ is the parameter of interests.

We aim to estimate $\boldsymbol{\lambda}$ by constructing a Quadratic Estimator for $\boldsymbol{\lambda}$. In other words, we need to design the matrix $\mathbf{A}$ and scalar b in the Quadratic Estimator defined in equation 4.1 that helps us to find $\boldsymbol{\lambda}$. According to [**HL09**], we take

$$\begin{aligned} \mathbf{A}_k &= \frac{1}{2}\mathbf{C}^{-1}\mathbf{P}_k\mathbf{C}^{-1} \\ b_k &= \frac{1}{2}\mathrm{tr}\mathbf{P}_k\mathbf{C}^{-1} \end{aligned}$$

and define the estimator $\tilde{\lambda}_k$

$$\tilde{\lambda}_k = \mathbf{x}^T\mathbf{A}_k\mathbf{x} - b_k \tag{4.3}$$

By direct calculations with the trace trick, the expectation of $\tilde{\lambda}_k$ over $\mathbf{x}$ is

$$\begin{aligned} \mathrm{E}[\tilde{\lambda}_k] &= \mathrm{E}[\mathbf{x}^T\mathbf{A}_k\mathbf{x}] - b_k \\ &= \frac{1}{2}\mathrm{tr}\mathrm{E}[\mathbf{x}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{P}_k\mathbf{C}^{-1}] - \frac{1}{2}\mathrm{tr}\mathbf{P}_k\mathbf{C}^{-1} \\ &= \frac{1}{2}\mathrm{tr}[\mathbf{C}\mathbf{C}^{-1}\mathbf{P}_k\mathbf{C}^{-1} + \sum_{k'}\lambda_{k'}\mathbf{P}_{k'}\mathbf{C}^{-1}\mathbf{P}_k\mathbf{C}^{-1}] - \frac{1}{2}\mathrm{tr}\mathbf{P}_k\mathbf{C}^{-1} \\ &= \frac{1}{2}\sum_{k'}\lambda_{k'}\mathrm{tr}[\mathbf{C}^{-1}\mathbf{P}_k\mathbf{C}^{-1}\mathbf{P}_{k'}] \\ &= [\mathbf{F}\boldsymbol{\lambda}]_k \end{aligned}$$

where $\mathbf{F} = (F_{kk'})$ is a matrix with entries

$$F_{kk'} = \frac{1}{2}\mathrm{tr}[\mathbf{C}^{-1}\mathbf{P}_k\mathbf{C}^{-1}\mathbf{P}_{k'}]. \tag{4.4}$$

We would like to solve the equation for $\boldsymbol{\lambda}$ by inverting the matrix $\mathbf{F}$. We define the unbiased estimator $\hat{\boldsymbol{\lambda}} = (\hat{\lambda}_k)_k$ to be

$$\hat{\boldsymbol{\lambda}} = \mathbf{F}^{-1}\tilde{\boldsymbol{\lambda}}. \tag{4.5}$$

so that $\mathrm{E}[\hat{\boldsymbol{\lambda}}] = \boldsymbol{\lambda}$. Beside the form of the covariance matrix, there is no assumption on the distribution of data $\mathbf{x}$.

The estimator $\hat{\boldsymbol{\lambda}}$ defined in equation 4.5 looks complicated at first but it is magically simplified once expectation is taken. Indeed the formula can be obtained from a one-step Fisher Scoring Algorithm on the log likelihood function of Gaussian distribution $N(\mathbf{0}, \mathbf{V}(\boldsymbol{\lambda}))$ where $\mathbf{V}$ is a covariance matrix parameterized with λ. Let $f(\mathbf{x}; \boldsymbol{\lambda})$ be the density function of the data $\mathbf{x}$. Suppose f is the density function of Gaussian with zero mean and covariance matrix $V(\boldsymbol{\lambda})$ which depends on the parameter $\boldsymbol{\lambda}$,

$$\log f(\mathbf{x}; \boldsymbol{\lambda}) := -\frac{1}{2}\mathbf{x}^T\mathbf{V}^{-1}(\boldsymbol{\lambda})\mathbf{x} - \frac{1}{2}\log\det\mathbf{V}(\boldsymbol{\lambda}).$$

and the gradient of $\log f$, denoated by $G(\boldsymbol{\lambda})$, and Fisher information matrix of f, denoted by $I(\boldsymbol{\lambda})$, are respectively

$$[G(\boldsymbol{\lambda})]_i := \frac{\partial}{\partial\lambda_i}\log f = \frac{1}{2}\mathbf{x}^T\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\lambda_i}\mathbf{V}^{-1}\mathbf{x} - \frac{1}{2}\mathrm{tr}\left[\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\lambda_i}\right]$$

and

$$[I(\boldsymbol{\lambda})]_{ij} := -\mathrm{E}\left[\frac{\partial^2}{\partial\lambda_i\partial\lambda_j}\log f\middle|\boldsymbol{\lambda}\right] = \frac{1}{2}\mathrm{tr}\left[\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\lambda_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\lambda_j}\right]$$

where the matrices $\mathbf{V}$ and $\frac{\partial\mathbf{V}}{\partial\lambda_i}$ are evaluated at $\boldsymbol{\lambda}$. The Fisher Scoring Algorithm, given $\boldsymbol{\lambda}_n$, is the update equation:

$$\boldsymbol{\lambda}_{n+1} = \boldsymbol{\lambda}_n + I(\boldsymbol{\lambda}_n)^{-1}G(\boldsymbol{\lambda}_n). \tag{4.6}$$

The covariance matrix of $\hat{\boldsymbol{\lambda}}$ is

$$\mathrm{Var}[\hat{\boldsymbol{\lambda}}] = \mathbf{F}^{-1} + O(\|\boldsymbol{\lambda}\|).$$

We claim that the estimator defined in equation 4.5 can be obtained by taking $\boldsymbol{\lambda}_1$ as the estimate with $\boldsymbol{\lambda} = 0$. Starting at $\boldsymbol{\lambda}_0 = \mathbf{0}$, the one-step Fisher Scoring Algorithm is

$$\begin{aligned}\boldsymbol{\lambda}_1 &= \boldsymbol{\lambda}_0 + I(\boldsymbol{\lambda}_0)^{-1} G(\boldsymbol{\lambda}_0) \\ &= I(\mathbf{0})^{-1} G(\mathbf{0})\end{aligned}$$

When $\mathbf{V}$ has a linear form $\mathbf{V}(\boldsymbol{\lambda}) = \mathbf{C} + \sum_k \lambda_k \mathbf{P}_k$ where $\mathbf{C}$ and $\mathbf{P}_k$'s are known matrices and λ_k's are parameters of interests, we can write down explicit formulas for $G(\mathbf{0})$ and $I(\mathbf{0})$. We have explicit form of $\mathbf{V}(\mathbf{0})$ and $\left.\frac{\partial \mathbf{V}}{\partial \lambda_k}\right|_{\boldsymbol{\lambda}=\mathbf{0}}$

$$\begin{aligned}\mathbf{V}(\mathbf{0}) &= \mathbf{C} \\ \left.\frac{\partial \mathbf{V}}{\partial \lambda_k}\right|_{\boldsymbol{\lambda}=\mathbf{0}} &= \mathbf{P}_k\end{aligned}$$

and thus

$$\begin{aligned}[G(\mathbf{0})]_k &= \frac{1}{2}\mathbf{x}^T \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \lambda_k} \mathbf{V}^{-1} \mathbf{x} - \frac{1}{2}\mathrm{tr}\left[\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \lambda_k}\right] \\ &= \frac{1}{2}\mathbf{x}^T \mathbf{C}^{-1} \mathbf{P}_k \mathbf{C}^{-1} \mathbf{x} - \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}\mathbf{P}_k\right] \\ &= \mathbf{x}^T \mathbf{A}_k \mathbf{x} - b_k\end{aligned}$$

and

$$\begin{aligned}[I(\mathbf{0})]_{ij} &= \frac{1}{2}\mathrm{tr}\left[\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \lambda_i} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \lambda_j}\right] \\ &= \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}\mathbf{P}_i \mathbf{C}^{-1}\mathbf{P}_j\right] \\ &= \mathbf{F}_{ij}.\end{aligned}$$

Therefore we have shown that $G(\mathbf{0})$ is indeed $\tilde{\boldsymbol{\lambda}}$ defined in equation 4.3 and $I(\mathbf{0})$ is nothing but $\mathbf{F}$ defined in equation 4.4.

In general, the covariance matrix may not have the linear form of the equation 4.2. However, when the covariance matrix is well-approximated by the linear form, i.e.

$$\mathbf{V}(\boldsymbol{\lambda}) \approx \mathbf{V}(\mathbf{0}) + \sum_k \lambda_k \left.\frac{\partial \mathbf{V}}{\partial \lambda_k}\right|_{\boldsymbol{\lambda}=\mathbf{0}},$$

the Quadratic Estimator is approximately unbiased for small magnitude $\|\boldsymbol{\lambda}\| \approx 0$. If the model has Gaussian distribution, the Quadratic Estimator is a one-step Fisher scoring algorithm which provides a good approximation of the maximum likelihood estimator when $\|\boldsymbol{\lambda}\|$ is small.

4.4. Quadratic Estimator for peculiar velocity CMB

We write down the likelihood function of the model of CMB under peculiar velocity. Denote the data of observed CMB spherical harmonic coefficients by $\mathbf{a}^{\text{obs}}$ and the primodial CMB spherical harmonic coefficients by $\mathbf{a}$. Let $\boldsymbol{\eta} = (\eta_i)$ be the rapidity of peculiar velocity. The conditional distribution of $\mathbf{a}^{\text{obs}}$ given $\mathbf{a}$ and distribution of $\mathbf{a}$ are Gaussian distributions:

$$\mathbf{a}^{\text{obs}}|\mathbf{a} \sim N\left(\exp\left(\sum_{i=1}^{d} \eta_i \mathbf{K}_i\right)\mathbf{a}, \boldsymbol{\Sigma}_{\text{noise}}\right)$$

$$\mathbf{a} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\text{pri}})$$

where $\boldsymbol{\Sigma}_{\text{pri}}$ and $\boldsymbol{\Sigma}_{\text{noise}}$ are known covariance matrices. $\mathbf{K}_i$ are known sparse anti-symmetric matrices that are derived from the differential equations of Flow. The matrix exponential $\exp\left(\sum_{i=1}^{d} \eta_i \mathbf{K}_i\right)$ is a linear operation on the vector of primodial CMB spherical harmonic coefficients by $\mathbf{a}$ and is parameterized by rapidity $\boldsymbol{\eta}$. The marginal distribution of $\mathbf{a}^{\text{obs}}$ is

$$\mathbf{a}^{\text{obs}} \sim N\left(\mathbf{0}, \exp\left(\sum_{i=1}^{d} \eta_i \mathbf{K}_i\right)\boldsymbol{\Sigma}_{\text{pri}} \exp\left(-\sum_{i=1}^{d} \eta_i \mathbf{K}_i\right) + \boldsymbol{\Sigma}_{\text{noise}}\right).$$

4.4.0.1. *Least Square Formulation.* We firstly show how to construct the Quadratic Estimator with formulation of the least square problem. In first order approximation of matrix exponential, $\mathbf{a}^{\text{obs}}$ can be approximated as

$$\mathbf{a}^{\text{obs}} = \left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{a} + \boldsymbol{\epsilon} + O(\|\boldsymbol{\eta}\|^2)$$

where $\mathbf{a} \sim N(\mathbf{0}, \mathbf{\Sigma}_{\text{pri}})$ and $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{\Sigma}_{\text{noise}})$. Expand the product $\mathbf{a}^{\text{obs}}\mathbf{a}^{\text{obs}T}$, we obtain

$$\begin{aligned}
\mathbf{a}^{\text{obs}}\mathbf{a}^{\text{obs}T} &= \left(\left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{a} + \boldsymbol{\epsilon}\right)\left(\left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{a} + \boldsymbol{\epsilon}\right)^T + O(\|\boldsymbol{\eta}\|^2) \\
&= \left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{a}\mathbf{a}^T\left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)^T \\
&\quad + \left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{a}\boldsymbol{\epsilon}^T + \boldsymbol{\epsilon}\mathbf{a}^T\left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)^T + \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T + O(\|\boldsymbol{\eta}\|^2)
\end{aligned}$$

Since $E[\mathbf{a}\mathbf{a}^T] = \mathbf{\Sigma}_{\text{pri}}$ and $E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T] = \mathbf{\Sigma}_{\text{noise}}$, we can obtain an expression of $E[\mathbf{a}^{\text{obs}}\mathbf{a}^{\text{obs}T}]$:

$$\begin{aligned}
E[\mathbf{a}^{\text{obs}}\mathbf{a}^{\text{obs}T}] &= \left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{\Sigma}_{\text{pri}}\left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)^T + \mathbf{\Sigma}_{\text{noise}} + O(\|\boldsymbol{\eta}\|^2) \\
&= \left(\mathbf{I} + \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{\Sigma}_{\text{pri}}\left(\mathbf{I} - \sum_{i=1}^{k} \eta_i \mathbf{K}_i\right) + \mathbf{\Sigma}_{\text{noise}} + O(\|\boldsymbol{\eta}\|^2) \\
&= \mathbf{\Sigma}_{\text{pri}} - \mathbf{\Sigma}_{\text{pri}}\left(\sum_{i=1}^{k} \eta_i \mathbf{K}_i\right) + \left(\sum_{i=1}^{k} \eta_i \mathbf{K}_i\right)\mathbf{\Sigma}_{\text{pri}} + \mathbf{\Sigma}_{\text{noise}} + O(\|\boldsymbol{\eta}\|^2) \\
&= \mathbf{\Sigma}_{\text{pri}} + \sum_{i=1}^{d} \eta_i \left(\mathbf{K}_i\mathbf{\Sigma}_{\text{pri}} - \mathbf{\Sigma}_{\text{pri}}\mathbf{K}_i\right) + \mathbf{\Sigma}_{\text{noise}} + O(\|\boldsymbol{\eta}\|_2^2).
\end{aligned}$$

The above expression provides a very good insight on the covariance structure of the observed CMB. The matrices $\mathbf{K}_i$'s are anti-symmetric. In the real Fourier basis, the diagonal elements of $\mathbf{K}_i$'s are zeros. $\mathbf{K}_i$'s are orthogonal, in the sense that $\text{tr}(\mathbf{K}_i^T\mathbf{K}_j) = 0$ if $i \neq j$). The matrix $\mathbf{\Sigma}_{\text{pri}}$ is diagonal. The matrices

$$\mathbf{K}_i\mathbf{\Sigma}_{\text{pri}} - \mathbf{\Sigma}_{\text{pri}}\mathbf{K}_i \quad \text{for } i = 1, \ldots, d$$

are symmetric and "orthogonal" to each other. Thus the sum

$$\mathbf{\Sigma}_{\text{pri}} + \mathbf{\Sigma}_{\text{noise}} + \sum_{i=1}^{d} \eta_i \left(\mathbf{K}_i\mathbf{\Sigma}_{\text{pri}} - \mathbf{\Sigma}_{\text{pri}}\mathbf{K}_i\right)$$

results in a matrix that

- the diagonal entry is $\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k}$ where $\Sigma_{\mathrm{pri},k}$ and $\Sigma_{\mathrm{obs},k}$ are (k,k) diagonal entry of $\mathbf{\Sigma}_{\mathrm{pri}}$ and $\mathbf{\Sigma}_{\mathrm{noise}}$ respectively.
- the off-diagonal entry is non-zero only if $||k| - |k'|| = 1$.
- For each off-diagonal non-zero entry, it is only linearly depending on one η_i.

In other words, for a_k being elements in vector $\mathbf{a}$, in first order approximation with respect to

- $\mathrm{E}((a_k^{\mathrm{obs}})^2) = \Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k}$
- $\mathrm{E}(a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}})$ is non-zero only if $||k| - |k'|| = 1$
- $\mathrm{E}(a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}})$, if non-zero, depends only either one parameter on η_1, η_2 or η_3 linearly.

We explicitly write down that $k \neq k'$

$$\mathrm{E}(a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}}) = \sum_{i=1}^{d} \eta_i [\mathbf{K}_i]_{kk'} (\Sigma_{\mathrm{pri},k'} - \Sigma_{\mathrm{pri},k}) + O(\|\boldsymbol{\eta}\|_2^2)$$

and that

$$\mathrm{Var}(a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}}) = (\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k})(\Sigma_{\mathrm{pri},k'} + \Sigma_{\mathrm{obs},k'}) + O(\|\boldsymbol{\eta}\|_2)$$

With the mean and variance approximated as above, we define a surrogate negative log-likelihood function, also known as the χ^2 functional, as:

$$\chi^2(\boldsymbol{\eta}, \mathbf{a}^{\mathrm{obs}}) = \sum_{k<k' \text{ and } [\mathbf{K}_i]_{kk'} \neq 0} \frac{(a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}} - \sum_{i=1}^{d} \eta_i [\mathbf{K}_i]_{kk'} (\Sigma_{\mathrm{pri},k'} - \Sigma_{\mathrm{pri},k}))^2}{(\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k})(\Sigma_{\mathrm{pri},k'} + \Sigma_{\mathrm{obs},k'})}$$

The minimizer of $\chi^2(\boldsymbol{\eta}, \mathbf{a}^{\mathrm{obs}})$ given $\mathbf{a}^{\mathrm{obs}}$

$$\hat{\boldsymbol{\eta}}_{\mathrm{QE}} = \arg\min_{\boldsymbol{\eta}} \chi^2(\boldsymbol{\eta}, \mathbf{a}^{\mathrm{obs}})$$

is called Quadratic Estimator of peculiar velocity.

We can enumerate the index $\{n\} = \{(k,k') : k \neq k'\}$ in certain order. Let $\mathbf{y}$ be the vector of $y_n = a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}}$. Let $\mathbf{X}$ be the matrix that $X_{ni} = [\mathbf{K}_i]_{kk'} (\Sigma_{\mathrm{pri},k'} - \Sigma_{\mathrm{pri},k})$. Let $\mathbf{W}$ be the vector that $W_{nn} = (\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k})(\Sigma_{\mathrm{pri},k'} + \Sigma_{\mathrm{obs},k'})$. The minimizer of $\chi^2(\boldsymbol{\eta}, \mathbf{a}^{\mathrm{obs}})$ is therefore

$$\hat{\boldsymbol{\eta}}_{\mathrm{QE}} = (\mathbf{X}^T \mathbf{W}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{-1} \mathbf{y}.$$

The variance of the Quadratic Estimator $\hat{\boldsymbol{\eta}}_{\mathrm{QE}}$ is

$$\mathrm{Var}(\hat{\boldsymbol{\eta}}_{\mathrm{QE}}) = (\mathbf{X}^T\mathbf{W}^{-1}\mathbf{X})^{-1}.$$

As $\mathbf{K}_i$'s are orthogonal in the sense that $\mathrm{tr}(\mathbf{K}_i^T\mathbf{K}_j)$ for $i \neq j$, we have that $([\mathbf{K}_i]_{kk'} : i = 1, \ldots, d)$ has only single or no non-zero entry for any (k, k'). In other words, $\mathbf{X}$ is an orthogonal matrix. The surrogate likelihood can be separate for each i:

$$\chi_i^2(\eta_i, \mathbf{a}^{\mathrm{obs}}) = \sum_{k<k' \text{ and } [\mathbf{K}_i]_{kk'}\neq 0} \frac{(a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}} - \eta_i[\mathbf{K}_i]_{kk'}(\Sigma_{\mathrm{pri},k'} - \Sigma_{\mathrm{pri},k}))^2}{(\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k})(\Sigma_{\mathrm{pri},k'} + \Sigma_{\mathrm{obs},k'})}$$

and the minimizer is

$$[\hat{\eta}_i]_{\mathrm{QE}} = \frac{\sum_{(k,k')\in S_i} \frac{[\mathbf{K}_i]_{kk'}(\Sigma_{\mathrm{pri},k'} - \Sigma_{\mathrm{pri},k})}{(\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k})(\Sigma_{\mathrm{pri},k'} + \Sigma_{\mathrm{obs},k'})} a_k^{\mathrm{obs}} a_{k'}^{\mathrm{obs}}}{\sum_{(k,k')\in S_i} \frac{([\mathbf{K}_i]_{kk'}(\Sigma_{\mathrm{pri},k'} - \Sigma_{\mathrm{pri},k}))^2}{(\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k})(\Sigma_{\mathrm{pri},k'} + \Sigma_{\mathrm{obs},k'})}}$$

where $S_i := \{(k, k') : k < k' \text{ and } [\mathbf{K}_i]_{kk'} \neq 0\}$ is the index set for summing up suitable (k, k'). $\mathbf{X}^T\mathbf{W}\mathbf{X}$ is d-by-d matrix diagonal. By symmetry of $\{\mathbf{K}_i\}$, $\mathbf{X}^T\mathbf{W}\mathbf{X}$ has same value in the diagonal entry. Thus the diagonal matrix values is just the reciprocal of the variance. We define the signal-to-noise ratio to be the reciprocal of the variance:

$$\mathrm{SNR}_i := \sum_{(k,k')\in S_i} \frac{([\mathbf{K}_i]_{kk'}(\Sigma_{\mathrm{pri},k'} - \Sigma_{\mathrm{pri},k}))^2}{(\Sigma_{\mathrm{pri},k} + \Sigma_{\mathrm{obs},k})(\Sigma_{\mathrm{pri},k'} + \Sigma_{\mathrm{obs},k'})}$$

Note that, by symmetry, the value of SNR_i is independent of i. We can refer the value as SNR. The standard error of $\hat{\eta}_i$ depends on the profile of $\boldsymbol{\Sigma}_{\mathrm{pri}}$ and $\boldsymbol{\Sigma}_{\mathrm{noise}}$. The approximated standard error of $\hat{\eta}_i$ is

$$\sqrt{\mathrm{Var}(\eta_i)} \approx \frac{1}{\sqrt{\mathrm{SNR}}}.$$

We show that the standard error of Quadratic Estimator is approximately the reciprocal of Fisher information of the model, which is the Cramér–Rao bound. The Quadratic Estimator is approximately unbiased and its standard error is approximately the inverse of Fisher information. Therefore, the performance of Quadratic Estimator is expected to be similar to that of the maximum likelihood estimator.

THEOREM 4.4.1. The Fisher information matrix $\mathcal{I}(\boldsymbol{\eta})$ of the Gaussian model

$$f(\mathbf{x};\boldsymbol{\eta},\boldsymbol{\Sigma}) = (2\pi)^{-1}|\boldsymbol{\Sigma}|^{-1/2} \times \exp\left(-\frac{1}{2}\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)^T(\boldsymbol{\Sigma})^{-1}\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)$$

is

$$[\mathcal{I}(\boldsymbol{\eta}=0)]_{ij} = -\frac{1}{2}\mathrm{tr}\left(\left[\mathbf{K}_i,\left[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}\right]\right]\boldsymbol{\Sigma}\right)$$

Further,

$$\begin{aligned}[\mathcal{I}(\boldsymbol{\eta})]_{ij} &= -\frac{1}{2}\mathrm{tr}\left(\left[\mathbf{K}_i,\left[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}\right]\right]\boldsymbol{\Sigma}\right)\\ &= \mathrm{tr}\left(\mathbf{K}_i\boldsymbol{\Sigma}\mathbf{K}_j\boldsymbol{\Sigma}^{-1}\right) - \mathrm{tr}\left(\mathbf{K}_i\mathbf{K}_j\right)\\ &= \begin{cases}\sum_{k<k'}[\mathbf{K}_i]^2_{kk'}\,\frac{([\boldsymbol{\Sigma}_i]_k-[\boldsymbol{\Sigma}_i]_{k'})^2}{[\boldsymbol{\Sigma}]_k[\boldsymbol{\Sigma}]_{k'}} & \text{if } i=j\\ 0 & \text{if } i\neq j\end{cases}\end{aligned}$$

which is the same formula for (approximated) standard error of Quadratic Estimator.

PROOF. Direct calculation shows that

$$\begin{aligned}\mathrm{Var}\left[\mathbf{x}\middle|\boldsymbol{\eta}\right] &= \exp\left(\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\boldsymbol{\Sigma}\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\\ \mathrm{E}\left[\mathbf{x}\left(\mathbf{x}\right)^T\middle|\boldsymbol{\eta}\right] &= \exp\left(\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\boldsymbol{\Sigma}\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\end{aligned}$$

hence

$$\mathrm{E}\left[\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)^T\middle|\boldsymbol{\eta}\right] = \boldsymbol{\Sigma}$$

Therefore, the fisher information can be simplified as

$$
\begin{aligned}
[\mathcal{I}(\boldsymbol{\eta})]_{ij} &= -\mathrm{E}\left[\frac{\partial^2}{\partial\eta_i\partial\eta_j}\log f\left(\mathbf{x};\boldsymbol{\eta}\right)\middle|\boldsymbol{\eta}=\mathbf{0}\right] \\
&= -\frac{1}{2}\mathrm{E}\left[\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)^T[\mathbf{K}_i,[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}]]\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)\middle|\boldsymbol{\eta}=\mathbf{0}\right] \\
&= -\frac{1}{2}\mathrm{E}\left[\mathrm{tr}\left[[\mathbf{K}_i,[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}]]\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)^T\right]\middle|\boldsymbol{\eta}=\mathbf{0}\right] \\
&= -\frac{1}{2}\mathrm{tr}\left[[\mathbf{K}_i,[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}]]\mathrm{E}\left[\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)\left(\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\mathbf{x}\right)^T\middle|\boldsymbol{\eta}=\mathbf{0}\right]\right] \\
&= -\frac{1}{2}\mathrm{tr}\left([\mathbf{K}_i,[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}]]\boldsymbol{\Sigma}\right).
\end{aligned}
$$

which does not involve $\boldsymbol{\eta}$. Further, we expand $[\mathbf{K}_i,[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}]]\boldsymbol{\Sigma}$ and get

$$
\begin{aligned}
&-\frac{1}{2}\mathrm{tr}\left([\mathbf{K}_i,[\mathbf{K}_j,\boldsymbol{\Sigma}^{-1}]]\boldsymbol{\Sigma}\right) \\
=&-\frac{1}{2}\mathrm{tr}\left(\mathbf{K}_i\mathbf{K}_j+\boldsymbol{\Sigma}^{-1}\mathbf{K}_j\mathbf{K}_i\boldsymbol{\Sigma}-\mathbf{K}_i\boldsymbol{\Sigma}^{-1}\mathbf{K}_j\boldsymbol{\Sigma}-\mathbf{K}_j\boldsymbol{\Sigma}^{-1}\mathbf{K}_i\boldsymbol{\Sigma}\right) \\
=&-\frac{1}{2}\mathrm{tr}\left(\mathbf{K}_i\mathbf{K}_j\right)-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{K}_j\mathbf{K}_i\boldsymbol{\Sigma}\right)+\frac{1}{2}\mathrm{tr}\left(\mathbf{K}_i\boldsymbol{\Sigma}^{-1}\mathbf{K}_j\boldsymbol{\Sigma}\right)+\frac{1}{2}\mathrm{tr}\left(\mathbf{K}_j\boldsymbol{\Sigma}^{-1}\mathbf{K}_i\boldsymbol{\Sigma}\right) \\
=&-\mathrm{tr}\left(\mathbf{K}_i\mathbf{K}_j\right)+\mathrm{tr}\left(\mathbf{K}_i\boldsymbol{\Sigma}^{-1}\mathbf{K}_j\boldsymbol{\Sigma}\right) \\
=&\mathrm{tr}\left(\mathbf{K}_i^T\mathbf{K}_j\right)-\mathrm{tr}\left((\boldsymbol{\Sigma}^{-1}\mathbf{K}_i)^T\mathbf{K}_j\boldsymbol{\Sigma}\right) \\
=&\langle\mathbf{K}_i,\mathbf{K}_j\rangle_{\mathrm{F}}-\langle\boldsymbol{\Sigma}^{-1}\mathbf{K}_i,\mathbf{K}_j\boldsymbol{\Sigma}\rangle_{\mathrm{F}}
\end{aligned}
$$

where $\langle\cdot,\cdot\rangle_{\mathrm{F}}$ is the Frobenius inner product. If $i\neq j$, then both terms are zero. If $i=j$, then

$$
\begin{aligned}
&=-2\sum_{k<k'}[\mathbf{K}_i]_{kk'}^2+\left(\sum_{k<k'}[\mathbf{K}_i]_{kk'}^2\frac{[\boldsymbol{\Sigma}]_k}{[\boldsymbol{\Sigma}]_{k'}}+\sum_{k<k'}[\mathbf{K}_i]_{kk'}^2\frac{[\boldsymbol{\Sigma}]_{k'}}{[\boldsymbol{\Sigma}]_k}\right) \\
&=\sum_{k<k'}[\mathbf{K}_i]_{kk'}^2\frac{\left([\boldsymbol{\Sigma}_i]_k-[\boldsymbol{\Sigma}_i]_{k'}\right)^2}{[\boldsymbol{\Sigma}]_k[\boldsymbol{\Sigma}]_{k'}}.
\end{aligned}
$$

□

More specifically for the CMB model on sphere, the observed spherical harmonic $\mathbf{a}^{\text{obs}} = (a_{\ell m}^{\text{obs}})$ has distribution

$$a_{\ell m}^{\text{obs}} | \mathbf{a} \sim N\left(\exp\left(\sum_{i=1}^{3} \eta_i \mathbf{K}_i\right), N_\ell\right)$$
$$a_{\ell m} \sim N(0, C_\ell).$$

To make formula simpler, we define the quantities

$$v_{i,\ell\ell' mm'} := [\mathbf{K}_i]_{\ell\ell' mm'}(C_\ell - C_{\ell'})$$
$$w_{\ell\ell' mm'} := (C_\ell + N_\ell)(C_{\ell'} + N_{\ell'})$$
$$W_\ell := \sum_m \frac{v_{3,\ell(\ell+1)mm}^2}{w_{\ell(\ell+1)mm}}. \tag{4.7}$$

The Quadratic Estimator of rapidity $\boldsymbol{\eta}$ is

$$[\hat{\eta}_i]_{\text{QE}} := \frac{\sum_{\ell<\ell'} \sum_{mm'} \frac{v_{i,\ell\ell' mm'}}{w_{\ell\ell' mm'}} a_{\ell m} a_{\ell' m'}}{\sum_{\ell<\ell'} W_\ell}.$$

Note that $\hat{\eta}_i$ is quadratic in $a_{\ell m} a_{\ell' m'}$. In [**ACM+11**], the signal-to-noise ratio is defined as

$$\text{SNR} := \sum_\ell \sum_m \frac{v_{3,\ell(\ell+1)mm}^2}{w_{\ell(\ell+1)mm}} = \sum_\ell W_\ell.$$

The three matrices $\mathbf{K}_1$, $\mathbf{K}_2$, and $\mathbf{K}_3$ are correspond to the x, y and z direction of the velocity. The sparse matrix $\mathbf{K}_3$ can be explicitly written down

$$[\mathbf{K}_3]_{\ell\ell' mm'} = \begin{cases} -\ell\sqrt{\frac{\ell^2 - m^2}{4\ell^2 - 1}} & \text{if } m = m' \text{ and } \ell' = \ell - 1 \\ (\ell+1)\sqrt{\frac{(\ell+1)^2 - m^2}{4(\ell+1)^2 - 1}} & \text{if } m = m' \text{ and } \ell' = \ell + 1 \\ 0 & \text{otherwise.} \end{cases}$$

Note that we define W_ℓ using $\mathbf{K}_3$ only. By symmetry, the sum W_ℓ is the same regardless whether we define it with $\mathbf{K}_1$, $\mathbf{K}_2$ or $\mathbf{K}_3$. For larger ℓ, W_ℓ is larger. It suggests that the corresponding a_ℓ contributes the more in the Quadratic Estimator. See Figure 4.1. The modes are corresponding to the change of C_ℓ. The weight is large when $|C_\ell - C_{\ell+1}|$ is large. The CMB power spectrum C_ℓ is

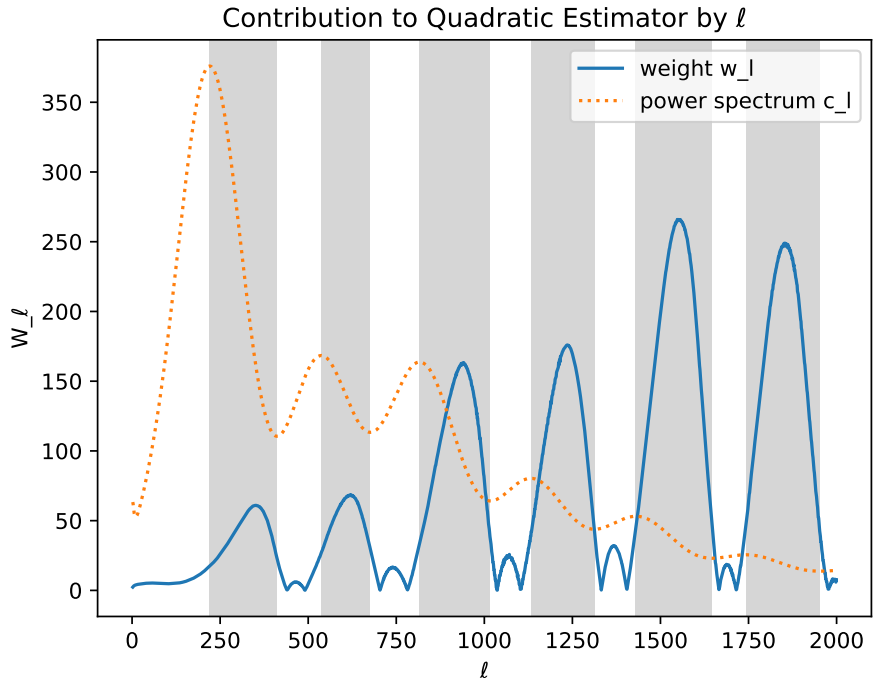


FIGURE 4.1. Weights in Quadratic Estimator, W_ℓ defined at equation (4.7). The values of W_ℓ depends highly on the power spectrum C_ℓ. The shaded area is highlighting the "modes" of W_ℓ which is corresponding to the decrease of power spectrum.

has "oscillations". See Figure 4.2. At the peaks ℓ, the data does not help us to determine. Away from those peaks, the data contribute the most. a_ℓ with larger difference $|C_{\ell+1} - C_\ell|$ contributes more to Quadratic Estimator.

Looking at the weights W_ℓ at Figure 4.1, the area under the curve is the SNR signal-to-noise ratio. Thus we can see the SNR is coming from different range of ℓ. The bigger area under the curve, the larger SNR, the smaller standard error. Under approximation, the variance of the $a_{\ell m}a_{\ell' m'}$ is

$$\mathrm{Var}[a_{\ell m}a_{\ell' m'}] \approx (C_\ell + N_\ell)(C_{\ell'} + N_{\ell'}) = w_{\ell\ell' mm'},$$

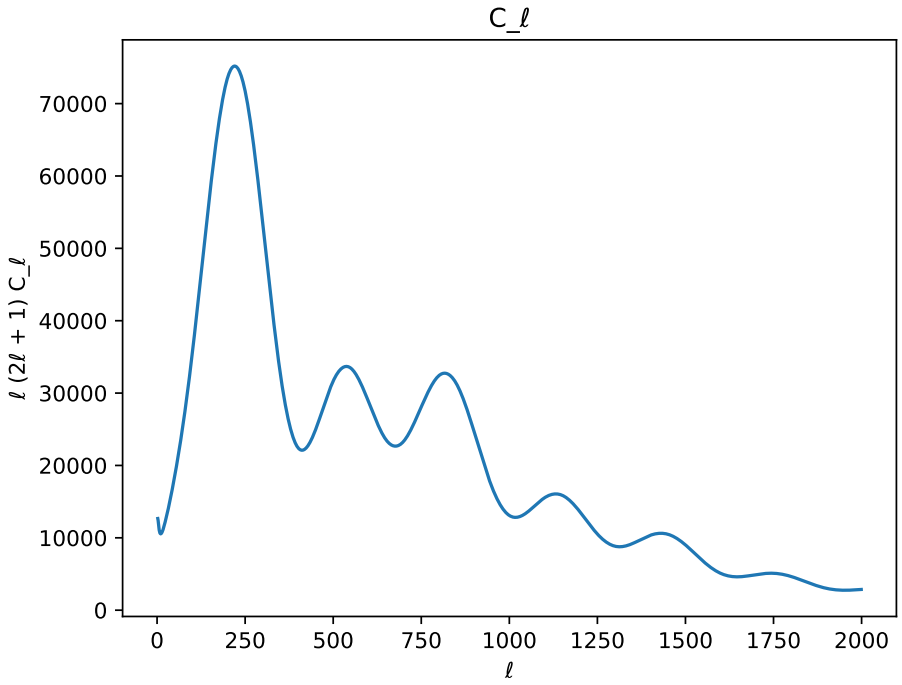


FIGURE 4.2. This graph displays the CMB power spectrum, C_ℓ, with respect to the multipole moment, ℓ. The spectrum is characterized by an oscillatory pattern, with the amplitude of oscillations representing the scale of C_ℓ. Both the peaks (local maxima) and the valleys (local minima) are regions where the CMB data are relatively flat, offering weaker information for parameter estimation. Between these extrema, where there is a significant change in $|C_\ell - C_{\ell+1}|$, the data are more influential. In these areas, the Quadratic Estimator is more sensitive, with coefficients a_ℓ showing larger differences $|C_{\ell+1} - C_\ell|$ having a greater impact on the estimation.

therefore

$$\begin{aligned}
\mathrm{Var}(\hat{\eta}_i) &= \frac{\sum_{\ell<\ell'} \sum_{mm'} \left(\frac{v_{i,\ell\ell' mm'}}{w_{\ell\ell' mm'}}\right)^2 \mathrm{Var}[a_{\ell m} a_{\ell' m'}]}{\left(\sum_{\ell<\ell'} \sum_{mm'} \frac{v^2_{i,\ell\ell' mm'}}{w_{\ell\ell' mm'}}\right)^2} \\
&\approx \frac{\sum_{\ell<\ell'} \sum_{mm'} \left(\frac{v_{i,\ell\ell' mm'}}{w_{\ell\ell' mm'}}\right)^2 w_{\ell\ell' mm'}}{\left(\sum_{\ell<\ell'} \sum_{mm'} \frac{v^2_{i,\ell\ell' mm'}}{w_{\ell\ell' mm'}}\right)^2} \\
&= \frac{1}{\left(\sum_{\ell<\ell'} \sum_{mm'} \frac{v^2_{i,\ell\ell' mm'}}{w_{\ell\ell' mm'}}\right)} \\
&= \frac{1}{\mathrm{SNR}}
\end{aligned}$$

and

$$\sigma_{\mathrm{QE}} = \sqrt{\mathrm{Var}(\hat{\eta}_i)} = \frac{1}{\sqrt{\mathrm{SNR}}}.$$

The standard error of Quadratic Estimator $\hat{\boldsymbol{\eta}}$ in each component is σ_{QE}. The Quadratic Estimator has distribution $\hat{\boldsymbol{\eta}} - \boldsymbol{\eta} \sim N(\mathbf{0}, \sigma_{\mathrm{QE}}^2 \mathbf{I})$ approximately. From numerical stimulation, this approximation is pretty good across different true values of η_i.

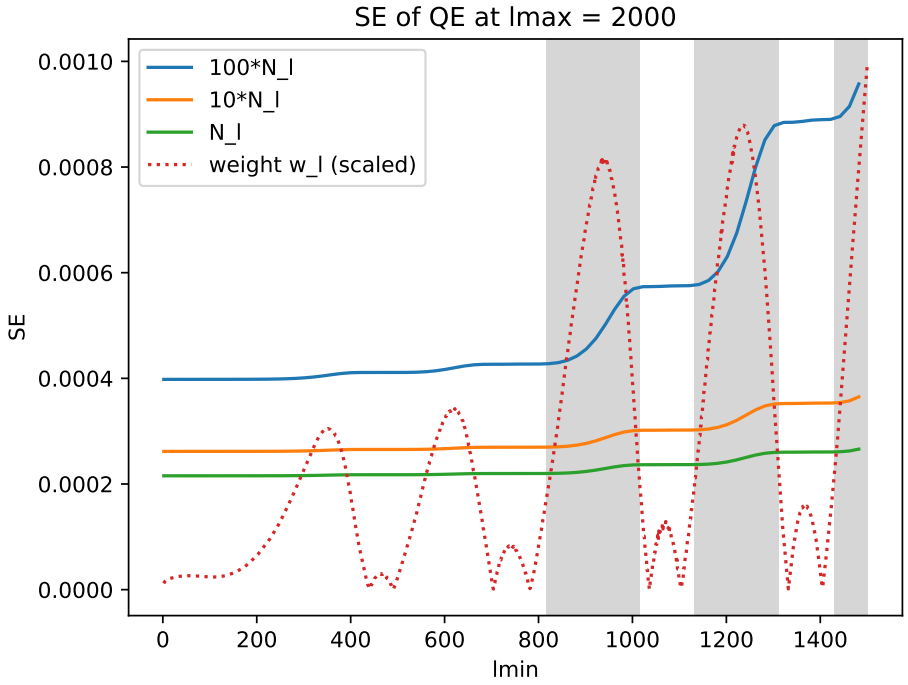


FIGURE 4.3. Standard error of Quadratic Estimator. The standard error (SE) of Quadratic Estimator (computed with $\ell \in [\ell_{\min}, \ell_{\max}]$). SE decreases significantly when going through weights W_ℓ. The shaded areas is highlighting the significant decrease in SE corresponding to the large weights W_ℓ.

4.4.0.2. *Fisher Scoring Algorithm.* We would like to write down the formula of the covariance matrix $\mathbf{V}(\boldsymbol{\eta})$ of the marginal distribution of the spherical harmonic coefficients $\mathbf{a}^{\mathrm{obs}}$ in the CMB peculiar velocity model.

Let $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{pri}} + \boldsymbol{\Sigma}_{\mathrm{noise}}$. Assume that $\boldsymbol{\Sigma}_{\mathrm{noise}}$ has the form $\sigma_{\mathrm{noise}}\mathbf{I}$ where σ_{noise} is a constant standard deviation of the observation noise and $\mathbf{I}$ is identity matrix. Let $\mathbf{B} = \exp\left(\sum_{i=1}^{d} \eta_i \mathbf{K}_i\right)$. We can combine $\boldsymbol{\Sigma}_{\mathrm{pri}}$ and $\boldsymbol{\Sigma}_{\mathrm{noise}}$ into one single covariance matrix $\boldsymbol{\Sigma}$ and obtain explicit formulas

for $\mathbf{V}$:

$$\begin{aligned}
\mathbf{V} &= \mathbf{B}\boldsymbol{\Sigma}_{\text{pri}}\mathbf{B}^T + \boldsymbol{\Sigma}_{\text{noise}} \\
&= \mathbf{B}\boldsymbol{\Sigma}_{\text{pri}}\mathbf{B}^{-1} + \mathbf{B}\boldsymbol{\Sigma}_{\text{noise}}\mathbf{B}^{-1} \\
&= \mathbf{B}\left(\boldsymbol{\Sigma}_{\text{pri}} + \boldsymbol{\Sigma}_{\text{noise}}\right)\mathbf{B}^{-1} \\
&= \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T \\
\mathbf{V}^{-1} &= \left(\mathbf{B}\boldsymbol{\Sigma}_{\text{pri}}\mathbf{B}^T + \boldsymbol{\Sigma}_{\text{noise}}\right)^{-1} \\
&= \mathbf{B}\left(\boldsymbol{\Sigma}_{\text{pri}} + \boldsymbol{\Sigma}_{\text{noise}}\right)^{-1}\mathbf{B}^{-1} \\
&= \mathbf{B}\boldsymbol{\Sigma}^{-1}\mathbf{B}^T.
\end{aligned}$$

We are ready to write down $\mathbf{V}$ and its derivatives as follows:

$$\begin{aligned}
\mathbf{V}(\boldsymbol{\eta}) &= \text{Var}\left[\mathbf{x}|\,\boldsymbol{\eta}\right] = \exp\left(\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right)\boldsymbol{\Sigma}\exp\left(-\sum_{i=1}^{d}\eta_i\mathbf{K}_i\right) \\
\mathbf{V}(\mathbf{0}) &= \boldsymbol{\Sigma} \\
\left.\frac{\partial\mathbf{V}}{\partial\eta_i}\right|_{\boldsymbol{\eta}=\mathbf{0}} &= [\mathbf{K}_i, \boldsymbol{\Sigma}_{\text{pri}}].
\end{aligned}$$

and also approximately

$$\mathbf{V}(\eta) \approx \boldsymbol{\Sigma} + \sum_{i=1}^{d}\eta_i[\mathbf{K}_i, \boldsymbol{\Sigma}_{\text{pri}}].$$

We may assume $\boldsymbol{\Sigma}_{\text{pri}}$ and $\boldsymbol{\Sigma}_{\text{obs}}$ are diagonal matrices. When $\boldsymbol{\lambda} = \mathbf{0}$, $b_k = \text{tr}\left[\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\lambda_i}\right] = 0$ which simplifies the formula of Quadratic Estimator to a pure quadratic term $\mathbf{x}^T\mathbf{A}_k\mathbf{x}$.

CHAPTER 5

Lie Group Formulation

5.1. Introduction

In this chapter, we propose an estimator for the peculiar velocity problem in the CMB. Our proposed estimator is based on maximum likelihood estimation and uses the Newton-Raphson method. Unlike the Quadratic Estimator, which relies on Fisher information, our proposed estimator utilizes observed information. To derive the estimator, we use Lie group theory to describe the peculiar velocity effect on the CMB. This approach not only enables us to efficiently compute the gradient and Hessian of the log-likelihood function, which are essential for the Newton-Raphson method, but also reveals the mathematical structure of the peculiar velocity effect on the CMB using the Lorentz Lie group, and show that the parameter space of peculiar velocity has the structure of a non-Euclidean manifold, specifically the Lorentz Lie group.

5.2. Space of covariance matrices

The primordial CMB is modeled as an isotropic Gaussian Random Field with a zero mean and a rotationally-invariant covariance function. Let $\mathbf{x}$ represent the spherical harmonic coefficients or Fourier coefficients of this isotropic Gaussian Random Field. We consider $\mathbf{x}$ as a random vector following a multivariate Gaussian distribution, $\mathbf{x} \sim N(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is the corresponding covariance matrix. In essence, the space of square-integrable Gaussain Random Fields on the sphere is isometric to the space of square-summable spherical harmonic coefficients, meaning they are equivalent up to a change of basis. Hence, we can use the terms "random field" and "spherical harmonic coefficients" interchangeably.

The requirement of rotational invariance on the covariance function means that the isotropic Gaussian random field should retain an identical distribution before and after applying a rotational transformation. In the language of spherical harmonics coefficients or Fourier coefficients, this condition translates to the invariance of the covariance matrix upon transformation, denoted as

$R\mathbf{x}$. Specifically, $R\mathbf{x}$ follows a Gaussian distribution with zero mean and covariance matrix $\mathbf{\Sigma}$, i.e., $R\mathbf{x} \sim N(\mathbf{0}, \mathbf{\Sigma})$, for any rotation operator R. The set of these rotation operators forms a group, labeled SO(3) for spherical harmonics and SO(2) for Fourier coefficients. These rotation operators form a group, denoted by SO(3) for the sphere and SO(2) for the circle. These rotation operators can be defined both in the space of the random field and in the spaces of spherical harmonic or Fourier coefficients. For convenience, we will use R to denote the rotation operator in both spaces. The covariance matrix $\mathbf{\Sigma}$ of an isotropic Gaussian random field must satisfy the property

$$R\mathbf{\Sigma}R^T = \mathbf{\Sigma},$$

for any rotation operator R. Since rotation operators are unitary (meaning their transpose or adjoint is also their inverse, $R^T = R^{-1}$), the transformation $R \to R\mathbf{\Sigma}R^T$ constitutes a group conjugation. In the context of group theory, a conjugation map is defined as $g \to ghg^{-1}$ for any $g, h \in G$, where G is a group. This conjugation map establishes a group isomorphism, which lets us link the space of covariance matrices with the rotation operator R. More formally, we link $R\mathbf{\Sigma}R^T$ with R, and endow the space $\{R\mathbf{\Sigma}R^T : R\}$ with the group structure of R. Furthermore, the inverse of a rotationally invariant covariance matrix is also rotationally invariant:

$$R\mathbf{\Sigma}^{-1}R^{-1} = \mathbf{\Sigma}^{-1}.$$

In the specific case of isotropic Gaussian random fields, the isomorphism is trivial because there is only one element $\mathbf{\Sigma}$, which serves as the "zero element" of the group.

In Chapter 3, we established that the effect of peculiar velocity can be mathematically formulated as a Lorentz Boost operator, which we denote as $B_{\boldsymbol{\eta}}$. This operator, parameterized by the rapidity vector $\boldsymbol{\eta}$, acts on the spherical harmonic coefficients of an isotropic Gaussian Random Field (GRF) with covariance matrix $\mathbf{\Sigma}$. Analogous to rotation operators, Lorentz Boost operators $B_{\boldsymbol{\eta}}$ act linearly on the spherical harmonic coefficients of the GRF. In a naive formulation, the parameterized space of spherical harmonic coefficients can be represented as

$$\{B_{\boldsymbol{\eta}}\mathbf{x} : \boldsymbol{\eta} \in \mathbb{R}^3\},$$

where $\mathbf{x} \sim N(\mathbf{0}, \mathbf{\Sigma})$ and $\mathbf{\Sigma}$ is a known covariance matrix satisfying the condition of rotation-invariance. Each transformed Gaussian random field $B_{\boldsymbol{\eta}}\mathbf{x}$ possesses a covariance matrix given by $B_{\boldsymbol{\eta}}\mathbf{\Sigma}B_{\boldsymbol{\eta}}^T$. Consequently, we obtain a set of Lorentz Boost operators acting on the Gaussian Random Field, denoted as $\{B_{\boldsymbol{\eta}} : \boldsymbol{\eta} \in \mathbb{R}^3\}$. These linear operators generate all potentially observable Gaussian Random Fields through the transformation $f(\mathbf{x}; \boldsymbol{\eta}) := f(B_{\boldsymbol{\eta}}\mathbf{x})$, whose probability density function is given by

$$f(\mathbf{x}; \boldsymbol{\eta}) = (2\pi)^{-\frac{k}{2}} \det(\mathbf{\Sigma}^{-1})^{\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{x}^T (B_{\boldsymbol{\eta}}\mathbf{\Sigma}B_{-\boldsymbol{\eta}})^{-1}\mathbf{x}\right).$$

The parameter space is the set of all possible rapidity vectors $\boldsymbol{\eta}$. As the Lorentz Boost operator $B_{\boldsymbol{\eta}}$ is unitary, the determinant of the covariance matrix remains invariant under transformations parameterized by $\boldsymbol{\eta}$. To formalize this notion, we introduce the set $\mathbf{L}(\mathbf{\Sigma})$, defined as

$$\mathbf{L}(\mathbf{\Sigma}) := \{B_{\boldsymbol{\eta}}\mathbf{\Sigma}B_{-\boldsymbol{\eta}} : \boldsymbol{\eta} \in \mathbb{R}^3\},$$

where $\mathbf{\Sigma}$ is a known rotation-invariant covariance matrix. This set essentially represents a parameterized collection of covariance matrices, which we can relate to the parameterized space of spherical harmonic coefficients $\{B_{\boldsymbol{\eta}}\mathbf{x} : \boldsymbol{\eta} \in \mathbb{R}^3\}$. Given that $B_{\boldsymbol{\eta}}$ is unitary, this set comprises conjugation maps of the form $\boldsymbol{\eta} \mapsto B_{\boldsymbol{\eta}}\mathbf{\Sigma}B_{\boldsymbol{\eta}}^{-1}$.

To develop algorithms for identifying the specific covariance matrix of the primordial CMB within the set $\mathbf{L}(\mathbf{\Sigma})$ that corresponds to the observed CMB, we need to traverse the Lorentz boost space $B_{\boldsymbol{\eta}}$. Mathematically, one would desire to apply a Lorentz Boost $B_{\boldsymbol{\eta}}$ to a given covariance matrix $\mathbf{S}$ such that $B_{\boldsymbol{\eta}} \cdot \mathbf{S} = B_{\boldsymbol{\eta}}\mathbf{S}B_{\boldsymbol{\eta}}^T$. However, the complications due to Wigner rotation must be accounted for. Operating solely within the set of pure boosts $\{B_{\boldsymbol{\eta}} : \boldsymbol{\eta} \in \mathbb{R}^3\}$ may not provide a comprehensive understanding of the parameter space due to the Wigner rotation. Wigner rotation is a physical consequence that arises when two non-collinear Lorentz boosts are composed, yielding a Lorentz transformation that is not purely a boost but also includes a rotation component. For instance, given two non-collinear rapidity vectors $\mathbf{u}$ and $\mathbf{v}$, the composite action of their respective Lorentz boosts on a random field can be written as $B_{\mathbf{v}}B_{\mathbf{u}}f = RB_{\mathbf{w}}f$, where $\mathbf{w}$ is not necessarily equal to $\mathbf{u} + \mathbf{v}$ and R is a rotation operator. Without loss of generality, we examine a specific case involving two rapidity vectors $\mathbf{u} = (\eta, 0, 0)$ and $\mathbf{v} = (\lambda\cos\phi, \lambda\sin\phi, 0)$. The composite effect of

these pure boosts is a combined rotation and boost given by:

$$B_{\mathbf{v}}B_{\mathbf{u}} = R(\epsilon)B_{\mathbf{w}},$$

where $\mathbf{w} = (\xi\cos\theta, \xi\sin\theta, 0)$ and $R(\epsilon)$ is a rotation along the z-axis by an angle ϵ in a counter-clockwise direction. The variables ξ, θ, and ϵ are determined by the equations:

$$\begin{aligned}
\cosh\xi &= \cosh\lambda\cosh\eta + \cos\phi\sinh\lambda\sinh\eta \\
\tan\theta &= \frac{\sin\phi\sinh\lambda}{\sinh\eta\cosh\lambda + \cos\phi\sinh\lambda\cosh\eta} \\
\tan\frac{\epsilon}{2} &= -\frac{\sin\phi\sinh\frac{\lambda}{2}\sinh\frac{\eta}{2}}{\cosh\frac{\lambda}{2}\cosh\frac{\eta}{2} + \cos\phi\sinh\frac{\lambda}{2}\sinh\frac{\eta}{2}}.
\end{aligned}$$

We may define the operation $\mathbf{v}\oplus\mathbf{u} = \mathbf{w}$ as the "addition" of rapidities, leading to the expression

$$B_{\mathbf{v}}B_{\mathbf{u}} = R(\epsilon)B_{\mathbf{v}\oplus\mathbf{u}}.$$

Additionally, through inversion and renaming, this can be rewritten as

$$B_{\mathbf{v}}B_{\mathbf{u}} = B_{\mathbf{u}\oplus\mathbf{v}}R(\epsilon).$$

Given that $\boldsymbol{\Sigma}$ is rotation-invariant, the rotation term R due to Wigner rotation is absorbed by $\boldsymbol{\Sigma}$, that is,

$$B_{\mathbf{v}}B_{\mathbf{u}}\boldsymbol{\Sigma}B_{-\mathbf{u}}B_{-\mathbf{v}} = B_{\mathbf{u}\oplus\mathbf{v}}\boldsymbol{\Sigma}B_{-(\mathbf{u}\oplus\mathbf{v})}.$$

According to the polar decomposition theorem [**Mor02**], any element of the Lorentz group can be uniquely decomposed into a product of rotation and pure boost. Consequently, $\mathbf{L}(\boldsymbol{\Sigma})$ can be understood as the coset space $\mathrm{SO}^+(3,1)/\mathrm{SO}(3)$. Thus, to navigate $\mathbf{L}(\boldsymbol{\Sigma})$, one should apply the conjugate action of pure Lorentz boosts $B_{\boldsymbol{\eta}}$ to the covariance matrices in $\mathbf{L}(\boldsymbol{\Sigma})$. Conceptually, $\mathbf{L}(\boldsymbol{\Sigma})$ should be treated as a homogeneous space under the group $\mathrm{SO}^+(3,1)$ with $\mathrm{SO}(3)$ as the stabilizer and $\mathbf{0}$ as the origin. In this framework, the rotational component within the Lorentz group becomes a nuisance parameter.

In summary, by employing the equations governing Wigner rotations, we can traverse the parameter space represented by $\mathbf{L}(\boldsymbol{\Sigma})$. Additionally, traversal is possible in the set of values defined

as

$$\{\mathbf{x}^T \left(B_{\boldsymbol{\eta}} \boldsymbol{\Sigma} B_{\boldsymbol{\eta}}^T\right)^{-1} \mathbf{x} : \boldsymbol{\eta} \in \mathbb{R}^3\}$$

by retaining the vectors $B_{\boldsymbol{\eta}}\mathbf{x}$ and subjecting them to a Lorentz boost $B_{\mathbf{u}}$, given some rapidity vector $\mathbf{u}$. In this formalism, the space $X = \mathbf{L}(\boldsymbol{\Sigma})$ is endowed with the group $\mathrm{SO}^+(3,1)$ and the stabilizer $\mathrm{SO}(3)$. The group action can be described as $gx \to x'$, where

$$g = B_{\mathbf{v}}$$

$$x = B_{\mathbf{u}} \boldsymbol{\Sigma} B_{\mathbf{u}}^T$$

$$x' = B_{\mathbf{u}\oplus\mathbf{v}} \boldsymbol{\Sigma} B_{\mathbf{u}\oplus\mathbf{v}}^T$$

$$B_{\mathbf{u}\oplus\mathbf{v}} R = B_{\mathbf{v}} B_{\mathbf{u}} \text{ for some rotation } R.$$

We then identify $\mathbf{L}(\boldsymbol{\Sigma})$ with the coset space $\mathrm{SO}^+(3,1)/\mathrm{SO}(3)$ via the relation

$$B_{\boldsymbol{\eta}} \boldsymbol{\Sigma} B_{\boldsymbol{\eta}}^T \longleftrightarrow B_{\boldsymbol{\eta}} \cdot \mathrm{SO}(3).$$

This mapping is well-defined due to the rotation-invariance of $\boldsymbol{\Sigma}$, as $R\boldsymbol{\Sigma}R^T = \boldsymbol{\Sigma}$ for any $R \in \mathrm{SO}(3)$. If $\boldsymbol{\Sigma}$ is trivial, meaning $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$ with $\sigma^2 > 0$, the set $\mathbf{L}(\boldsymbol{\Sigma})$ would contain only $\boldsymbol{\Sigma}$, and thus we make the assumption that $\boldsymbol{\Sigma}$ is non-trivial. With this identification, the theory of Lie Groups becomes applicable for traversing the parameter space. Specifically, the quotient space $\mathrm{SO}^+(3,1)/\mathrm{SO}(3)$ is an oriented hyperbolic space. Furthermore, $\mathbf{L}(\boldsymbol{\Sigma})$, characterized as $\mathrm{SO}^+(1,3)/\mathrm{SO}(3)$, is a Riemannian symmetric space of negative curvature. This implies that the parameter space under consideration is a non-Euclidean space.

In our analysis, observational noise is modeled as additive Gaussian noise, represented by the covariance matrix $\sigma^2_{\text{noise}}\mathbf{I}$. Here, σ^2_{noise} is a positive scalar representing the variance of the noise. Consequently, the observed CMB temperature map becomes a Gaussian random vector with a covariance matrix given by:

$$B_{\boldsymbol{\eta}} \boldsymbol{\Sigma} B_{\boldsymbol{\eta}}^T + \sigma^2_{\text{noise}}\mathbf{I}.$$

Since $B_{\boldsymbol{\eta}}$ is an unitary operator, i.e. $B_{\boldsymbol{\eta}}^T = B_{\boldsymbol{\eta}}^{-1}$, the additive Gaussian noise term can be incorporated into the primordial CMB. The covariance matrix then simplifies to:

$$B_{\boldsymbol{\eta}}(\boldsymbol{\Sigma} + \sigma^2_{\text{noise}}\mathbf{I})B_{\boldsymbol{\eta}}^T.$$

Importantly, this resulting covariance matrix, $\boldsymbol{\Sigma} + \sigma^2_{\text{noise}}\mathbf{I}$, remains rotationally invariant. Under this framework, the spherical harmonic coefficients of the primordial CMB can be modeled as a Gaussian random vector with the aforementioned combined covariance matrix. This interpretation is based on the assumption that the additive Gaussian noise is white, meaning its covariance matrix is proportional to the identity matrix, $\sigma^2_{\text{noise}}\mathbf{I}$.

Furthermore, we extend our analysis to the space of covariance matrices generated by the Lorentz boost operator, denoted by $\mathbf{L}(B_{\boldsymbol{\eta}}\boldsymbol{\Sigma}B_{\boldsymbol{\eta}}^T)$, for a rapidity vector $\boldsymbol{\eta}$ and a rotationally invariant covariance matrix $\boldsymbol{\Sigma}$ both are being given. This allows us to consider the covariance matrix of not just an Isotropic Gaussian Random Field but also its transformations under Lorentz Boosts. The advantage of this formulation is its flexibility, as we can choose any Lorentz-Boosted covariance matrix, $B_{\boldsymbol{\eta}}\boldsymbol{\Sigma}B_{\boldsymbol{\eta}}^T$, as the origin of this space. To explore this idea further, let's consider the nested covariance matrix space:

$$\mathbf{L}(\mathbf{L}(\boldsymbol{\Sigma})) = \{B_{\boldsymbol{\eta}}CB_{\boldsymbol{\eta}}^T : \boldsymbol{\eta} \in \mathbb{R}^3, C \in \mathbf{L}(\boldsymbol{\Sigma})\}.$$

A natural question arises: Is the operation of taking the Lorentz-Boosted covariance matrices idempotent, i.e., does $\mathbf{L}(\mathbf{L}(\boldsymbol{\Sigma})) = \mathbf{L}(\boldsymbol{\Sigma})$ hold for all $\boldsymbol{\Sigma}$? This equality holds when $\boldsymbol{\Sigma}$ is rotationally invariant. However, if $\boldsymbol{\Sigma}$ is not rotationally invariant, this equality may not hold in general. When f is an isotropic Gaussian random field, i.e. the corresponding covariance matrix $\boldsymbol{\Sigma}$ is rotationally invariant, both Rf and f share the same distribution. This allows us to absorb the rotation operator R into $\boldsymbol{\Sigma}$, implying that the covariance matrix of the spherical harmonic coefficients of $B\boldsymbol{\eta}_2 B\boldsymbol{\eta}_1 f$ can be expressed as a covariance matrix of $B\boldsymbol{\eta}_3\boldsymbol{\Sigma}B^T\boldsymbol{\eta}_3$ for some $\boldsymbol{\eta}_3$. This further implies that $\mathbf{L}(\mathbf{L}(\boldsymbol{\Sigma})) \subseteq \mathbf{L}(\boldsymbol{\Sigma})$, and therefore $\mathbf{L}(\mathbf{L}(\boldsymbol{\Sigma})) = \mathbf{L}(\boldsymbol{\Sigma})$ when $\boldsymbol{\eta} = 0$.

5.3. Lorentz Lie Group

This section aims to introduce the foundational concepts of Lie algebra and Lie group, particularly focusing on their relevance to the Lorentz Boost operator. The set of Lorentz Boost operators, parameterized by rapidity, forms a Lie group. The corresponding Lie algebra is identified as the infinitesimal perturbation of the Lorentz Boost. Understanding these perturbations is crucial for differentiating the likelihood functions with respect to the Lorentz Boost, thereby facilitating the computation of the maximum likelihood estimator.

Let E denote a vector space. We first turn our attention to the space of linear operators from E to itself, denoted $\mathrm{L}(E)$, which also forms a vector space. A critical component in our algorithm is the Lie bracket, which provides formulas for differentiating the exponential map.

Definition 5.3.1. The Lie bracket $[\cdot,\cdot] : \mathrm{L}(E)\times\mathrm{L}(E) \to \mathrm{L}(E)$ is defined by $[X,Y] = XY - YX$ for any $X, Y \in \mathrm{L}(E)$.

Consider a vector subspace $\mathfrak{g}$ of $\mathrm{L}(E)$, such as the space of anti-symmetric matrices or $\mathrm{L}(E)$ itself. The Lie bracket is an alternating bilinear map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying $[X,Y] = -[Y,X]$ for $X, Y \in \mathfrak{g}$. This bracket must also satisfy the Jacobi identity, given by $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ for $X, Y, Z \in \mathfrak{g}$. We can verify by direct computation that the previous definition of Lie bracket satisfies Jacobi identity.

Proposition 5.3.1. The Lie bracket complies with the Jacobi identity, $[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]] = 0$, for any $X, Y, Z \in \mathrm{L}(E)$.

By definition, a Lie algebra is a vector space $\mathfrak{g}$ equipped with a Lie bracket. In the context of this study, it serves as the space of infinitesimal perturbations for the Lorentz Boost operator.

Continuing from our discussion on the Lie algebra $\mathfrak{g}$ and its role in the Lorentz Boost operator, we now focus on the 'little adjoint operator' in Lie algebra. This operator is essential for differentiating the Lorentz Boost operator.

Definition 5.3.2. The little adjoint operator $\mathrm{ad} : \mathfrak{g} \to \mathrm{L}(\mathfrak{g})$ is a linear operation defined as follows:

$$\mathrm{ad}(X)Y := [X,Y] = XY - YX,$$

where $X, Y \in \mathfrak{g}$. Equivalently, for any vector $\mathbf{v} \in E$,

$$(\mathrm{ad}(X)Y)\,\mathbf{v} = XY\mathbf{v} - YX\mathbf{v}.$$

Since $\mathfrak{g}$ is itself a vector space, the Lie bracket can be defined on $\mathrm{L}(\mathfrak{g})$ in a manner analogous to $\mathrm{L}(E)$. An important property, stemming from the Jacobi identity, is that the little adjoint operator preserves the Lie bracket. This can be expressed as:

$$[\mathrm{ad}(X), \mathrm{ad}(Y)](Z) = \mathrm{ad}([X, Y])Z.$$

We now transition to the concept of a Lie group G. A Lie group is both a group and a differentiable manifold, where the group operations of multiplication and inversion are smooth maps. Smoothness of the group multiplication means that the mapping $(x, y) \mapsto x^{-1}y$ is smooth for any $x, y \in G$.

Definition 5.3.3. Let $\mathrm{GL}(E)$ represent the space of invertible linear operators from E to itself. We define the exponential map $\exp : \mathrm{L}(E) \to \mathrm{GL}(E)$, when E is finite dimensional, as:

$$\exp \mathbf{X} := \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{X}^k,$$

where $\mathbf{X} \in \mathrm{L}(E)$.

Next, we introduce the concept of the Adjoint action, which serves as a homomorphism between the spaces $\mathrm{GL}(E)$ and $\mathrm{GL}(\mathrm{L}(E))$.

Definition 5.3.4. The Adjoint action is defined as:

$$\mathrm{Ad} : \mathrm{GL}(E) \to \mathrm{GL}(\mathrm{L}(E))$$

by

$$(\mathrm{Ad}(\mathbf{A})\mathbf{S})\,\mathbf{v} = \mathbf{A}\mathbf{S}\mathbf{A}^{-1}\mathbf{v},$$

for every $\mathbf{S} \in \mathrm{L}(E)$, $\mathbf{A} \in \mathrm{GL}(E)$, and $\mathbf{v} \in E$.

After defining the Adjoint action, it's important to note that $\mathrm{Ad}(\mathbf{A}) \in \mathrm{GL}(L(E))$ is a group homomorphism. Specifically, the Adjoint action follows the equation:

$$\mathrm{Ad}(\mathbf{A}_1)(\mathbf{S})\mathrm{Ad}(\mathbf{A}_2)(\mathbf{S}) = \mathrm{Ad}(\mathbf{A}_1\mathbf{A}_2)(\mathbf{S}),$$

This shows that the Adjoint action preserves the group structure.

Before we move to the theorem about the Adjoint and exponential maps working together, let's talk about the link between Lie groups and Lie algebras. For any Lie group G, there exists a corresponding Lie algebra, denoted $\mathrm{Lie}(G)$. In formal terms, you can think of the Lie group G as generated by the set $\exp(X) : X \in \mathrm{Lie}(G)$.

THEOREM 5.3.1. The Adjoint map and the exponential map commute:

$$\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}(X)).$$

Note that if $\mathrm{ad}(X) \in \mathrm{L}(\mathrm{L}(E))$, then $\exp(\mathrm{ad}(X)) \in \mathrm{GL}(\mathrm{L}(E))$. This theorem can be visually represented by the following commutative diagram:

$$\begin{array}{ccc} \mathrm{L}(E) & \xrightarrow{\mathrm{ad}} & \mathrm{L}(\mathrm{L}(E)) \\ \downarrow{\scriptstyle \exp} & & \downarrow{\scriptstyle \exp} \\ \mathrm{GL}(E) & \xrightarrow{\mathrm{Ad}} & \mathrm{GL}(\mathrm{L}(E)) \end{array}$$

In Chapter 2, we discussed the toy model of peculiar velocity affecting the CMB on a circle. We now extend this discussion to comprehend the behavior of CMB velocity in terms of Lie groups. Our exploration will first consider a subgroup of the Lorentz group, before moving on to the more general spherical model. The Lorentz subgroup corresponding to the circle CMB peculiar velocity model has a Lie algebra generated by the set $\{\mathbf{J}, \mathbf{K}_x, \mathbf{K}_y\}$. The structure constants are as follows:

$$[\mathbf{K}_x, \mathbf{K}_y] = -\mathbf{J},$$

$$[\mathbf{J}, \mathbf{K}_x] = \mathbf{K}_y,$$

$$[\mathbf{J}, \mathbf{K}_y] = -\mathbf{K}_x.$$

The Jacobi identity can be verified through direct computation:

$$[\mathbf{K}_x, [\mathbf{K}_y, \mathbf{J}]] + [\mathbf{K}_y, [\mathbf{J}, \mathbf{K}_x]] + [\mathbf{J}, [\mathbf{K}_x, \mathbf{K}_y]] = [\mathbf{K}_x, \mathbf{K}_x] + [\mathbf{K}_y, \mathbf{K}_y] + [\mathbf{J}, -\mathbf{J}] = \mathbf{0}.$$

Here, we utilize the property that the Lie bracket is alternating, meaning $[X, X] = 0$ for all X in the Lie algebra.

For the spherical model, the corresponding Lie algebra is generated by $\{\mathbf{J}_x, \mathbf{J}_y, \mathbf{J}_z, \mathbf{K}_x, \mathbf{K}_y, \mathbf{K}_z\}$. The structure constants are given by:

$$\begin{aligned}
[\mathbf{K}_i, \mathbf{K}_j] &= -\sum_k \epsilon_{ijk} \mathbf{J}_k, \\
[\mathbf{J}_i, \mathbf{J}_j] &= \sum_k \epsilon_{ijk} \mathbf{J}_k, \\
[\mathbf{J}_i, \mathbf{K}_j] &= \sum_k \epsilon_{ijk} \mathbf{K}_k.
\end{aligned}$$

Here, ϵ_{ijk} is the Levi-Civita symbol, defined by:

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \text{ is an even permutation of } (x,y,z), \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (x,y,z). \end{cases}$$

5.4. Maximum Likelihood Over Lie Groups

Consider a multivariate Gaussian distribution with zero mean and a rotation-invariant covariance matrix, $\mathbf{\Sigma}$. The Lorentz Boost is characterized by a rapidity vector, $\boldsymbol{\eta}$. The likelihood function for this distribution is given by:

$$f(\mathbf{x}; \boldsymbol{\eta}) = (2\pi)^{-\frac{k}{2}} \det(\mathbf{\Sigma})^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \mathbf{x}^T \left(B_{\boldsymbol{\eta}} \mathbf{\Sigma} B_{-\boldsymbol{\eta}} \right)^{-1} \mathbf{x} \right), \tag{5.1}$$

where k denotes the dimension of the random vector $\mathbf{x}$. The objective is to find the value of $\boldsymbol{\eta}$ that minimizes the negative log-likelihood. This corresponds to minimizing the following quadratic form involving $\boldsymbol{\eta}$:

$$\arg\min_{\boldsymbol{\eta}} (B_{-\boldsymbol{\eta}} \mathbf{x})^T \mathbf{\Sigma}^{-1} (B_{-\boldsymbol{\eta}} \mathbf{x}).$$

Here, we assume that both the data $\mathbf{x}$ and the covariance matrix $\mathbf{\Sigma}$ are given. To improve computational efficiency, we first transform the data vector $\mathbf{x}$ by applying the inverse Lorentz boost

$B_{-\boldsymbol{\eta}}$, resulting in $\mathbf{y} = B_{-\boldsymbol{\eta}}\mathbf{x}$. Afterward, the quantity $\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}$ can be computed without explicitly calculating the inverse of the entire matrix $B_{\boldsymbol{\eta}}\boldsymbol{\Sigma}B_{-\boldsymbol{\eta}}$. In the special case of an Isotropic Gaussian Random Field, the covariance matrix $\boldsymbol{\Sigma}$ is diagonal, simplifying computation of its inverse $\boldsymbol{\Sigma}^{-1}$. By structuring the problem this way, we avoid computational bottlenecks and facilitate an efficient solution for $\boldsymbol{\eta}$.

We aim to minimize the quadratic form $(B_{-\boldsymbol{\eta}}\mathbf{x})^T\boldsymbol{\Sigma}^{-1}(B_{-\boldsymbol{\eta}}\mathbf{x})$ over $\boldsymbol{\eta} \in \mathbb{R}^3$. This can be achieved iteratively by applying successive Lorentz boosts. Starting with an initial element $(B_{-\boldsymbol{\eta}_0}\mathbf{x})^T\boldsymbol{\Sigma}^{-1}(B_{-\boldsymbol{\eta}_0}\mathbf{x})$ for some $\boldsymbol{\eta}_0$, we look for an incremental change $\Delta\boldsymbol{\eta}$ such that

$$(B_{-\Delta\boldsymbol{\eta}}B_{-\boldsymbol{\eta}_0}\mathbf{x})^T\boldsymbol{\Sigma}^{-1}(B_{-\Delta\boldsymbol{\eta}}B_{-\boldsymbol{\eta}_0}\mathbf{x}) < (B_{-\boldsymbol{\eta}_0}\mathbf{x})^T\boldsymbol{\Sigma}^{-1}(B_{-\boldsymbol{\eta}_0}\mathbf{x}).$$

To streamline the notation, let $\mathbf{y} = B_{-\boldsymbol{\eta}}\mathbf{x}$ and define a function $g(\mathbf{v})$ as

$$g(\mathbf{v}) = (B_{\mathbf{v}}\mathbf{y})^T\boldsymbol{\Sigma}^{-1}(B_{\mathbf{v}}\mathbf{y}).$$

Our objective then becomes finding a vector $\mathbf{v}$ such that $g(\mathbf{v}) < g(\mathbf{0})$. Once such a $\mathbf{v} = -\Delta\boldsymbol{\eta}$ is identified, we update $\mathbf{y}$ using the Wigner rotation formula

$$B_{-\Delta\boldsymbol{\eta}}B_{-\boldsymbol{\eta}_0} = RB_{-(\Delta\boldsymbol{\eta}\oplus\boldsymbol{\eta}_0)},$$

where R is some rotation matrix. This results in $\mathbf{y} = RB_{-(\Delta\boldsymbol{\eta}\oplus\boldsymbol{\eta}_0)}\mathbf{x}$. It's important to note that the presence of the rotation R does not influence the evaluation of the quadratic form $\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}$. This is because $\boldsymbol{\Sigma}^{-1}$ is rotation-invariant, ensuring that $\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y} = (B_{-(\Delta\boldsymbol{\eta}\oplus\boldsymbol{\eta}_0)}\mathbf{x})^T\boldsymbol{\Sigma}^{-1}(B_{-(\Delta\boldsymbol{\eta}\oplus\boldsymbol{\eta}_0)}\mathbf{x})$.

To determine the optimal $\mathbf{v}$ that minimizes $g(\mathbf{v})$, we differentiate $g(\mathbf{v})$ with respect to $\mathbf{v}$. We initiate the process with $\boldsymbol{\eta}_0 = \mathbf{0}$ and aim to find a sequence of updates $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ that iteratively minimize $\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}$, where $\mathbf{y} = B_{\mathbf{v}_n}\cdots B_{\mathbf{v}_2}B_{\mathbf{v}_1}\mathbf{x}$ represents the transformed data at each step. After obtaining this sequence, we determine a $\boldsymbol{\eta} \in \mathbb{R}^3$ such that $RB_{-\boldsymbol{\eta}} = B_{\mathbf{v}_n}\cdots B_{\mathbf{v}_2}B_{\mathbf{v}_1}$ for some rotation matrix R. Applying the Wigner rotation formula, we obtain:

$$\boldsymbol{\eta} = -(((\mathbf{v}_n \oplus \mathbf{v}_{n-1}) \oplus \mathbf{v}_{n-2})\cdots \oplus \mathbf{v}_1).$$

To find the minimizer $\arg\min_{\mathbf{v}} g(\mathbf{v})$, where $g(\mathbf{v}) = (B_{\mathbf{v}}\mathbf{y})^T\boldsymbol{\Sigma}^{-1}(B_{\mathbf{v}}\mathbf{y})$, we differentiate $g(\mathbf{v})$ with respect to $\mathbf{v}$. The gradient and Hessian of $g(\mathbf{v})$ can be explicitly computed as follows:

$$\frac{\partial}{\partial v_i} g\bigg|_{\mathbf{v}=\mathbf{0}} = \mathbf{y}^T[K_i, \boldsymbol{\Sigma}^{-1}]\mathbf{y},$$
$$\frac{\partial^2}{\partial v_i \partial v_j} g\bigg|_{\mathbf{v}=\mathbf{0}} = \mathbf{y}^T[K_j, [K_i, \boldsymbol{\Sigma}^{-1}]]\mathbf{y},$$

where $[A, B] = AB - BA$ represents the matrix commutator.

PROBLEM 5.4.1. Optimization Problem of Lorentz Boost

Objective Function: For $\boldsymbol{\eta} \in \mathbb{R}^3$, define the objective function $F : \mathbb{R}^3 \to \mathbb{R}$ as

$$F(\boldsymbol{\eta}) := (B_{-\boldsymbol{\eta}}\mathbf{x})^T\boldsymbol{\Sigma}^{-1}(B_{-\boldsymbol{\eta}}\mathbf{x})$$

Optimization Problem: The problem aims to find an optimal $\hat{\boldsymbol{\eta}} \in \mathbb{R}^3$ that minimizes the objective function F. This is formally expressed as

$$\hat{\boldsymbol{\eta}} = \arg\min_{\boldsymbol{\eta}\in\mathbb{R}^3} F(\boldsymbol{\eta})$$

ALGORITHM 5.4.1. Newton–Raphson Method for Optimizing Problem 5.4.1

Input: $\mathbf{x}$, $\boldsymbol{\Sigma}$, and $\{K_i : i = 1, 2, 3\}$

Output: $\boldsymbol{\eta}$

Initialization: $\mathbf{y} \leftarrow \mathbf{x}$, $\mathbf{v} \leftarrow \mathbf{0}$

Algorithm Steps: (1) **Update Vector y:**

$$\mathbf{y} \leftarrow B_{\mathbf{v}}\mathbf{y}$$

(2) **Compute Objective Function g:**

$$g(\mathbf{v}) = \mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}$$

(3) **Compute Gradient and Hessian:** Define the gradient vector $\mathbf{g}$ and the Hessian matrix $\mathbf{H}$ as follows:

$$\mathbf{g} = \begin{pmatrix} \frac{\partial}{\partial v_1} g\Big|_{\mathbf{v}=\mathbf{0}} \\ \vdots \\ \frac{\partial}{\partial v_i} g\Big|_{\mathbf{v}=\mathbf{0}} \end{pmatrix},$$

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2}{\partial v_1^2} g\Big|_{\mathbf{v}=\mathbf{0}} & \cdots & \frac{\partial^2}{\partial v_1 \partial v_i} g\Big|_{\mathbf{v}=\mathbf{0}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial v_i \partial v_1} g\Big|_{\mathbf{v}=\mathbf{0}} & \cdots & \frac{\partial^2}{\partial v_i^2} g\Big|_{\mathbf{v}=\mathbf{0}} \end{pmatrix}$$

(4) **Update Vector v:** Let

$$\mathbf{v}(\lambda) = -\lambda \mathbf{H}^{-1} \mathbf{g}$$

where $\lambda > 0$. Search for a λ such that $g(\mathbf{v}(\lambda))$ is minimized. Update $\mathbf{v} \leftarrow \mathbf{v}(\lambda)$.

(5) **Convergence Check:** If $\|\mathbf{v}\|$ is below a predetermined threshold, terminate the algorithm; otherwise, go back to the first step.

Problem 5.4.2. Optimization Problem in Exponential Map

Objective Function: Let $\mathrm{L}(E)$ be a space of linear operators on the vector space E. Let $\{K_j\}_{j=1}^d \subset \mathrm{L}(E)$ be a basis of $\mathrm{L}(E)$. Consider a linear functional $F : \mathrm{L}(E) \to \mathbb{R}$. The function F and the basis set $\{K_j\}$ satisfy, for any $X, Y \in \mathrm{span}\{K_j\}$, there exists $W \in \mathrm{span}\{K_j\}$ such that

$$F(\exp W) = F(\exp(X) \exp(Y))$$

Optimization Problem: The objective is to find an optimal $\hat{X} \in \mathrm{span}\{K_j\}$ that minimizes $F(\exp X)$, formally defined as

$$\hat{X} = \underset{X \in \mathrm{span}\{K_j\}}{\arg\min} F(\exp X)$$

Algorithm 5.4.2. Newton Method for Optimization Problem 5.4.2

Input: Tolerance threshold, Initial guess for X_{new}

Output: Optimal X_{new}

Initialization: (1) Set a tolerance threshold.

(2) Initialize $X_{\text{new}} \in \text{span}\{K_j\}$, e.g., $\mathbf{0}$.

(3) Let $F_X(Y) := F(\exp(X)\exp(Y))$.

Algorithm Steps: (1) **Assign X:**

$$X = X_{\text{new}}$$

(2) **Compute Gradient and Hessian:** Compute the gradient $\mathbf{g}$ of $F_X(Y)$ with respect to Y as

$$\mathbf{g} = \begin{pmatrix} F(\exp(X)K_1) \\ \vdots \\ F(\exp(X)K_j) \\ \vdots \end{pmatrix}.$$

Compute the Hessian $\mathbf{H}$ as

$$\mathbf{H} = \begin{pmatrix} F(\exp(X)K_1K_1) & \cdots & F(\exp(X)K_1K_d) \\ \vdots & \ddots & \vdots \\ F(\exp(X)K_dK_1) & \cdots & F(\exp(X)K_dK_d) \end{pmatrix}.$$

(3) **Update Y:** Update Y using $\delta\mathbf{t} = -\mathbf{H}^{-1}\mathbf{g}$, i.e.,

$$Y = \sum_{j=1}^{d} \delta t_j K_j.$$

(4) **Update X_{new}:** Update X_{new} to satisfy $F(\exp X_{\text{new}}) = F(\exp(X)\exp(Y))$.

(5) **Convergence Check:** Check for convergence: If $|F(\exp X_{\text{new}}) - F(\exp X)|$ is less than the tolerance, terminate. Otherwise, return to the first step.

PROBLEM 5.4.3. Optimization Problem in Exponential Map with Log-linear Objective.

Objective Function: Given linear functions $F_i : \mathrm{L}(E) \to \mathbb{R}$, and a basis set $\{K_j\}_{j=1}^{d} \subset L(E)$, the objective function in this case is $\sum_i \log F_i(\exp X)$.

Assumption: For any $X, Y \in \text{span}\{K_j\}$, there exists a $W \in \text{span}\{K_j\}$ such that

$$F_i(\exp W) = F_i(\exp(X)\exp(Y)).$$

Optimization Objective: Find an optimal $\hat{X} \in \text{span}\{K_j\}$ that minimizes the log-linear function. Formally,

$$\hat{X} = \underset{X \in \text{span}\{K_j\}}{\arg\min} \sum_{i=1}^{n} \log F_i(\exp X).$$

ALGORITHM 5.4.3. Newton method for the optimization problem 5.4.3. Define $F_X(Y) := \sum_n \log F_n(\exp(X)\exp(Y))$. Initialize $X_{\text{new}} \in \text{span}\{K_j\}$.

Algorithm Steps: (1) **Assign X:** Update $X = X_{\text{new}} \in \text{span}\{K_j\}$

(2) **Compute Gradient and Hessian:** Compute the gradient of $F_X(Y)$ w.r.t. Y, namely $\mathbf{g}$:

$$\mathbf{g} = \sum_n \begin{pmatrix} \frac{F_n(\exp(X)K_1)}{F_n(\exp X)} \\ \vdots \\ \frac{F_n(\exp(X)K_j)}{F_n(\exp X)} \\ \vdots \end{pmatrix}$$

and Hessian

$$\mathbf{H} = \sum_n \left[\frac{F_n(\exp X)F_n(\exp(X)K_iK_j) - F_n(\exp(X)K_i)F_n(\exp(X)K_j)}{(F_n(\exp X))^2} \right].$$

Then take $Y = \sum_{j=1}^{d} \delta t_j K_j$ where $(\delta t_1, \ldots, \delta t_d) = -\mathbf{H}^{-1}\mathbf{g}$.

(3) **Update X_{new}:** Set X_{new} such that $F_n(\exp X_{\text{new}}) = F_n(\exp(X)\exp(Y))$.

(4) **Convergence Check:** Check if $|F_X(Y) - F_X(\mathbf{0})| <$ tolerance. If not, go back to first step.

PROBLEM 5.4.4. Optimization problem in Adjoint exponential map

Objective Function: Let $F : \text{L}(E) \to \mathbb{R}$ be a linear function. Let $\{K_j\}_{j=1}^{d} \subset L(E)$ be a basis set. Assume that F satisfies the condition that, for any $X \in \text{span}\{K_j\}$ and

$Y \in \text{span}\{K_j\}$, there exists $W \in \text{span}\{K_j\}$ such that

$$F(\text{Ad}\exp W) = F(\text{Ad}\exp(X)\text{Ad}\exp(Y)).$$

Optimization Problem: We aim to find $\hat{X} \in L(E)$ that optimizes $F(\text{Ad}\exp X)$, i.e.

$$\hat{X} = \underset{X \in \text{span}\{K_j\}}{\arg\min} \; F(\text{Ad}\exp X)$$

ALGORITHM 5.4.4. Newton method for the optimization problem 5.4.4.

Define $F_X(Y) := F(\exp(X)\exp(Y))$. Initialize $X_{\text{new}} \in \text{span}\{K_j\}$.

Algorithm Steps: (1) Update $X = X_{\text{new}} \in \text{span}\{K_j\}$

(2) Compute the gradient of $F_X(Y)$ w.r.t. Y, namely $\mathbf{g}$:

$$\mathbf{g} = \begin{pmatrix} F((\text{Ad}\exp X)\text{ad}(K_1)) \\ \vdots \\ F((\text{Ad}\exp X)\text{ad}(K_j)) \\ \vdots \end{pmatrix}$$

and Hessian

$$\mathbf{H} = \begin{pmatrix} F((\text{Ad}\exp X)\text{ad}(K_1)\text{ad}(K_1)) & & \ddots & \\ & \ddots & & F((\text{Ad}\exp X)\text{ad}(K_i)\text{ad}(K_j)) \\ & & & \ddots \end{pmatrix}.$$

Then take $Y = \sum_{j=1}^{d} \delta t_j K_j$ where $(\delta t_1, \ldots, \delta t_d) = -\mathbf{H}^{-1}\mathbf{g}$.

(3) Find X_{new} such that $F(\exp X_{\text{new}}) = F(\exp(X)\exp(Y))$.

(4) Return to step 1.

5.4.1. Standard Error. The Fisher information $\mathcal{I}(\boldsymbol{\eta})$ of the likelihood function 5.1 is given by

$$[\mathcal{I}(\boldsymbol{\eta})]_{ij} = -\frac{1}{2}\text{tr}\left(\left[\mathbf{K}_i, \left[\mathbf{K}_j, \boldsymbol{\Sigma}^{-1}\right]\right]\boldsymbol{\Sigma}\right)$$

Note that the Fisher Information does not depend on the parameter $\boldsymbol{\eta}$.

When Σ is diagonal and $\mathbf{K}_i$'s are orthogonal in Frobenius inner product, i.e. $\mathrm{tr}(\mathbf{K}_i\mathbf{K}_j) = 0$ for $i \neq j$, the formula can be further simplified as

$$[\mathcal{I}(\boldsymbol{\eta})]_{ij} = \begin{cases} \sum_{k<k'} [\mathbf{K}_i]^2_{kk'} \frac{([\boldsymbol{\Sigma}_i]_k - [\boldsymbol{\Sigma}_i]_{k'})^2}{[\boldsymbol{\Sigma}]_k [\boldsymbol{\Sigma}]_{k'}} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Approximately, the MLE $\hat{\boldsymbol{\eta}}$ has distribution $N(\boldsymbol{\eta}, \mathcal{I}(\boldsymbol{\eta})^{-1})$.

5.5. Newton-Wigner Iteration

This section introduces the Newton-Wigner iteration method as an alternative to the Quadratic Estimator. The method consists of two components:

(1) Newton Iteration: Utilizing the derived formulas for the gradient and Hessian of the log-likelihood, we apply the Newton method for optimization when the noise is constant across the frequency k.

(2) Wigner Rotation: The primary computational bottleneck in evaluating the gradient and Hessian is the matrix exponential. To address this, we include a Wigner rotation step that allows us to update the matrix-exponential-vector component in an efficient manner.

5.5.1. Preliminaries. Let us define the action of the matrix exponential on the data vector $\mathbf{a}^{\mathrm{obs}}$ as:

$$\mathbf{q}(\boldsymbol{\eta}) = \exp\left(-\sum_{i=1}^{d} \eta_i \mathbf{K}_i\right) \mathbf{a}^{\mathrm{obs}},$$

and the quadratic forms as:

$$Q(\boldsymbol{\eta}, \mathbf{S}) = \mathbf{q}(\boldsymbol{\eta})^T \mathbf{S} \mathbf{q}(\boldsymbol{\eta}),$$

where $\mathbf{S}$ could be $\boldsymbol{\Sigma}^{-1}, [\mathbf{K}_i, \boldsymbol{\Sigma}^{-1}]$, or $[\mathbf{K}_i, [\mathbf{K}_j, \boldsymbol{\Sigma}^{-1}]]$.

The objective function $F(\boldsymbol{\eta})$, its gradient $\nabla F(\boldsymbol{\eta})$, and its Hessian $H(\boldsymbol{\eta})$ are given by:

$$F(\boldsymbol{\eta}) = Q(\boldsymbol{\eta}, \boldsymbol{\Sigma}^{-1}),$$

$$\nabla_i F(\boldsymbol{\eta}) = Q(\boldsymbol{\eta}, [\mathbf{K}_i, \boldsymbol{\Sigma}^{-1}]),$$

$$H_{ij}(\boldsymbol{\eta}) = Q(\boldsymbol{\eta}, [\mathbf{K}_i, [\mathbf{K}_j, \boldsymbol{\Sigma}^{-1}]])$$

5.5.2. Newton Method. The Newton method to optimize $F(\boldsymbol{\eta})$ is:

$$\boldsymbol{\eta}_{n+1} = \boldsymbol{\eta}_n - [H(\boldsymbol{\eta}_n)]^{-1} \nabla F(\boldsymbol{\eta}_n), \quad n = 0, 1, 2, \ldots$$

with the initial guess $\boldsymbol{\eta}_0 = \mathbf{0}$.

However, for $\boldsymbol{\eta}_n$ with large magnitudes, evaluating $\mathbf{q}(\boldsymbol{\eta})$ directly using the matrix exponential can be computationally expensive.

5.5.3. Wigner Rotation. Given $\boldsymbol{\eta}_n$, we aim to efficiently update it to $\boldsymbol{\eta}_{n+1}$. Specifically, we seek a low-magnitude Lorentz Boost Lie algebra Δ such that:

$$\mathbf{q}(\boldsymbol{\eta}_{n+1}) = \exp(\Delta \cdot \mathbf{K})\mathbf{q}(\boldsymbol{\eta}_n).$$

However, if $\boldsymbol{\eta}_n$ and $\boldsymbol{\eta}_{n+1}$ are not parallel, then a Δ like this cannot exist, because of Wigner rotation. Mathematically, this is captured by:

$$\exp(\boldsymbol{\eta}_{n+1} \cdot \mathbf{K}) \exp(-\boldsymbol{\eta}_n \cdot \mathbf{K}) = \mathbf{R} \exp(\Delta \cdot \mathbf{K})$$

where $\mathbf{R}$ is a rotation matrix. Multiplying both sides by $\mathbf{R}^{-1}$ from the left, we have

$$\mathbf{R}^{-1}\mathbf{q}(\boldsymbol{\eta}_{n+1}) = \exp(\Delta \cdot \mathbf{K})\mathbf{q}(\boldsymbol{\eta}_{n+1}).$$

The extra rotation $\mathbf{R}^{-1}$ does not affect the value of the negative log likelihood function for an Isotropic Gaussian Random Field, since $\mathbf{R}\boldsymbol{\Sigma}\mathbf{R}^{-1} = \boldsymbol{\Sigma}$ for any rotation $\mathbf{R}$. This is because

$$(\mathbf{R}^{-1}\mathbf{q}(\boldsymbol{\eta}_{n+1}))^T \boldsymbol{\Sigma} (\mathbf{R}^{-1}\mathbf{q}(\boldsymbol{\eta}_{n+1})) = \mathbf{q}(\boldsymbol{\eta}_{n+1})^T \boldsymbol{\Sigma} \mathbf{q}(\boldsymbol{\eta}_{n+1}).$$

To obtain such Δ, we use the following formulaes: For any $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$

$$\exp(\boldsymbol{\eta}_2 \cdot \mathbf{K}) \exp(\boldsymbol{\eta}_1 \cdot \mathbf{K}) = \mathbf{R} \exp(\boldsymbol{\eta}_3 \cdot \mathbf{K})$$

where for some rotation $\mathbf{R}$ and $\boldsymbol{\eta}_3$ can be obtained by the formula of its norm and unit direction:

$$\cosh\|\boldsymbol{\eta}_3\| = \cosh\|\boldsymbol{\eta}_1\|\cosh\|\boldsymbol{\eta}_2\| + \left(\frac{\boldsymbol{\eta}_1}{\|\boldsymbol{\eta}_1\|}\cdot\frac{\boldsymbol{\eta}_2}{\|\boldsymbol{\eta}_2\|}\right)\sinh\|\boldsymbol{\eta}_1\|\sinh\|\boldsymbol{\eta}_2\|$$

$$\frac{\boldsymbol{\eta}_3}{\|\boldsymbol{\eta}_3\|} = \frac{1}{\sinh\|\boldsymbol{\eta}_3\|}\left(\sinh\|\boldsymbol{\eta}_1\|\cosh\|\boldsymbol{\eta}_2\|\frac{\boldsymbol{\eta}_1}{\|\boldsymbol{\eta}_1\|} + \sinh\|\boldsymbol{\eta}_2\|\frac{\boldsymbol{\eta}_2}{\|\boldsymbol{\eta}_2\|}\right.$$
$$\left. + \frac{\sinh^2\|\boldsymbol{\eta}_1\|\sinh\|\boldsymbol{\eta}_2\|}{\cosh\|\boldsymbol{\eta}_1\|+1}\left(\frac{\boldsymbol{\eta}_1}{\|\boldsymbol{\eta}_1\|}\cdot\frac{\boldsymbol{\eta}_2}{\|\boldsymbol{\eta}_2\|}\right)\frac{\boldsymbol{\eta}_1}{\|\boldsymbol{\eta}_1\|}\right).$$

We omit the explicit formula for the rotation $\mathbf{R}$ as it is not required in our computations. Setting $\boldsymbol{\eta}_1 = -\boldsymbol{\eta}_n$ and $\boldsymbol{\eta}_2 = -\boldsymbol{\eta}_{n+1}$, we obtain $\Delta = \boldsymbol{\eta}_3$.

5.5.4. Newton-Wigner Iteration. The Newton-Wigner iteration combines the Newton method with the Wigner Rotation to optimize the log-likelihood function $\log f(\mathbf{a}^{\text{obs}};\boldsymbol{\eta})$. The steps are as follows:

(1) Calculate the preliminary update $\boldsymbol{\eta}^*$ via the standard Newton update:

$$\boldsymbol{\eta}^* = \boldsymbol{\eta} - [H(\boldsymbol{\eta})]^{-1}\nabla F(\boldsymbol{\eta}).$$

(2) Use Wigner rotation to find Δ so that

$$\exp(\boldsymbol{\eta}^*\cdot\mathbf{K})\exp(-\boldsymbol{\eta}\cdot\mathbf{K}) = \mathbf{R}\exp(\Delta\cdot\mathbf{K})$$

where $\mathbf{R}$ is the Wigner rotation term. With small enough updates, this enables more efficient computation.

(3) Refine $\mathbf{q}(\boldsymbol{\eta}^*) = \exp(\Delta\cdot\mathbf{K})\mathbf{q}(\boldsymbol{\eta})$.

5.6. Comparison to other works

In this section, we compare our Lie group framework with the Bayesian estimation approach used in [**SSM$^+$21**], which is conducted within the bipolar spherical harmonic decomposition framework.

In [**SSM$^+$21**], the covariance function of the Gaussian Random Field on a sphere is described as a bipolar spherical harmonic decomposition, as indicated in [**JRS12**]. Peculiar velocity is characterized using spherical harmonic coefficients of the dipole field. This parameterization offers

advantages in working with the bipolar spherical harmonic decomposition of the Gaussian Random Field's covariance function, as it simplifies the derivation of linear approximations of the bipolar spherical harmonic coefficients from the spherical harmonic coefficients of the dipole field. Consequently, the bipolar spherical harmonic coefficients of the covariance function are parameterized by the spherical harmonic coefficients of the peculiar velocity dipole field. Thus, the likelihood function of the Gaussian Random Field for observed CMB images is formulated in terms of the spherical harmonic coefficients of the peculiar velocity dipole field. Following the formulation of this likelihood function, Bayesian sampling techniques are applied, specifically employing the Hamiltonian Monte Carlo method for sampling the spherical harmonic coefficients of the peculiar velocity dipole field.

The representation of peculiar velocity through spherical harmonic coefficients of the dipole field is effectively the same as the rapidity parameterization in our framework, as both parameterizations transform peculiar velocity into $\mathbb{R}^3$. In first order approximation, the rapidity and the spherical harmonic coefficients of the dipole field are related by a linear transformation. The advantage of parameterization using rapidity is that we can naturally leverage the mathematical results in Lorentz geometry which describes special relativity. Nevertheless, rapidity is a more natural parameterization as it is with the Lorentz geometry which describes special relativity. Expressing the effect of peculiar velocity on the Gaussian Random Field on the sphere in terms of rapidity simplifies many of the complicated-looking coefficients in the formulas. The study in [**SSM$^+$21**] involves extensive calculations in mathematical physics, which may present a barrier to readers without a strong background in the subject. Our rapidity-based parameterization leads to simpler and more elegant formulas without relying on such intensive calculations. It highlights the mathematical elegance of Einstein's theory of relativity using Lorentz geometry. Moreover, the parameterization in terms of rapidity is exact, as it does not rely on linear approximations. It retains all information in higher order, enabling higher order differentiation, such as Hessian. In contrast, the study in [**SSM$^+$21**] is limited by its use of linear approximations, which restricts the ability to perform higher-order differentiations.

CHAPTER 6

Numerical Experiments

6.1. Introduction

This chapter is organized into three main sections. First, we employ numerical simulations to illustrate the effects of peculiar velocity on the CMB using Isotropic Gaussian Random Fields and Lorentz Boosts. Second, we analyze the impact of Lorentz Boosts on the statistical properties of the CMB Gaussian Random Field. Finally, we conduct numerical experiments to evaluate the performance of our newly proposed Newton Method-based estimator for peculiar velocity. We assess the accuracy of this method by calculating the Root-Mean-Square (RMS) error in simulations and comparing it with the Quadratic Estimator.

This chapter contains a series of figures designed to illustrate the impacts of peculiar velocity on CMB observations. Figure 6.1 and Figure 6.2 focus on light aberration effects. Figure 6.3 elaborates on the Doppler effect and how it manifests on a spherical image under different rapidity magnitudes. Figure 6.4 provides an integrated perspective, combining both light aberration and Doppler effects. Figures 6.5 and 6.6 apply these phenomena to a simulated Gaussian Random Field, demonstrating realistic effects under Lorentz Boost. Figures 6.7 and 6.8 explore the implications of varying the maximum frequency of spherical harmonics, denoted as $\ell_{\max}$, in simulating Lorentz Boost.

6.2. Illustrations of peculiar velocity effect

Figure 6.1 illustrates the aberration of light due to peculiar velocity. Aberration, a consequence of the finite speed of light in conjunction with the observer's motion, leads to an apparent shift in the incoming light's direction. This is demonstrated through the progressive displacement of patterns on the celestial sphere in the illustrations in figure 6.1, as the observer moves relative to the CMB rest frame. The Lorentz Boost has rapidity $\boldsymbol{\eta}$ with same direction $(x = 1.0, y = 1.0, z = 1.0)$ but increasing magnitudes $\|\boldsymbol{\eta}\| = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. For $\|\boldsymbol{\eta}\| = 0.0$, it is essentially no

Lorentz boost, so the image is the original image without Lorentz Boost. The dots keep the perfect circular shape throughout the aberration; however, the size of the dots changes. This type of image transformation is a conformal transformation on sphere. For any two cross lines on any location on this sphere, the angle between the cross lines will stay the same after the transformation, i.e. angle-preserving. Thus, any dots with a perfect circular shape will retain their perfect circular shape after the transformation, not becoming an ellipse. Figure 6.2 is the same as Figure 6.1, but in Mollweide projection, showing the front and back of the sphere as a whole. Although the circular shape is not displayed optimally, it shows the magnification and shrinking of the circles. Counter-intuitively, the shrinking occurs in front of the moving observers, and the magnification occurs at the back, as if the dots are moving from the back to the front.

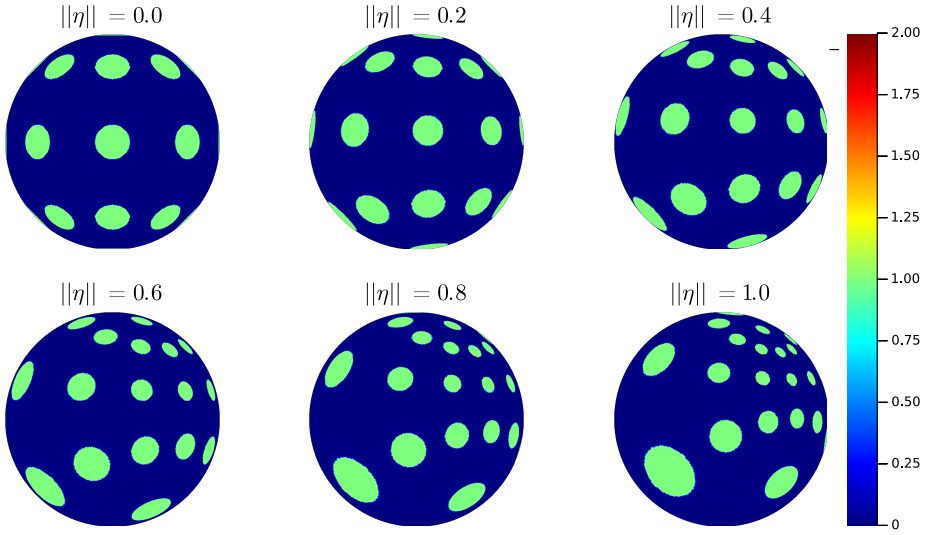


FIGURE 6.1. Illustration of light aberration due to peculiar velocity. The figure demonstrates how the observer's motion causes an apparent shift in the incoming light's direction on a celestial sphere. Rapidity $\boldsymbol{\eta} = \frac{\|\boldsymbol{\eta}\|}{\sqrt{3}}(x = 1.0, y = 1.0, z = 1.0)$ for $\|\boldsymbol{\eta}\| = 0.0, 0, 2, 0.4, 0.6, 0.8, 1.0$. Top left is the primordial image as it is $\|\boldsymbol{\eta}\| = 0.0$. $(x = 1.0, y = 1.0, z = 1.0)$ is point toward the reader to the top right corner direction. The dots gradually shift to the top right corner, and shrinking at the direction, and magnifying on the other side. The dots keep the circular shape as it is conformal mapping.

Figure 6.3 explores the Doppler effect. As the observer moves towards or away from a source of CMB radiation, the detected frequency of the radiation changes, leading to a Doppler shift. This is represented through a gradual change in color across the spherical images, with blue denoting a

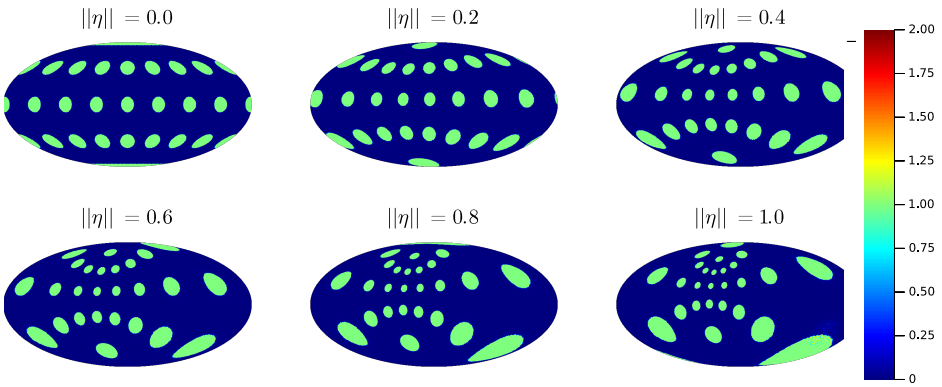


FIGURE 6.2. Mollweide projection of light aberration effects in Figure 6.1, offering a comprehensive view of the celestial sphere and emphasizing the size changes in the dots. Note that the Mollweide projection is not conformal, therefore the dot shape does not preserve the circular shape.

blueshift (towards observer, higher frequency) and red indicating a redshift (away from observer, lower frequency). The Doppler effect corresponds to the same sequence of rapidity with the direction $(x = 1.0, y = 1.0, z = 1.0)$ but increasing magnitudes $\|\boldsymbol{\eta}\| = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. At the start, the neutral temperature is 1.0 everywhere on the sphere. As the magnitude of rapidity increases, the head $(x = 1, y = 1, z = 1)$ gets hotter, while the tail $(x = -1, y = -1, z = -1)$ gets colder. The increase and decrease are multiplicative; that is, if the starting magnitude is any number a, then the same point will have magnitude pa where p is the multiplicative number. The cold spot in the illustration is actually close to 0, but not negative. In other words, the tail has a magnitude closer to 0. The fluctuations are suppressed at the tail and amplified at the head.

Figure 6.4 is a sequence of illustrations which shows combined effects, simultaneously showcasing light aberration and the Doppler effect, i.e. the Lorentz Boost. The resulting pattern and color shifts exemplify the intricate interplay of these phenomena as induced by the observer's peculiar velocity. The images offer a holistic perspective on how our motion relative to the CMB rest frame can significantly influence our observations of the CMB temperature map. At the start, there are dots of temperature 1.0 while in the background rest are zero 0.0 temperature. The light aberration moves the dots from the tail toward the head. The dots are magnified near the tails and sinks near the head. The dots get colder near the tail, while hotter near the head. Figure 6.5 and 6.6 are another series of illustrations in which Lorentz Boost is applied on simulated Gaussian Random

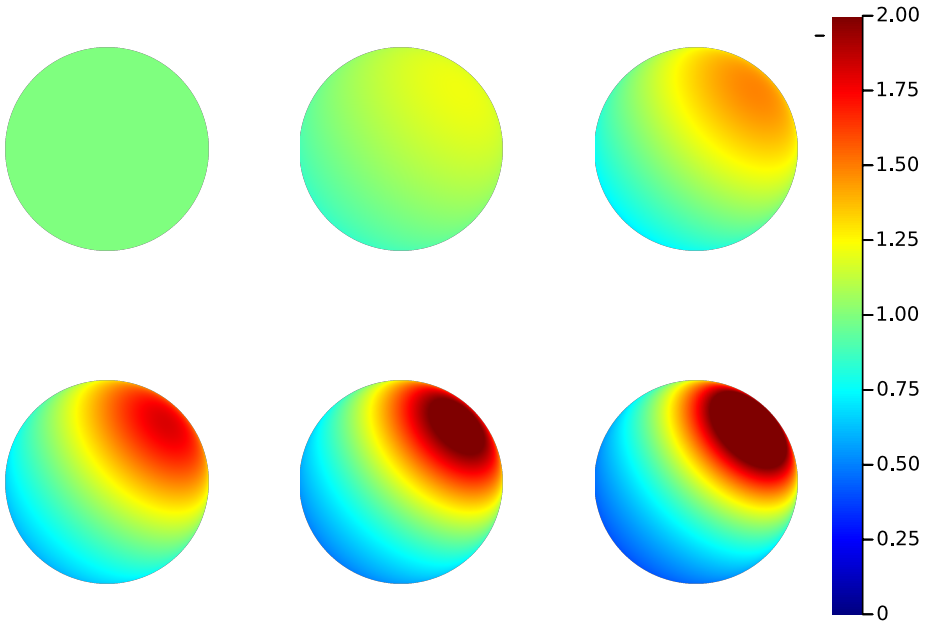


FIGURE 6.3. Illustration of the Doppler effect on the CMB sphere. Rapidity $\boldsymbol{\eta} = \frac{\|\boldsymbol{\eta}\|}{\sqrt{3}}(x = 1.0, y = 1.0, z = 1.0)$ for $\|\boldsymbol{\eta}\| = 0.0, 0, 2, 0.4, 0.6, 0.8, 1.0$. Top left is the primordial image as it is $\|\boldsymbol{\eta}\| = 0.0$. $(x = 1.0, y = 1.0, z = 1.0)$ is point toward the reader to the top right corner direction. Higher values (red in color bar) indicate blueshift (higher frequency), while Lowever values (blue in color bar) indicate redshift (lower frequency). Green in color bar is the neutral value.

Field. The Gaussian Random Field simulated with maximum spherical harmonic frequency at $\ell = 10$ in order to have visible hot spots and cold spots in the illustrations. Under the Lorentz Boost, the hot and cold spots sink towards the head, and magnifies at the tail, and the spots move along from tail to head. The magnitude of the hot and cold spots magnifies at the head while diminish at the tail. It shows the more realistic effects of the Lorentz Boost on CMB temperature map. However, in reality, the magnitude of the peculiar velocity is much smaller, at magnitude around 10^{-3}, so the light aberration is not actually visible. However, the Doppler effect is still visible when the CMB is assumed to be roughly constant.

We wish to underscore the critical role of the maximum frequency of spherical harmonics, denoted as $\ell_{\max}$, in simulating Lorentz Boost. The coefficients at any given frequency can spill over to adjacent frequencies, both lower and higher. This 'leakage' results in the dispersion of coefficient magnitudes across these nearby frequencies.

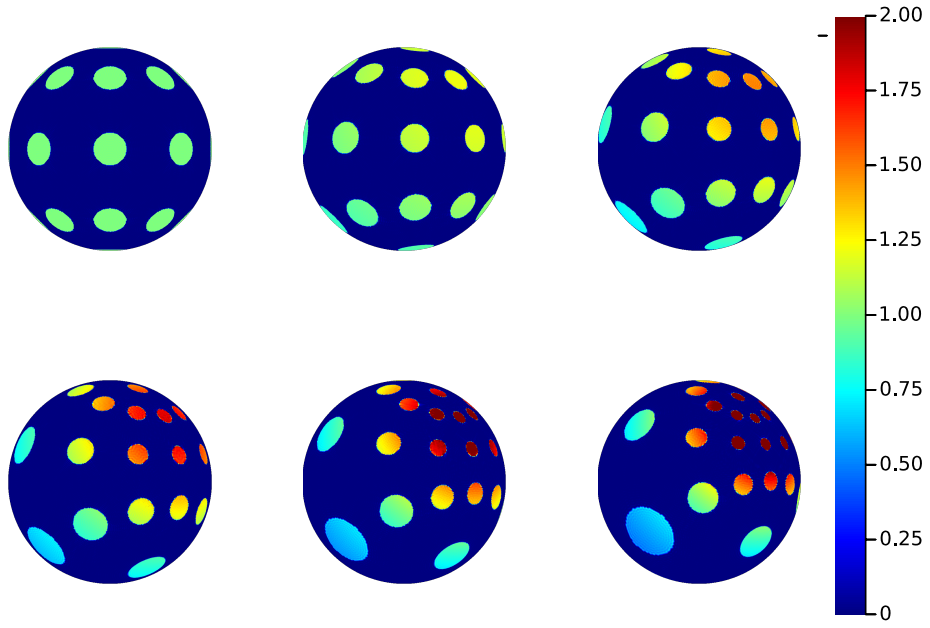


FIGURE 6.4. Combined effects of light aberration and the Doppler effect under Lorentz Boost, illustrating how these phenomena impact CMB observations.

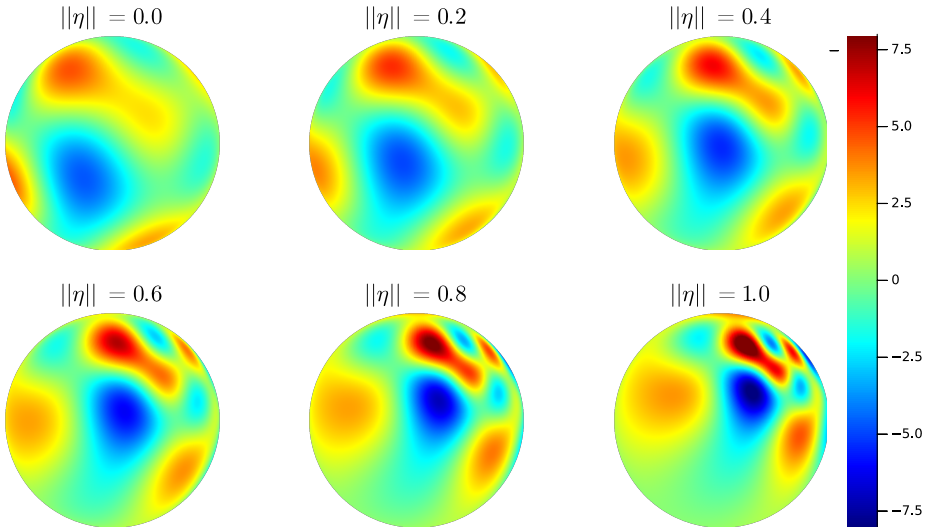


FIGURE 6.5. Application of Lorentz Boost on a simulated Gaussian Random Field with maximum spherical harmonic frequency at $\ell = 10$. The figure showcases the motion of hot and cold spots. Rapidity $\boldsymbol{\eta} = \frac{\|\boldsymbol{\eta}\|}{\sqrt{3}}(x = 1.0, y = 1.0, z = 1.0)$ for $\|\boldsymbol{\eta}\| = 0.0, 0, 2, 0.4, 0.6, 0.8, 1.0$. Top left is the primordial image as it is $\|\boldsymbol{\eta}\| = 0.0$. $(x = 1.0, y = 1.0, z = 1.0)$.

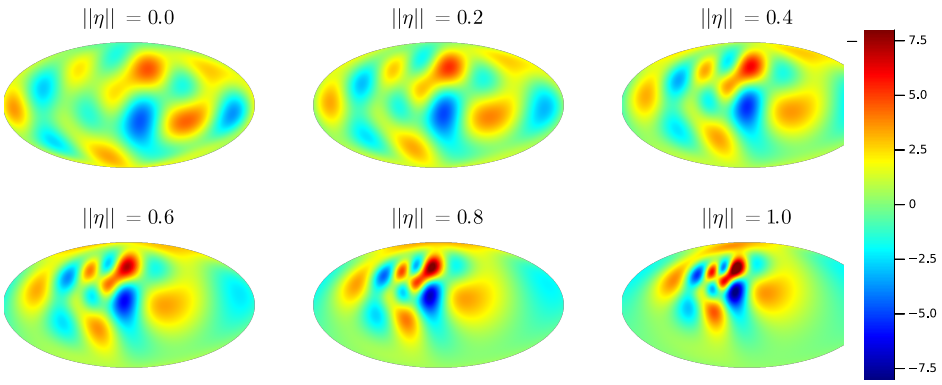


FIGURE 6.6. Mollweide projection of Lorentz Boost effects on a Gaussian Random Field in Figure 6.5, displaying the front and back of the celestial sphere.

Trimming the sparse matrix K while preserving its anti-symmetry ensures that the associated differential equation still maintains the norm of the Gaussian random field. Consequently, the total magnitudes remain constant. If we allocate enough buffer space for this leakage, it can contribute to a flatter spectrum.

Figures 6.7 and 6.8 display Lorentz Boosts applied to the same Gaussian Random Field. However, in these simulations, we intentionally used a particularly small value for $\ell_{\max}$, which showed insufficient for accurately representing the Lorentz Boost phenomenon. This inadequacy is evident as the hot and cold spots fail to effectively sink and magnify. This demonstration indicates the need for a larger $\ell_{\max}$ value in the simulation of Lorentz Boost.

6.3. Effect of Lorentz Boost on the variance of Gaussian Random Field

6.3.1. Spectrum of Gaussian Random Field. To investigate how Lorentz boosts affect the variance of a Gaussian Random Field, we perform numerical experiments. Specifically, we study the variance of the boosted Gaussian Random Field $a^{\text{obs}} = \exp(\boldsymbol{\eta} \cdot \mathbf{K})\,\mathbf{a}$, which we denote as $\sigma_\ell^2(\boldsymbol{\eta}) = \text{Var}(a_{\ell m}^{\text{obs}})$. This variance can be evaluated numerically by using the equation:

$$\sigma_\ell^2(\boldsymbol{\eta}) := \left(\exp(\boldsymbol{\eta} \cdot \mathbf{K})\,\mathbf{e}_{\ell 0}\right)^T \mathbf{S} \left(\exp(\boldsymbol{\eta} \cdot \mathbf{K})\,\mathbf{e}_{\ell 0}\right)$$

where $\mathbf{e}_{\ell m}$ is a unit vector having $e_{\ell m} = 1$ and zeros elsewhere, $\mathbf{S}$ is a diagonal matrix containing entries C_ℓ, and the Lorentz Boost Lie algebra linear operators $\mathbf{K} = (\mathbf{K}_x, \mathbf{K}_y, \mathbf{K}_z)$ are explicitly defined in Equation 3.3 in Chapter 3.

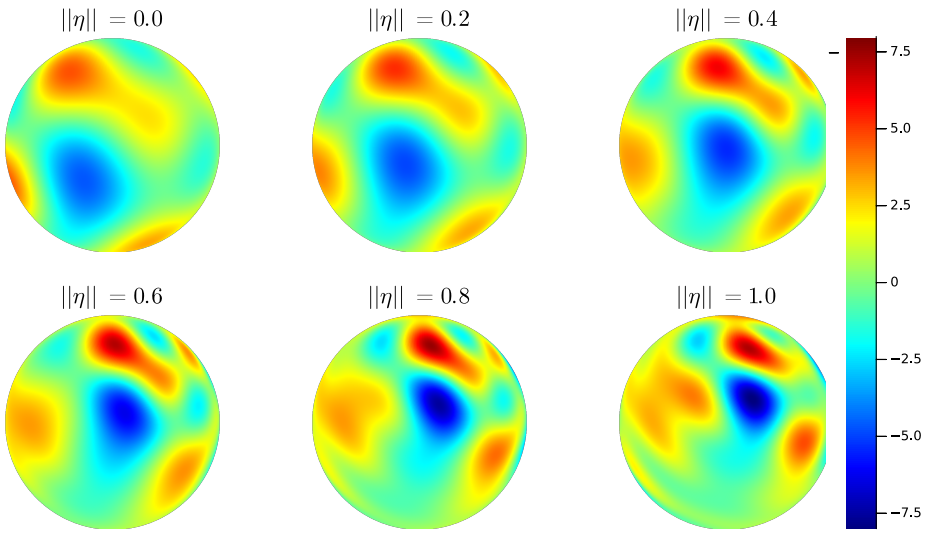


FIGURE 6.7. Impact of using a smaller $\ell_{\max}$ on simulating Lorentz Boost, demonstrating limitations in capturing the effects accurately. Rapidity $\boldsymbol{\eta} = \frac{\|\boldsymbol{\eta}\|}{\sqrt{3}}(x = 1.0, y = 1.0, z = 1.0)$ for $\|\boldsymbol{\eta}\| = 0.0, 0, 2, 0.4, 0.6, 0.8, 1.0$. Top left is the primordial image as it is $\|\boldsymbol{\eta}\| = 0.0$. $(x = 1.0, y = 1.0, z = 1.0)$. If the truncation frequency $\ell_{\max}$ is not sufficiently high, the Lorentz Boost fails to establish the correct relativistic light aberration effect. This is because the fine details of the image, which involve higher frequency information, are not adequately captured.

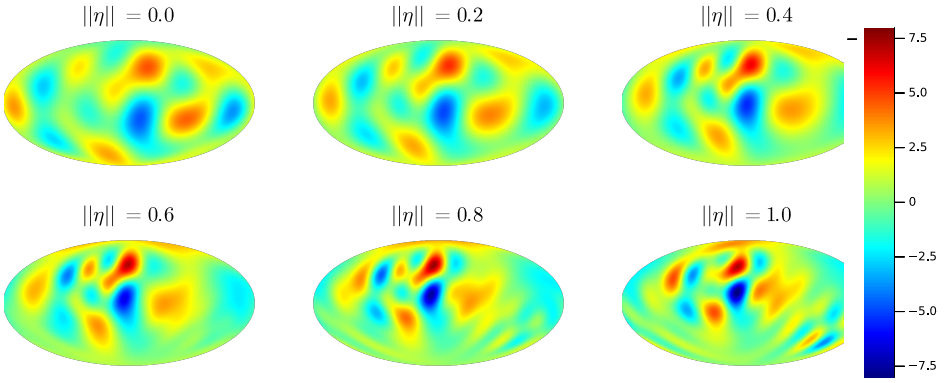


FIGURE 6.8. Mollweide projection using a smaller $\ell_{\max}$ in Figure 6.7, showing its inadequacy in representing Lorentz Boost effects comprehensively. Rapidity $\boldsymbol{\eta} = \frac{\|\boldsymbol{\eta}\|}{\sqrt{3}}(x = 1.0, y = 1.0, z = 1.0)$ for $\|\boldsymbol{\eta}\| = 0.0, 0, 2, 0.4, 0.6, 0.8, 1.0$. Top left is the primordial image as it is $\|\boldsymbol{\eta}\| = 0.0$. $(x = 1.0, y = 1.0, z = 1.0)$.

The Lorentz Boost introduces covariance between the spherical harmonic coefficients $\{a_{\ell m} : m = -\ell, \ldots, \ell\}$ and $\{a_{\ell+1,m} : m = -\ell, \ldots, \ell\}$, but here we focus on examining the variance

$\sigma_\ell^2(\boldsymbol{\eta})$ as a function of the rapidity magnitude $\|\boldsymbol{\eta}\|$. In this study, we use a predefined spectrum $\{C_\ell : \ell = 0, 1, 2, \ldots\}$ described by

$$C_\ell = \begin{cases} \frac{1}{100^2} & \text{if } \ell < 100 \\ \frac{1}{\ell(\ell+1)} + \frac{1}{10}\sin(\frac{2\pi}{50}\ell) & \text{if } 100 \leq \ell \leq 300 \\ \frac{1}{300^2} & \text{if } \ell > 300 \end{cases}$$

When $\boldsymbol{\eta} = \mathbf{0}$, the variance $\sigma_\ell^2(\boldsymbol{\eta})$ simplifies to C_ℓ. The sine function is included to mimic the angular peaks in the CMB power spectrum, which tend to be smoothed out upon applying a Lorentz Boost.

The blue line in Figure 6.9 represents the original power spectrum C_ℓ when $|\boldsymbol{\eta}| = 0$. The orange line shows the variance after applying a Lorentz Boost with a rapidity magnitude of $|\boldsymbol{\eta}| = 0.05$, and the green line corresponds to $|\boldsymbol{\eta}| = 0.1$. As the magnitude of the Lorentz boost increases, the power spectrum becomes smoother, with the peaks being reduced. This smoothing effect is due to the redistribution of power across different ℓ values by the Lorentz Boost.

6.3.2. Truncation of Lorentz Boost. A crucial feature of our formulation is that the matrix representation of the Lorentz Boost, when truncated to a finite dimension, remains unitary. This is beneficial for numerical simulations, which inherently require truncation. The Lorentz Boost operator's unitarity is preserved because its Lie algebra components $\mathbf{K}$ (i.e., $\mathbf{K}_x, \mathbf{K}_y, \mathbf{K}_z$) remain anti-symmetric even after truncation, thereby preserving the L^2 norm of the random field.

Unitarity implies that the transpose of the Lorentz Boost operator serves as its inverse. As a result, Lorentz Boosting leaves the spectrum unchanged when the variance is constant across ℓ. This property is advantageous when incorporating additive white constant Gaussian noise into our models, as it simplifies the associated mathematics.

When the unitary Lorentz Boost operator is applied to a power spectrum that is constant at high frequency ℓ, the spectrum remains the same constant at high frequency ℓ. In this case, the Lorentz Boost operator has no impact, leading to identical outcomes for both finite and infinite ℓ. This serves as an ideal scenario for numerical simulations, as the finite-dimensional representation accurately captures the behavior of the infinite-dimensional system. In our simulations therefore, we assume that the primordial spectrum approaches zero at high frequency ℓ values, ensuring that

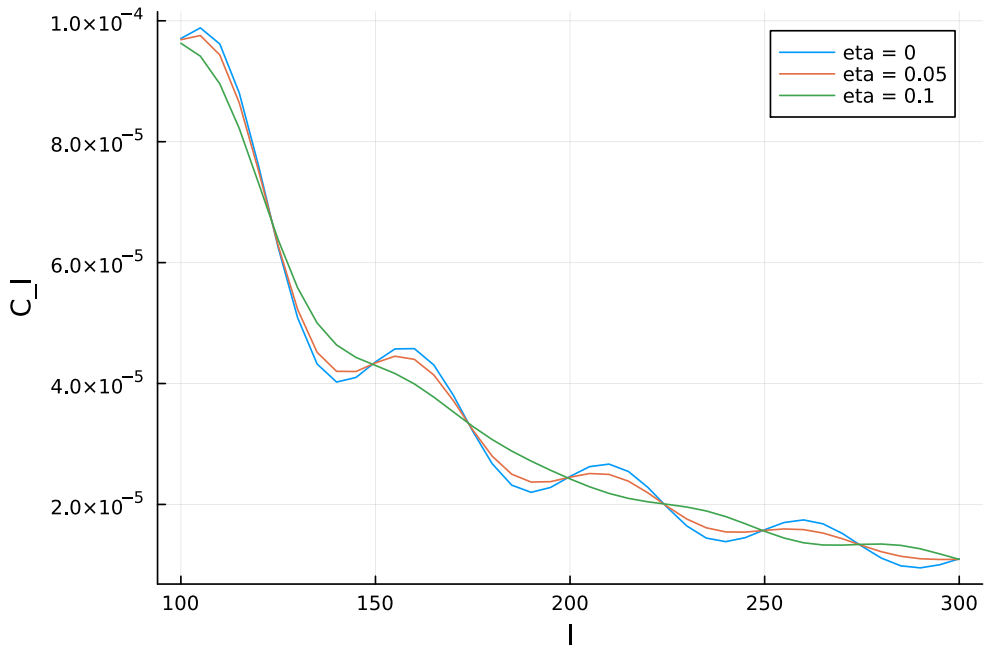


FIGURE 6.9. The effect of Lorentz Boost on the power spectrum C_ℓ for different rapidity magnitudes $\|\boldsymbol{\eta}\|$. The blue line represents the original power spectrum with $\|\boldsymbol{\eta}\| = 0$, showing significant oscillations. The orange line corresponds to $\|\boldsymbol{\eta}\| = 0.05$, demonstrating slight smoothing. The green line for $\|\boldsymbol{\eta}\| = 0.1$ shows pronounced smoothing and reduced peak values, indicative of effective power redistribution across ℓ values.

the total power spectrum (including noise) remains constant at sufficiently high frequency ℓ. By selecting a sufficiently large truncation frequency $\ell_{\max}$ that enters this zone of constant variance, the finite-dimensional simulation becomes an accurate representation of the real, infinite-dimensional situation. To substantiate this, we performed simulations with a Gaussian Random Field truncated at $\ell_{\max} = 1200$. The corresponding C_ℓ values were defined as

$$C_\ell = \begin{cases} \frac{1}{100^2} & \text{if } \ell < 100 \\ \frac{1}{\ell(\ell+1)} & \text{if } 100 \leq \ell \leq 1100 \\ \frac{1}{1100^2} & \text{if } \ell > 1100 \end{cases}$$

The numerical results are summarized in Figure 6.10. The blue line represents $\sigma^2(\boldsymbol{\eta})$ when $\|\boldsymbol{\eta}\| = 0$, which essentially corresponds to the original power spectrum C_ℓ. By design, this value becomes constant after $\ell = 1100$. When the Lorentz Boost is applied with $\|\boldsymbol{\eta}\| = 0.05$, the resulting curve,

shown as the orange line, becomes more linear. Notably, even when the magnitude of $\|\boldsymbol{\eta}\|$ is increased to 0.1, the value of $\sigma^2_\ell(\boldsymbol{\eta})$ remains constant at the tail of the spectrum, specifically at $\ell = 1200$.

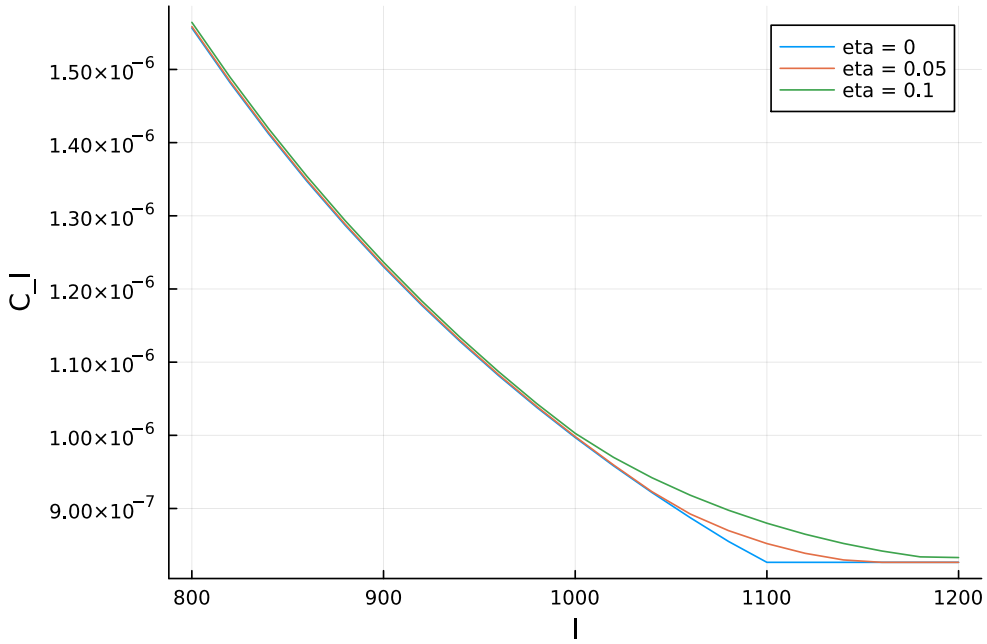


FIGURE 6.10. Impact of Lorentz Boost on a truncated power spectrum C_ℓ. The blue line indicates the variance $\sigma^2(\boldsymbol{\eta})$ at $\|\boldsymbol{\eta}\| = 0$, showing constancy beyond $\ell = 1100$. The orange and green lines represent the variances with Lorentz Boost applied for $\|\boldsymbol{\eta}\| = 0.05$ and $\|\boldsymbol{\eta}\| = 0.1$ respectively, highlighting the preservation of constant variance at the tail.

However, what occurs when truncation does not happen in a region of constant variance? Truncating at $\ell_{\max}$ in the Lorentz Boost Lie algebra $\mathbf{K}$ introduces an artifact, which manifests as an increase in power in C_ℓ at $\ell_{\max}$. The Lorentz Boost redistributes power across the spectrum up to the truncation frequency and causes it to rebound, resulting in a spurious increase in power at $\ell_{\max}$ where the variance accumulates. The maximum likelihood estimator is sensitive to this artifact, as it is based on the probability density function, which is affected by the truncation. Consequently, the maximum likelihood estimator may appear more accurate in simulations than in real-world data, where this truncation artifact is absent. This truncation artifact significantly influences the performance of our Newton Method-based estimator for peculiar velocity. Specifically, this spike

at the truncation frequency disproportionately amplifies the information, causing the estimator to yield an unrealistically accurate estimate of peculiar velocity.

In summary, the Lorentz Boost changes the statistical properties of the CMB Gaussian Random Field in subtle ways. Using a high truncation frequency can reduce these changes, but it's important to be aware of the limits and issues that arise when working in a space with a finite number of dimensions. Future work should aim to develop better methods for estimating peculiar velocity that take these truncation issues into account.

6.4. Root-Mean-Square Error of simulations

In this study, we conduct numerical experiments to assess the effectiveness of our Newton-Method-based estimator for the peculiar velocity in CMB. The synthetic CMB temperature maps are generated using the power spectrum produced by the Code for Anisotropies in the Microwave Background (CAMB). This ensures that our synthetic data align with contemporary scientific models.

We begin by creating simulated primordial CMB temperature maps, where each map is modeled as an isotropic Gaussian random field. The primordial spectrum C_ℓ is assumed to be zero for $\ell \geq \ell_{\max}$. This assumption ensures that the total power spectrum (signal plus noise) is approximately constant for ℓ near $\ell_{\max}$, thus enabling valid truncation of spherical harmonics at this multipole moment $\ell_{\max}$. These simulated maps are then subjected to Lorentz-boosted with varying directions and magnitudes to incorporate the effects of peculiar velocities. Additionally, Gaussian noise with a uniform variance is introduced to mimic realistic observational errors. Finally, we apply our Newton-Method-based estimator to these synthetically generated maps to estimate peculiar velocities.

With the true parameters known from the simulations, we can calculate the Root-Mean-Square (RMS) errors to evaluate the accuracy of our estimates. These errors serve as comprehensive metrics to assess the performance of the proposed estimator and validate our research methodology.

6.4.1. Theoretical RMS. The RMS error of our estimator can be theoretically computed using Fisher information, aligning with the RMS error for the Quadratic Estimator described in Chapter 4, as detailed in Chapter 5. Figure 6.11 shows the theoretical RMS error across various

data resolutions for $\ell_{\max}$ ranging from 400 to 3200. We also vary the noise level with Var(noise) $= \frac{1}{l}$ for l ranging from 200 to 3000. As Figure 6.11 illustrates, the RMS error decreases as the noise level reduces, although a lower limit exists due to the intrinsic uncertainty in the Gaussian random field. Figure 6.12 represents the same information as Figure 6.11, but it scales the data in terms

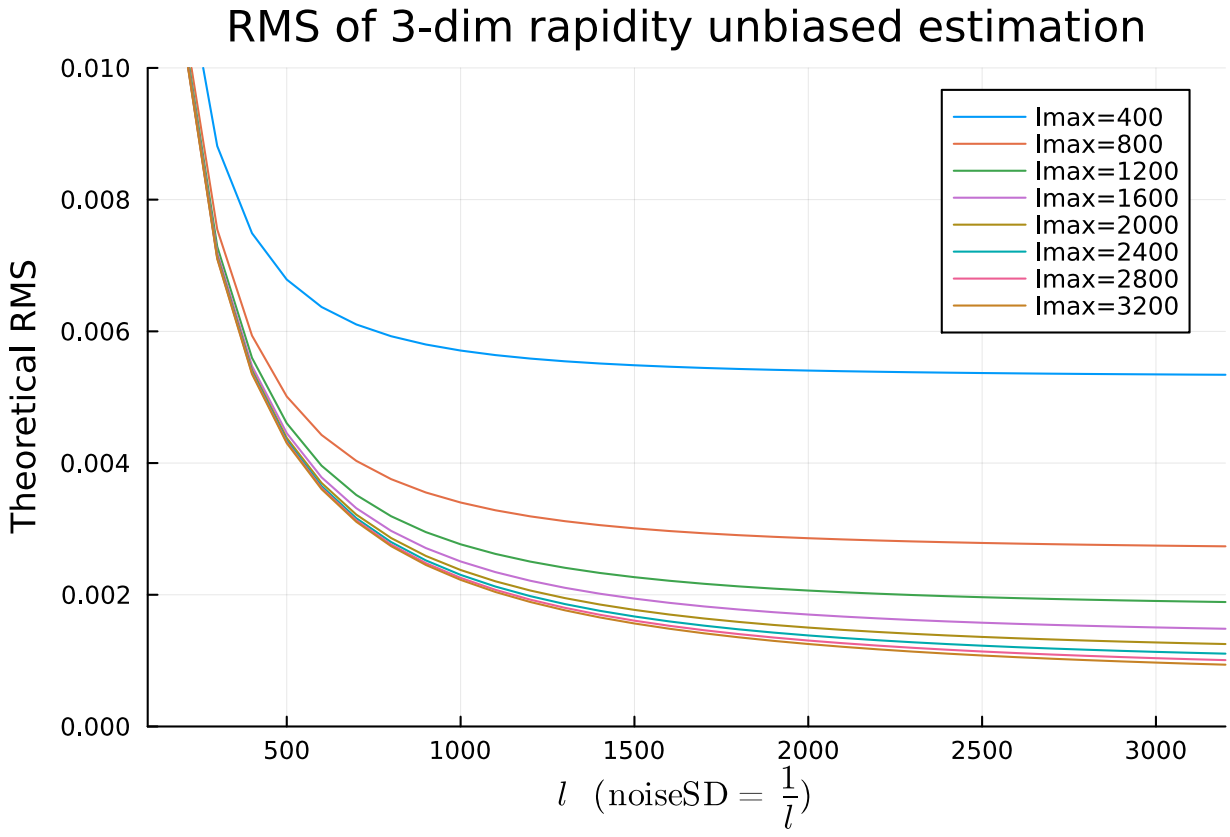


FIGURE 6.11. Theoretical Root-Mean-Square (RMS) error of the proposed Newton-Method-based estimator across various data resolutions. The x-axis indicates different levels of noise variance (Var(noise) $= \frac{1}{l}$) for l ranging from 200 to 3000, and the curves correspond to different maximum multipole moments $\ell_{\max}$. The figure illustrates the diminishing gains in precision as noise decreases, showcasing an intrinsic limit of accuracy dictated by the Gaussian random field.

of "noise RMS". The figure shows the RMS ratio between the estimator and the noise. A steeper slope indicates that the estimator becomes more effective when noise is reducing. Also indicated in the figure, CMB maps with higher $\ell_{\max}$ values provide more efficient RMS reduction.

6.4.2. Empirical RMS. We conduct empirical simulations to measure the RMS error for both the Quadratic Estimator and the Newton method. Figure 6.13 presents the empirical RMS

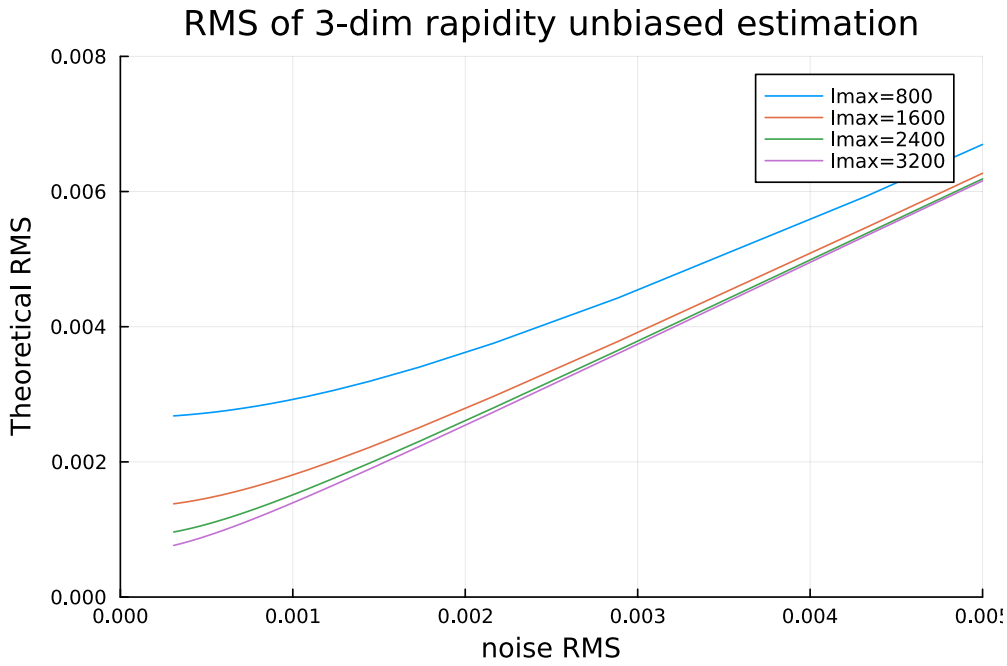


FIGURE 6.12. Ratio of RMS errors of the Newton-Method-based estimator to the noise RMS, evaluated across varying levels of noise and maximum multipole moments ($\ell_{\max}$). A steeper slope indicates greater efficiency in noise reduction. The plot reveals that estimators with higher $\ell_{\max}$ values achieve better efficiency in noise reduction.

error as a function of the norm of the true peculiar velocity, $\|\boldsymbol{\eta}\|$. The shaded regions correspond to two standard errors around the RMS, giving a sense of the uncertainty in our estimates. For small values of $|\boldsymbol{\eta}|$, both the Quadratic Estimator and the Newton method agree well with the theoretical RMS. In particular, our Newton-Method-based estimator performs consistently well across all $\|\boldsymbol{\eta}\|$ values, corroborating the theoretical RMS derived from Fisher information.

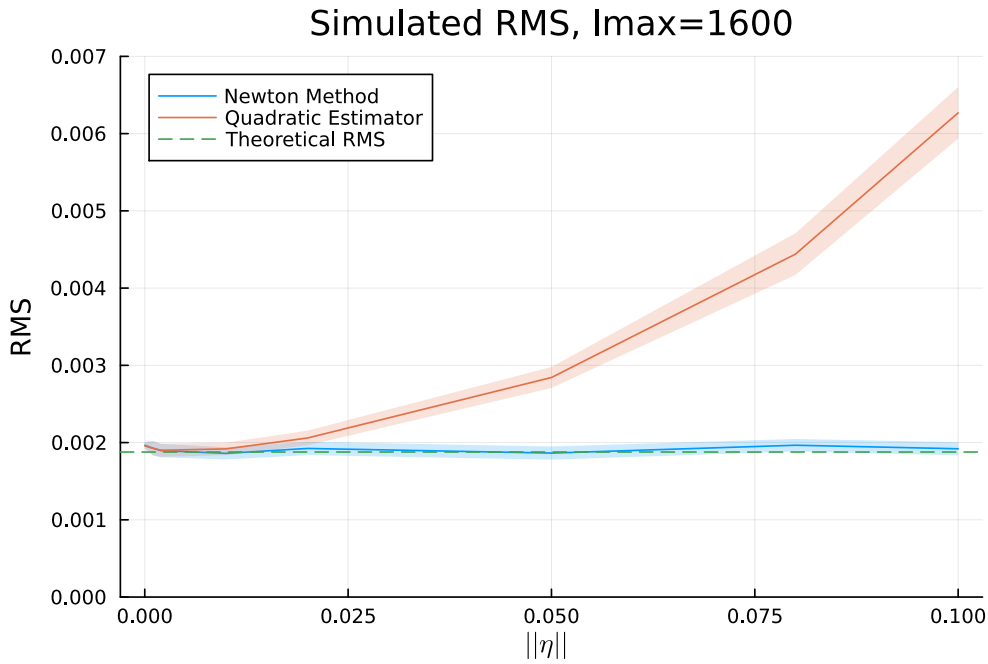


FIGURE 6.13. Empirical Root-Mean-Square (RMS) error based on simulated data, plotted against the norm of the true simulated rapidity ($\|\boldsymbol{\eta}\|$). Shaded areas represent the uncertainty within two standard errors of the sample RMS. The figure demonstrates that while the Quadratic Estimator performs well for smaller rapidity norms, the Newton-Method-based estimator shows consistent performance across all norms, aligning well with the theoretical RMS derived from Fisher information.

Bibliography

[ACM+11] L. Amendola, R. Catena, I. Masina, A. Notari, M. Quartin, and C. Quercellini, *Measuring our peculiar velocity on the CMB with high-multipole off-diagonal correlations*, Journal of Cosmology and Astroparticle Physics **2011** (2011), no. 07, 027–027.

[CvL02] A. Challinor and F. van Leeuwen, *Peculiar velocity effects in high-resolution microwave background experiments*, Phys. Rev. D **65** (2002), 103001.

[DC14] L. Dai and J. Chluba, *New operator approach to the cmb aberration kernels in harmonic space*, Physical Review D **89** (2014), no. 12, 123504.

[DPRW65] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson, *Cosmic black-body radiation.*, The Astrophysical Journal **142** (1965), 414–419.

[Fix09] D. J. Fixsen, *THE TEMPERATURE OF THE COSMIC MICROWAVE BACKGROUND*, The Astrophysical Journal **707** (2009), no. 2, 916–920.

[HL09] D. Hanson and A. Lewis, *Estimators for cmb statistical anisotropy*, Physical Review D **80** (2009), no. 6, 063004.

[HWH+09] G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, N. Jarosik, and et al., *Five-year wilkinson microwave anisotropy probe observations: Data processing, sky maps, and basic results*, The Astrophysical Journal Supplement Series **180** (2009), no. 2, 225–245.

[JRS12] N. Joshi, A. Rotti, and T. Souradeep, *Statistics of bipolar representation of cmb maps*, Phys. Rev. D **85** (2012), 043004.

[KK11] A. Kosowsky and T. Kahniashvili, *Signature of local motion in the microwave sky*, Phys. Rev. Lett. **106** (2011), 191301.

[KR08] P. J. Kostelec and D. N. Rockmore, *Ffts on the rotation group*, Journal of Fourier analysis and applications **14** (2008), no. 2, 145–179.

[LC11] A. Lewis and A. Challinor, *CAMB: Code for Anisotropies in the Microwave Background*, Astrophysics Source Code Library, record ascl:1102.026, February 2011, p. ascl:1102.026, 1102.026.

[Lee12] J. Lee, *Introduction to smooth manifolds*, Graduate Texts in Mathematics, Springer New York, 2012.

[Lin96] C. H. Lineweaver, *The cmb dipole: The most recent measurement and some history*, 1996, astro-ph/9609034.

[LS15] A. Lang and C. Schwab, *Isotropic Gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations*, The Annals of Applied Probability **25** (2015), no. 6, 3047 – 3094.

[LTT18] N. N. Leonenko, M. S. Taqqu, and G. H. Terdik, *Estimation of the covariance function of Gaussian isotropic random fields on spheres, related Rosenblatt-type distributions and the cosmic variance problem*, Electronic Journal of Statistics **12** (2018), no. 2, 3114 – 3146.

[MFS+99] J. C. Mather, D. J. Fixsen, R. A. Shafer, C. Mosier, and D. T. Wilkinson, *Calibrator design for thecobefar-infrared absolute spectrophotometer (firas)*, The Astrophysical Journal **512** (1999), no. 2, 511–520.

[Mor02] V. Moretti, *The interplay of the polar decomposition theorem and the lorentz group*, arXiv preprint math-ph/0211047 (2002).

[PW65] A. A. Penzias and R. W. Wilson, *A measurement of excess antenna temperature at 4080 mc/s.*, The Astrophysical Journal **142** (1965), 419–421.

[PYSS10] T. S. Pereira, A. Yoho, M. Stuke, and G. D. Starkman, *Effects of a cut, lorentz-boosted sky on the angular power spectrum*, 2010, 1009.4937.

[SSM+21] S. Saha, S. Shaikh, S. Mukherjee, T. Souradeep, and B. D. Wandelt, *Bayesian estimation of our local motion from the planck-2018 cmb temperature map*, Journal of Cosmology and Astroparticle Physics **2021** (2021), no. 10, 072.

[Teg97] M. Tegmark, *How to measure cmb power spectra without losing information*, Phys. Rev. D **55** (1997), 5895–5907.