HELSINKI INSTITUTE OF PHYSICS INTERNAL REPORT SERIES

HIP-2005-04

Map-Making and Power Spectrum Estimation for Cosmic Microwave Background Temperature Anisotropies

Torsti Poutanen

Helsinki Institute of Physics, and
Division of Theoretical Physics, Department of Physical Sciences
Faculty of Science
University of Helsinki
P.O. Box 64, FIN-00014 University of Helsinki
Finland

ACADEMIC DISSERTATION
To be presented for public criticism, with the permission of the Faculty of Science of the University of Helsinki, in Auditorium D101 at Physicum, Gustaf Hälltrömin katu 2, on November 24, 2005, at 2 p.m..

Helsinki 2005

ISBN 952-10-1693-0 (printed version)
ISSN 1455-0563
Helsinki 2005
Yliopistopaino

ISBN 952-10-1694-9 (pdf version)
http://ethesis.helsinki.fi
Helsinki 2005
Helsingin yliopiston verkkojulkaisut

Acknowledgements

This thesis is based on research carried out at Helsinki Institute of Physics and at the Theoretical Physics Division of the Department of Physical Sciences at the University of Helsinki during the years 2002-2005. I wish to thank the personnel of both institutes for a pleasant working atmosphere. I am thankful to the Väisälä Foundation and Academy of Finland for their financial support, and CSC - Scientific Computing Ltd. for computational resources.

I thank Prof. Kari Enqvist for his open-mindedness to accept me, who had no previous background in cosmology, to work in his Cosmology Group. I acknowledge his support during the entire time that I have pursued my thesis in his group.

My warmest thanks go to my supervisor and colleague Dr. Hannu Kurki-Suonio. I greatly appreciate his deep interest to my work and the countless hours that he has spent discussing CMB data processing with me. I have felt priviledged to have such a devoted supervisor. His lecture notes on CMB physics were a great help when I wrote Chapter 2 of this thesis.

Dr. Elina Keihänen has been my closest coworker in the practical development of the CMB data processing methods. I thank her for many good advice and numerous tips in Fortran. I thank Ville Heikkilä for his help in CMB simulations and Vesa Muhonen and Jussi Väliviita for discussions on various CMB topics.

I am grateful to Davide Maino for his help in destriping and cooperation in papers **I - III**. I thank Carlo Burigana for cooperation in paper **I**.

I thank the referees of this thesis, Dr. Anthony Challinor and Dr. Ben Wandelt, for the careful reading of the manuscript.

Part of my thesis has benefitted from the collaboration of the Planck Working Group 3 (WG3). I thank Mark Ashdown, Carlo Baccigalupi, Amedeo Balbi, Julian Borrill, Chris Cantalupo, Giancarlo de Gasperis, Kris Górski, Eric Hivon, Charles Lawrence, Paolo Natoli, Martin Reinecke, and Radek Stompor for cooperation in WG3. I acknowledge National Energy Research Scientific Computing Center of U.S Department of Energy for providing computational resources to WG3.

This work has made use of the PLANCK satellite simulation package (Level S), which is assembled by the Max Planck Institute for Astrophysics PLANCK Analysis Centre (MPAC). I acknowledge the use of the CMBFAST code for the computation of the theoretical CMB angular power spectra. Some of the results in this thesis have been derived using the HEALPix package ([51],[52]).

Helsinki 2005

Torsti Poutanen

Poutanen, Torsti: Map-Making and Power Spectrum Estimation for Cosmic Microwave Background Temperature Anisotropies, University of Helsinki, 2005, 144 p., Helsinki Institute of Physics Internal Report Series, HIP-2005-04, ISSN 1455-0563, ISBN 952-10-1693-0 (printed version), ISBN 952-10-1694-9 (pdf version).
Classification (INSPEC): A9575P, A9870V, A9880L
Keywords: cosmology, cosmic microwave background radiation, data analysis

Abstract

The universe is filled with cosmic microwave background (CMB) radiation. We can detect these photons in every direction with nearly the same intensity. The tiny intensity variation is called the CMB temperature anisotropy and it reflects the inhomogeneities of the universe during the photon decoupling. Cosmological theories, that describe the early universe, predict the statistical properties of these anisotropies. Therefore CMB observations are important when determining the values of the cosmological parameters of these theories. In this thesis we restrict ourselves to temperature observations and do not consider polarization here.

The CMB temperature anisotropy can be displayed as a pixelized map over the celestial sphere. Map-making and angular power spectrum estimation are important steps in the data processing of a CMB experiment. Destriping is an efficient method to reduce the level of correlated $(1/f)$ noise in the observed data. In this thesis we have developed a maximum-likelihood approach to destriping and use these methods to make maps from the simulated observations of the Low Frequency Instrument (LFI) of the Planck satellite. We compare these output maps to the output maps of general least squares (GLS) map-making algorithms. Under the assumption of Gaussian distributed noise GLS algorithms are implementations of the maximum-likelihood map-making. Therefore their output maps fall close to the minimum variance map. Destriped maps are not optimal in this maximum-likelihood sense. Our results reveal, however, that the difference in the map noise between destriped and GLS output maps is very small in these cases. The map-making methods cause error in the signal part of the output maps too. The source of this error is the subpixel structure of the signal. Its coupling to the output map varies in different map-making methods. This error was larger for GLS than for destriping, but in both cases it was clearly smaller than the level of the noise in the maps.

In this thesis we also studied the angular power spectrum estimation from the output maps of destriping. We found that the map-making error due to pixelization noise had an insignificant effect in the power spectrum estimates. We noticed, however, that the non-uniform distribution of observations in the output map pixels caused high-ℓ excess power in the power spectrum estimates. We corrected for this by subtracting a signal bias whose value we estimated using Monte Carlo simulations. The angular power spectrum estimates, that we obtained, were unbiased and their errors were close to their theoretical expectations.

List of Papers

This thesis consists of three research papers and an introduction, which provides relevant background and additional discussions.

I E. Keihänen, H. Kurki-Suonio, T. Poutanen, D. Maino, and C. Burigana

A MAXIMUM LIKELIHOOD APPROACH TO THE DESTRIPING TECHNIQUE

A&A **428**, 287-298 (2004). Reference [1] in the bibliography.

II T. Poutanen, D. Maino, H. Kurki-Suonio, E. Keihänen, and E. Hivon

COSMIC MICROWAVE BACKGROUND POWER SPECTRUM ESTIMATION WITH THE DESTRIPING TECHNIQUE

MNRAS **353**, 43-58 (2004). Reference [2] in the bibliography.

III T. Poutanen, G. de Gasperis, E. Hivon, H. Kurki-Suonio, A. Balbi, J. Borrill, C. Cantalupo, O. Doré, E. Keihänen, C. Lawrence, D. Maino, P. Natoli, S. Prunet, R. Stompor, and R. Teyssier

COMPARISON OF MAP-MAKING ALGORITHMS FOR CMB EXPERIMENTS

To be published in A&A, [astro-ph/0501504] (2005). Reference [3] in the bibliography.

The contribution of the author to the publications

Paper I: This is the first paper that Keihänen, Kurki-Suonio and I published on CMB data processing. Although the destriping codes used in our paper were not written by me, I had made my own test codes using similar principles. That experience allowed me to contribute in the development of the destriping theory and in the analysis of the results presented in the paper. I made inputs to the maximum-likelihood and to the eigenvalue analysis of the destriping problem.

Paper II: In this paper we studied the angular power spectrum estimation from the CMB temperature maps that were made using destriping. The items affecting

the accuracy of the power spectrum estimates were analyzed using Monte Carlo simulations. I implemented the Monte Carlo pipeline, ran all the simulations and calculated the results shown in the paper. The destriping code used in the study was a code described in paper [1]. I wrote most of the the paper and developed the theory in its Appendix A.

Paper III: This paper is a result of the Planck WG3 collaboration. The maps for this study were made by research groups participating in WG3. I made the destriped maps using the code described in paper [1]. The methods used in the map and angular power spectrum comparison were mainly designed by me and I produced all the results. I wrote large parts of the paper and I developed a significant part of the theory in its Appendices A and B.

Contents

Acknowledgements iii

Abstract v

List of Papers vi

1 Introduction 1

2 CMB Physics 3
2.1 Introduction . . . 3
2.2 Perturbed Universe . . . 4
2.3 Boltzmann Equation for the CMB Photons . . . 9
2.3.1 Photon Distribution and Brightness Functions . . . 9
2.3.2 Collisionless Boltzmann Equation . . . 10
2.3.3 Collision Term . . . 12
2.3.4 Boltzmann Hierarchy . . . 14
2.4 Boltzmann Equations for the Other Particle Species . . . 16
2.5 C_ℓ Spectrum . . . 17
2.6 Line-of-Sight Integration . . . 19
2.7 C_ℓ Spectrum Today . . . 21
2.7.1 Large Scales - Sachs-Wolfe Effect . . . 23
2.7.2 Small Scales - Acoustic Oscillations . . . 23
2.8 C_ℓ and Cosmological Parameters . . . 27

3 Overview of a Satellite CMB Experiment 31
3.1 Instrumentation . . . 31
3.1.1 Low Frequency Instrument . . . 34
3.1.2 High Frequency Instrument . . . 37
3.1.3 Telescope Beams . . . 37
3.1.4 Detected Signal . . . 41
3.1.5 Radiometer Noise Model . . . 44
3.1.6 Radiometer Performance Goals . . . 46
3.2 Scanning Strategy . . . 46
3.3 Foregrounds . . . 51

3.3.1 Satellite Motion Induced Signal . . . 51
3.3.2 Galactic Emissions . . . 51
3.3.3 Extragalactic Emissions . . . 52
3.4 Time-ordered Data . . . 53
3.5 Data Processing . . . 56
3.5.1 Data Processing Pipeline . . . 56
3.5.2 Sky Pixelization . . . 57
3.5.3 Level S . . . 57

4 Map-making **60**
4.1 Map-making Problem . . . 60
4.2 Minimum Variance Map . . . 61
4.3 GLS Map-making Algorithms . . . 63
4.4 Destriping . . . 66
4.4.1 Destriping Principle . . . 66
4.4.2 Maximum Likelihood Approach to Destriping . . . 69
4.4.3 Covariance of the Basefunction Amplitudes . . . 75
4.5 Comparison of GLS and Destriping . . . 77
4.5.1 Relation between GLS and Destriping Algorithms . . . 78
4.5.2 Comparison of Temperature Maps . . . 80
4.6 Deconvolution Map-making and Destriping . . . 84

5 Angular Power Spectrum Estimation **89**
5.1 Estimation Problem . . . 89
5.2 Maximum Likelihood Estimator . . . 90
5.3 Quadratic Estimator . . . 92
5.4 Pseudo-C_ℓ Estimator . . . 95
5.5 Power Spectrum Estimation and Destriping . . . 97
5.5.1 Simulation Description . . . 97
5.5.2 Mode Coupling Matrix . . . 101
5.5.3 Pixel Window and Pointing Distribution Effects . . . 104
5.5.4 Noise Bias . . . 109
5.5.5 Covariance Matrix of the Power Spectrum Estimate . . . 112

6 Conclusions **116**

A Quadratic Power Spectrum Estimator for Signal and Noise Dominated Maps **118**
A.1 Signal Dominated Map . . . 118
A.2 Noise Dominated Map . . . 121

B Beam Deconvolution **122**

Bibliography **127**

List of Figures

2.1 Acoustic oscillations of the photon brightness function 26
2.2 C_ℓ and cosmological parameters - 1 28
2.3 C_ℓ and cosmological parameters - 2 29

3.1 Schematic view of the PLANCK satellite 32
3.2 Detailed view of the PLANCK focal plane 33
3.3 Foot-print of the focal plane in the sky 34
3.4 Front-end unit of the LFI radiometer 36
3.5 Back-end unit of the LFI radiometer 36
3.6 $1/f$ noise reduction in the LFI radiometer 37
3.7 HFI focal plane unit 38
3.8 Elliptic Gaussian beam pointing to the north pole 40
3.9 Spherical harmonic expansion coefficients of an elliptic Gaussian beam pointing to the north pole 41
3.10 Radiometer noise PSD 45
3.11 Radiometer noise stream 47
3.12 PLANCK in L2 orbit 49
3.13 Coverage maps of two scanning strategies 50
3.14 Map of the CMB dipole and galaxy 52
3.15 TOD of the CMB dipole and galaxy 53
3.16 Intensities of the galactic and CMB emissions 54
3.17 Observed TOD 56
3.18 Data processing pipeline 58

4.1 Noise filters 64
4.2 Uniform baselines modelling the $1/f$ noise 67
4.3 Noise maps before and after destriping 72
4.4 Angular power spectra of the instrument noise maps 73
4.5 Observed temperature map after destriping 80
4.6 Map-making error due to pixelization noise 84

5.1 Hit count maps in the power spectrum study 100
5.2 Angular power spectra of the sky coverage maps 102
5.3 Mode coupling matrix 103

5.4 Spectrum of the binned noiseless map 108
5.5 Ratio of power spectra with signal bias removed 109
5.6 Noise bias 111
5.7 Normalized estimator variance for different pixel weights 115

List of Tables

2.1 Cosmological parameters 27

3.1 Antenna temperature of the CMB photons 43
3.2 Goals of the PLANCK payload 48
3.3 Amount of data produced by the PLANCK detectors 55
3.4 Angular sizes of the HEALPix pixels 59

4.1 Simulation parameters of the map comparison 81
4.2 Results of the comparison of the temperature maps 82
4.3 Signal component of the reconstruction error map 83

5.1 Simulation parameters of the power spectrum estimation 99

Chapter 1

Introduction

The universe is filled with cosmic microwave background (CMB) radiation. We can detect these photons in every direction with nearly the same intensity. They have a blackbody spectrum of mean temperature $T_0 = 2.725$ K ([4]). The CMB radiation was first observed in 1965 ([5]). There is a small temperature difference δT between photons coming from different directions. This is called the CMB temperature anisotropy. Its typical magnitude is $\delta T/T_0 \lesssim 10^{-4}$. The CMB anisotropy was first observed by COBE satellite in 1992 ([6]).

The CMB photons were generated during photon decoupling, where photons became free from the baryon interactions due to the formation of stable atoms. Photon decoupling took place about 380 000 years after the big bang ([7]). The CMB temperature anisotropy, that we see today, reflects the inhomogeneities of the universe during the photon decoupling and, to a small extent, also later inhomogeneities. Theories, that describe the early universe, predict the statistical properties of the CMB anisotropies that we observe today. Therefore CMB observations are important when we choose between these theories and determine the values of the cosmological parameters.

The CMB temperature anisotropy is usually displayed in a pixelized map over the celestial sphere. The angular power spectrum of that map is a quantity whose value is predicted by the cosmological theories. Map-making and angular power spectrum estimation from the observed data are important steps in the data processing of a CMB experiment. In this thesis we restrict ourselves to the map-making and power spectrum estimation from the temperature anisotropy observations. We do not consider polarization here.

We introduce the reader to the CMB data processing from two directions. In Chapter 2 we give a brief discussion on CMB physics and show how the angular power spectrum, that we observe today, depends on the inhomogeneities of the universe during photon decoupling. In Chapter 3 we describe the relevant aspects of a satellite CMB experiment that we need to know in order to understand the observed data that we use as an input in the map-making. We use the Planck experiment as an example here ([8], [9]).

The noise of the detectors is an important source of error in a CMB experiment. In this thesis we assume that the detector noise is a sum of correlated $(1/f)$ and non-correlated (white) noise. The correlated part of the noise, when coupled to the observing strategy, will lead to stripes in the final maps. Destriping is a map-making method which, in its simplest form, models the correlated noise with uniform baselines and uses the observations of pixels, that are monitored several times during the mission, to solve the estimates of the baseline magnitudes. When the baselines have been subtracted from the observed data, the cleaned data can be averaged (binned) to the output map pixels.

Destriping was first introduced to the CMB map-making in the early planning phases of the PLANCK experiment ([68]). It is an efficient map-making algorithm which requires no prior knowledge on the instrument noise. However, it is not optimal in the sense that it does not produce a minimum variance map. Under the assumption of Gaussian distributed noise the maps produced by the general least squares (GLS) map-making algorithms ([56] - [61]) produce maps that fall closer to the minimum variance map.

In this thesis we developed a maximum likelihood approach to destriping and implemented map-making codes that use these principles (paper **I** [1]). Compared to the destriping methods of other authors ([69] - [73]) our method differs in the weights it gives to the observed pixels. Our weights arise from the maximum-likelihood analysis while the weights of the other authors are more heuristic. We compare in paper **I** ([1]) the maps of our weights to the maps of the weights of the other authors. Using simulated observations of a PLANCK detector we also compared the output maps of our implementation of the destriping method and two implementations of the GLS method (paper **III** [3]).

A discussion of the map-making problem and map-making algorithms is given in Chapter 4 of this thesis. The results of the above map comparisons are given there as well.

In this thesis we also studied angular power spectrum estimation from the output maps of destriping (paper **II** [2]). We used Monte Carlo methods to reveal the accuracy of our power spectrum estimates. A thorough discussion of the angular power spectrum estimation and the results of the study (paper **II** [2]) are given in Chapter 5 of this thesis. Appendices A and B provide some further details to Chapter 5.

Chapters 2 and 3 are introductions to the map-making and angular power spectrum estimation, which are the main research topics of this thesis. They are discussed in Chapters 4 and 5 and in papers **I - III** of this thesis. We did not carry out any research on CMB physics in this thesis. Therefore Chapter 2 just gives a brief overview of this topic.

Chapter 2

CMB Physics

2.1 Introduction

When the universe had cooled to $T \approx 3000$ K, protons and electrons were able to form stable hydrogen atoms. Before that all matter was in a plasma state and the photon mean free path was short due to Compton scattering from the charged particles and ions (atomic nuclei). Frequent scattering maintained the photons and baryons in a thermal equilibrium and they can be described as a single "baryon-photon" fluid at that time. The build-up of hydrogen atoms decreased the density of free electrons and the mean free path of the photons was rapidly increased allowing them to propagate nearly freely. The epoch when the photons became free from the baryon-photon fluid is called the *photon decoupling*. It occurred about 380 000 years after the big bang at redshift $z_{\mathrm{dec}} \approx 1089$ ([7]). It was not an instantaneous event but it lasted about 120 000 years ([7]). The value of a quantity at the photon decoupling is denoted with a subscript "dec" in this chapter.

Photons liberated at decoupling have propagated to us and we can detect them today. We can see them coming to us from all directions with a nearly constant intensity. The stream of these photons is the CMB radiation. Before the CMB was detected by Penzias and Wilson in 1965 ([5]) its existence was theoretically predicted (see e.g. [10] - [14]). The mean temperature of the CMB photons, that we see today, has redshifted to $T_0 = 2.725$ K ([4]).

Already before the detection of the CMB it was anticipated that the energy density (ρ) of the early universe (before photon decoupling) was not perfectly homogeneous but it had some small non-homogeneous perturbations (see e.g. [15] - [17])

$$\rho(t, \mathbf{x}) = \bar{\rho}(t) + \delta\rho(t, \mathbf{x}). \tag{2.1}$$

Here $\bar{\rho}(t)$ is the homogeneous part of the energy density. It has the same value in every space-point of the universe and it depends only on time (t). The non-homogeneous energy-density perturbation $\delta\rho(t, \mathbf{x})$ depends both on space (indicated by 3-vector $\mathbf{x}$) and time. It is assumed that the energy density perturbations were

small ($|\delta\rho(t,\mathbf{x})| \ll \bar{\rho}(t)$) in the early universe. The regions of energy overdensity of the early universe have grown to the galaxies that we see today.

The energy density perturbations caused perturbations in the intensity of CMB photons. These CMB perturbations were predicted theoretically (see e.g. [15] - [17]) and they were detected by the COBE satellite in 1992 ([6]). According to the COBE observations the CMB photons coming to us from different directions have slightly different temperatures. The relative differences are $10^{-5} \dots 10^{-4}$. This variation (relative to the mean temperature T_0) is called the CMB anisotropy signal. It has a distinct value in every point of the celestial sphere and the observations of the CMB anisotropy field are usually displayed as a map over the sphere.

It is widely believed that the universe underwent a period of an exponential expansion in its early history ($\lesssim 10^{-32}$ s after the big bang). This expansion is called inflation. One or more scalar fields (inflaton fields) could have caused the rapid expansion. Vacuum fluctuations of those fields lead to perturbations in the energy-momentum tensor. The inflation model predicts that the energy density perturbations were initiated during the inflation period ($\sim 10^{-32}$ s after the big bang).

In this chapter of this thesis we apply units where $c = \hbar = k_{\mathrm{B}} = 1$.

2.2 Perturbed Universe

Because we assume that the early universe was nearly homogeneous with some small non-homogeneous perturbations, it can be described with a metric $g_{\mu\nu}$ that is a sum of a metric of a homogeneous spacetime and a small non-homogeneous perturbation

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}. \tag{2.2}$$

The homogeneous universe with the metric $\bar{g}_{\mu\nu}$ is called the background universe. The metric of the perturbed universe is $g_{\mu\nu}(t,\mathbf{x})$.

Because we assume that the perturbations are small (relative to their background values), our perturbed universe can be approximately described by first order perturbation theory of general relativity. In this theory every quantity is a sum of its background quantity and a small perturbation: $h = \bar{h} + \delta h$. We denote with an overbar the quantities of the background universe. They depend on time only. We will drop from all equations the terms which are of order $\mathcal{O}(\delta h^2)$ or higher. Thus the perturbation of a quantity depends on the perturbations of the other quantities through linear equations.

In general relativity the development of the universe is determined by the Einstein equations

$$G^{\mu}_{\nu} = 8\pi G T^{\mu}_{\nu}. \tag{2.3}$$

Here T^{μ}_{ν} is the energy-momentum tensor, G is the gravitational constant and G^{μ}_{ν} is the Einstein tensor. The latter depends on the metric tensor and its first and second derivatives with respect to the spacetime coordinates x^{μ}. The Einstein equations are valid separately in the perturbed and in the background universes. Subtracting the

Einstein equations of these universes from each other leads to the Einstein equations between the perturbations δG^μ_ν and δT^μ_ν.

The fact that $G^\mu_{\nu;\mu} = 0$ (subscript ";μ" means a covariant derivation with respect to x^μ) leads to energy-momentum continuity equations $T^\mu_{\nu;\mu} = 0$. They apply in both perturbed and background universes. Energy-momentum continuity equations can be used instead of some of the Einstein equations.

In this chapter we assume, for simplicity, that the background universe is a flat Friedmann-Robertson-Walker (FRW) universe. Its metric in comoving coordinates (t, x, y, z) is

$$ds^2 = \bar{g}_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a^2(t)[dx^2 + dy^2 + dz^2]. \tag{2.4}$$

In this chapter we apply a summation rule where we sum over the repeated indeces (Einstein summation rule). The function $a(t)$ is the scale factor.

We will often use the conformal time η instead of the coordinate time t. They are related by $d\eta = dt/a(t)$. Using conformal time the background metric is

$$ds^2 = a^2(\eta)[-d\eta^2 + dx^2 + dy^2 + dz^2]. \tag{2.5}$$

We assume that the background energy-momentum tensor can be approximated by the perfect fluid energy-momentum tensor

$$T^\mu_\nu = diag(-\bar{\rho}, \bar{p}, \bar{p}, \bar{p}), \tag{2.6}$$

where $\bar{p}$ is the background pressure of the fluid. The equations for the scale factor can be obtained from the $0-0$ and $i-i$ components of the background Einstein equations. These equations (for the scale factor) are called the Friedmann equations. We give them here using the conformal time.

$$\mathcal{H}^2 \equiv \left(\frac{a'}{a}\right)^2 = \frac{8\pi G}{3}\bar{\rho}a^2 \tag{2.7}$$

and

$$\mathcal{H}' = -\frac{4\pi G}{3}(\bar{\rho} + 3\bar{p})a^2, \tag{2.8}$$

where $' \equiv d/d\eta$ and $\mathcal{H} \equiv a'/a$ is the comoving Hubble parameter. Its relation to the ordinary Hubble parameter H is $\mathcal{H} = aH$. Unless otherwise noted we normalize the scale factor to have value $a_0 = 1$ today (subscript "0" of a quantity means its value today). In this normalization a comoving value of a quantity equals its value today. The continuity equation $\bar{T}^\mu_{0;\mu} = 0$ of the background universe gives

$$\bar{\rho}' = -3\mathcal{H}(\bar{\rho} + \bar{p}). \tag{2.9}$$

A thorough discussion of the perturbation theory in the FRW universe is given in e.g. [18] - [20]. We just give some main results here.

The perturbations in the metric and energy-momentum tensors can be scalar, vector or tensor type depending on how they behave in the 3-space coordinate rotations. In linear perturbation theory these three types of perturbations evolve independent from each other. The scalar perturbations are the most important, because they couple to the energy density and pressure perturbations and are responsible for the gravitational growth of the overdense regions of the early universe. We will consider only the scalar perturbations in this chapter and ignore the vector and tensor perturbations.

In the scalar perturbation theory the perturbation B_i of a 3-vector quantity (i indexes the space coordinates) is derived from a scalar perturbation B: $B_i = -B_{,i}$ ("$,i$" is a partial derivative with respect to x^i). The perturbation E_{ij} of a 3-space traceless tensor quantity is derived from a scalar perturbation as well: $E_{ij} = E_{,ij} - \frac{1}{3}\delta_{ij}\nabla^2 E$. Here δ_{ij} is the Kronecker delta.

There are several coordinate systems in the perturbed universe that are close to each other and that we could therefore use. Transformations between such coordinate systems are called gauge transformations. For practical calculations we need to select a coordinate system. In this thesis we use the *conformal-Newtonian* gauge. In this gauge the metric is

$$ds^2 = a^2(\eta)[-(1+2\Phi)d\eta^2 + (1-2\Psi)(dx^2+dy^2+dz^2)]. \tag{2.10}$$

Here Φ and Ψ are small perturbations from the background metric. They are functions of the spacetime coordinates $(\eta, \mathbf{x})$ and are called Bardeen potentials ([19]).

The energy-momentum tensor of the perturbed universe is (scalar perturbations in the conformal-Newtonian gauge)

$$T^\mu_\nu = \begin{bmatrix} -\bar{\rho} & \\ & \bar{p}\delta_{ij} \end{bmatrix} + \begin{bmatrix} -\delta\rho & -(\bar{\rho}+\bar{p})v_{,i} \\ (\bar{\rho}+\bar{p})v_{,i} & \delta p\delta_{ij} + \bar{p}\Pi_{,ij} - \frac{1}{3}\bar{p}\delta_{ij}\nabla^2\Pi \end{bmatrix} \tag{2.11}$$

Here index μ refers to the rows and index ν refers to the columns of the energy-momentum tensor. The quantities $\delta\rho$ and δp are the energy density and pressure perturbations and $-v_{,i}$ is the velocity perturbation (derived from the scalar perturbation v) of the fluid. The background value of the velocity perturbation is zero. The quantity $\Sigma_{ij} = \bar{p}\Pi_{,ij} - \frac{1}{3}\bar{p}\delta_{ij}\nabla^2\Pi$ is the anisotropic stress. Its background value is zero as well.

Because we assume a flat background universe, the perturbations can be Fourier expanded using the plane waves $e^{i\mathbf{k}\cdot\mathbf{x}}$ (complete orthogonal set in a flat universe). We can write for the energy density perturbations ([20])

$$\delta\rho(\eta,\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int d^3k\,\delta\rho(\eta,\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{2.12}$$

Here $\mathbf{k}$ is the comoving wavevector. Its magnitude k ($k \equiv |\mathbf{k}|$) can be given in terms of the comoving wavelength λ: $k = 2\pi/\lambda$. The physical wavelength is $\lambda_{\text{phys}} = a\lambda$ and the physical wavevector is $\mathbf{k}_{\text{phys}} = \mathbf{k}a^{-1}$. The physical wavelength of a Fourier mode

$\mathbf{k}$ increases as the universe expands. The comoving wavelength of a Fourier mode $\mathbf{k}$ is called a *scale*. The unit vector in the direction of $\mathbf{k}$ is $\hat{\mathbf{k}}$ ($\mathbf{k} = k\hat{\mathbf{k}}$) and $\hat{k}^i$ is the component of $\hat{\mathbf{k}}$ in the x^i direction. Similar Fourier expansions exist for δp, Ψ and Φ.

Following [20] we define for the Fourier expansions of the scalar velocity perturbation v and the scalar anisotropic stress Π

$$v(\eta, \mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int d^3k \frac{v(\eta, \mathbf{k})}{k} e^{i\mathbf{k}\cdot\mathbf{x}} \tag{2.13}$$

and

$$\Pi(\eta, \mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int d^3k \frac{\Pi(\eta, \mathbf{k})}{k^2} e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{2.14}$$

This way $v(\eta, \mathbf{k})$ and $\Pi(\eta, \mathbf{k})$ will have the same dimensions and magnitudes as the perturbations ($v_i = -v_{,i}$ and $\Sigma_{ij}/\bar{p}$) themselves have. In first order perturbation theory each Fourier mode evolves independently. We can find a solution to each Fourier mode separately and obtain the total perturbations from Eqs. (2.12) - (2.14).

We can now apply the Fourier expansion to the Einstein equations and we obtain for the Fourier modes of the perturbations

$$\mathcal{H}^{-1}\Psi' + \Phi + \frac{1}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \Psi = -\frac{1}{2}\delta \tag{2.15}$$

$$\mathcal{H}^{-1}\Psi' + \Phi = -\frac{3}{2}(1+w)\frac{\mathcal{H}}{k}v \tag{2.16}$$

$$\mathcal{H}^{-2}\Psi'' + \mathcal{H}^{-1}(\Phi' + 2\Psi') - 3w\Phi - \frac{1}{3}\left(\frac{k}{\mathcal{H}}\right)^2 (\Phi - \Psi) = \frac{3}{2}\frac{\delta p}{\bar{\rho}} \tag{2.17}$$

$$\left(\frac{k}{\mathcal{H}}\right)^2 (\Psi - \Phi) = 3w\Pi \tag{2.18}$$

We have defined the density contrast $\delta \equiv \delta\rho/\bar{\rho}$ and the equation-of-state parameter $w \equiv \bar{p}/\bar{\rho}$. For later convenience we also define the "speed of sound" c_s of the fluid: $c_\mathrm{s}^2 \equiv \bar{p}'/\bar{\rho}'$. Note that all perturbations are functions of $(\eta, \mathbf{k})$.

We can Fourier expand the continuity equations as well and obtain the following evolution equations for the Fourier modes of the perturbations

$$\delta' = (1+w)(-v + 3\Phi') + 3\mathcal{H}\left(w\delta - \frac{\delta p}{\bar{\rho}}\right) \tag{2.19}$$

$$v' = -\mathcal{H}(1-3w)v - \frac{w'}{1+w}v + \frac{\delta p}{\bar{\rho}+\bar{p}} - \frac{2}{3}\frac{w}{1+w}\Pi + \Phi \tag{2.20}$$

The fluctuations in the local curvature of the perturbed universe are characterised by the *curvature perturbation* $\mathcal{R}(\eta, \mathbf{x})$. In the conformal-Newtonian gauge it is defined as

$$\mathcal{R} = -\Psi - \frac{2}{3(1+w)}(\mathcal{H}^{-1}\Psi' + \Phi). \tag{2.21}$$

It gives the local curvature of a $t = const$ slice in the comoving gauge, i.e. one where the fluid flow is orthogonal to this slice. Although we do not work here in the comoving gauge, $\mathcal{R}$ is a useful quantity to describe primordial perturbations, since, as explained below, for adiabatic perturbations it remains constant at scales outside horizon ("outside horizon" means here that $\lambda \gg \mathcal{H}^{-1}$ which is equivalent to $k \ll \mathcal{H}$).

Using Einstein equations an evolution equation can be derived for the Fourier modes of the curvature perturbation

$$\frac{3}{2}(1+w)\mathcal{H}^{-1}\mathcal{R}' = \left(\frac{k}{\mathcal{H}}\right)^2 \left[c_{\mathrm{s}}^2\Psi + \frac{1}{3}(\Psi - \Phi)\right] + \frac{9}{2}c_{\mathrm{s}}^2(1+w)\mathcal{H}\left(\frac{\delta p}{\bar{p}'} - \frac{\delta\rho}{\bar{\rho}'}\right). \quad (2.22)$$

In this thesis we assume that the perturbations are initially adiabatic. Adiabatic perturbations are predicted by the simple inflation models and are consistent with observations. For adiabatic perturbations $\delta p/\bar{p}' = \delta\rho/\bar{\rho}'$ and they remain adiabatic while outside horizon. Thus the second term on the right hand side of Eq. (2.22) is zero.

At the end of the inflation all relevant scales were outside horizon. For these scales the first term on the right hand side of Eq. (2.22) is small. Therefore, for the adiabatic curvature perturbations, whose scales are outside horizon, $\mathcal{R}' = 0$. This means that the curvature perturbations of those scales remain constant (in time) as long as they are outside horizon.

To solve the perturbations of the interesting quantities at some time (e.g. today) we need to set the initial conditions. They are usually specified during the early radiation-dominated era ($t = t_{\mathrm{rad}}$), when all interesting scales were well outside the horizon ($k \ll \mathcal{H}$). We can assume that at this time neutrino decoupling, $\mathrm{e}^+\mathrm{e}^-$ annihilation and nucleosynthesis were over and the temperature of the universe was, e.g., T $\approx 10^7$ K. Because all relevant scales are outside horizon, the Fourier modes of the curvature perturbation have constant values which we denote by $\mathcal{R}_{\mathbf{k}}(\mathrm{rad})$. They are called the *primordial* values of the curvature perturbation. They are our initial conditions. The magnitude of the first order perturbation of a quantity at some time later than t_{rad} depends linearly on $\mathcal{R}_{\mathbf{k}}(\mathrm{rad})$. Because at $t > t_{\mathrm{rad}}$ the horizon expands more rapidly than the perturbation scales, the scales enter the horizon one by one (smaller scales enter earlier than the larger ones) and their $\mathcal{R}$ evolve thereafter.

Present theories, like inflation, for the origin of the perturbations assume that their primordial values have been produced by random processes. Therefore these theories do not give us the primordial values themselves but their statistical properties. It is usually assumed that $\mathcal{R}_{\mathbf{k}}(\mathrm{rad})$ are zero mean, complex Gaussian random variables with a power spectrum ([20])

$$\langle \mathcal{R}_{\mathbf{k}}(\mathrm{rad})\mathcal{R}^*_{\mathbf{k}'}(\mathrm{rad})\rangle = \frac{2\pi^2}{k^3}\mathcal{P}_{\mathcal{R}}(k)\delta(\mathbf{k}-\mathbf{k}'). \quad (2.23)$$

Here asterisk indicates a complex conjugate and $\langle\cdot\rangle$ is an ensemble average. Typical inflation theories predict that $\mathcal{P}_{\mathcal{R}}(k) \propto k^{n-1}$, where $n \approx 1$ (scale invariant spectrum). The parameter n is called the spectral index.

2.3 Boltzmann Equation for the CMB Photons

Using Eqs. (2.15) - (2.20) and the initial conditions, the evolution of the fluid perturbations can be determined. These equations govern the fluid as a whole but they are not able to tell us anything about the evolution of different particle species of the fluid. At $t > t_{\rm rad}$ we assume that the universe contains photons, baryons (electrons, protons, neutrons, ions, atoms), cold dark matter (CDM) and neutrinos. Although electrons are not really baryons (they are leptons), they are tightly coupled to baryons, forming a single fluid component with them, and therefore it is customary in cosmology to include electrons under the term "baryons". Photons and baryons interact with Compton scattering, charged baryons are coupled with an electromagnetic interaction, neutrons and protons with the strong interaction, and all particles contribute to the gravity and are affected by it.

To account for the particle interactions we need to discuss the distribution functions of each particle species and examine their Boltzmann equations that govern their evolution. Boltzmann equations of different particle species are discussed in e.g. [20] - [24]. The issues discussed in the remaining parts of this chapter are extracted from these references. This section covers the Boltzmann equation for the CMB photons. We are restricted to the CMB temperature anisotropies. The Boltzmann equations for polarization are discussed in e.g. [25].

2.3.1 Photon Distribution and Brightness Functions

The photon distribution function $f(t, \mathbf{x}, \mathbf{p})$ is defined in 6-dimensional phase space so that $\frac{g}{(2\pi)^3} f$ is the number of photons in a phase space element $d^3x d^3p$. Here $g = 2$ is the number of spin states, (t,$\mathbf{x}$) are the comoving coordinates and $\mathbf{p}$ is the photon momentum in the locally orthonormal coordinates of the comoving observer (i.e., one who is at rest in the (t,$\mathbf{x}$) coordinate system). The distribution function evolves in time due to freely falling motion of the photons and due to collisions with charged baryons. The evolution is governed by the Boltzmann equation

$$\frac{df}{dt} = C[f], \tag{2.24}$$

where $C[f]$ is the collision term. The collisionless Boltzmann equation ($df/dt = 0$) is discussed in Sect. 2.3.2 and the collision term is given in Sect. 2.3.3 of this thesis.

In first order perturbation theory the distribution function is a sum of the background function and a perturbation. The background distribution function can be approximated by the distribution function of the thermal equilibrium (blackbody distribution function)

$$\bar{f}(t, p) = \frac{1}{e^{p/T(t)} - 1}. \tag{2.25}$$

Here $p = |\mathbf{p}|$, which equals the photon energy (in the locally orthonormal coordinate system). Today the value of the temperature $T(t)$ is $T(t_0) = T_0 = 2.725$ K. The

perturbed distribution function can be given in the following form

$$f(t, \mathbf{x}, p, \hat{\mathbf{p}}) = \frac{1}{\exp\left(\frac{p}{T(t)[1+\Theta(t,\mathbf{x},\hat{\mathbf{p}})]}\right) - 1}. \tag{2.26}$$

Here $\mathbf{p} = p\hat{\mathbf{p}}$ and $\hat{\mathbf{p}}$ is the unit vector in the direction of the photon momentum. The function $\Theta(t, \mathbf{x}, \hat{\mathbf{p}})$ is called the *brightness* function. It does not depend on the full momentum, just on its direction. This reflects the usual approximation (justified in Sect. 2.3.3), where the perturbation depends on the direction of the photon momentum, but it is not a deviation from the blackbody spectrum ([24]).

The perturbed distribution function can be expanded (using Taylor series) in the sum of the background function and a perturbation

$$f(t, \mathbf{x}, p, \hat{\mathbf{p}}) = \bar{f}(t, p) + \delta f(t, \mathbf{x}, \hat{\mathbf{p}}) = \bar{f} + \frac{\partial \bar{f}}{\partial T} T\Theta = \bar{f} - p\frac{\partial \bar{f}}{\partial p}\Theta. \tag{2.27}$$

The last form is obtained after applying the identity $T\partial\bar{f}/\partial T = -p\partial\bar{f}/\partial p$.

2.3.2 Collisionless Boltzmann Equation

The derivative of the photon distribution function can be written as

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}\frac{dx^i}{dt} + \frac{\partial f}{\partial p}\frac{dp}{dt} + \frac{\text{"}\partial f}{\partial \hat{\mathbf{p}}}\frac{d\hat{\mathbf{p}}\text{"}}{dt} \tag{2.28}$$

The background distribution function does not depend on the photon direction and the photons do not change direction in the background universe. Therefore the terms $\partial f/\partial\hat{\mathbf{p}}$ and $d\hat{\mathbf{p}}/dt$ are perturbations and their product can be ignored (in first order perturbation theory).

The relation between the photon 4-momentum (P^0, P^i) in the coordinate basis and its momentum (p, p^i) in the locally orthonormal frame of the comoving observer is (no summation over repeated indeces)

$$p = \sqrt{|g_{00}|}P^0 \tag{2.29}$$

and

$$p^i = \sqrt{|g_{ii}|}P^i \tag{2.30}$$

Using $p^i = p\hat{p}^i$ ($\hat{p}^i$ is the i^{th} component of the unit vector $\hat{\mathbf{p}}$ of the photon momentum and $\delta_{ij}\hat{p}^i\hat{p}^j = 1$) and the conformal-Newtonian metric (Eq. (2.10)), we obtain

$$P^0 = a^{-1}(1 - \Phi)p \tag{2.31}$$

and

$$P^i = a^{-1}(1 + \Psi)p\hat{p}^i. \tag{2.32}$$

For the second term of Eq. (2.28) we need to evaluate dx^i/dt.

$$\frac{dx^i}{dt} = \frac{1}{a}\frac{dx^i}{d\eta} = \frac{1}{a}\frac{dx^i}{d\lambda}\frac{d\lambda}{d\eta} = \frac{1}{a}\frac{P^i}{P^0} = \frac{1}{a}\frac{1+\Psi}{1-\Phi}\hat{p}^i = a^{-1}(1+\Phi+\Psi)\hat{p}^i. \tag{2.33}$$

Here λ is the affine parameter of the photon trajectory ($P^\mu = dx^\mu/d\lambda$) and we used Eqs. (2.31) and (2.32) in place of P^0 and P^i. Because the background distribution function does not depend on x^i, the term $\partial f/\partial x^i$ is a perturbation and we can thus drop Φ and Ψ from Eq. (2.33) when using it in the second term of Eq. (2.28). We obtain for the second term

$$\frac{\partial f}{\partial x^i}\frac{dx^i}{dt} = \frac{\hat{p}^i}{a}\frac{\partial f}{\partial x^i}. \tag{2.34}$$

The term dp/dt required in the third term of Eq. (2.28) can be evaluated using the photon geodesic equation

$$\frac{dP^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta}P^\alpha P^\beta = 0. \tag{2.35}$$

Here $\Gamma^\mu_{\alpha\beta}$ are the Christoffel symbols of the metric (Eq. (2.10)). Dividing both sides of Eq. (2.35) with $P^0 = d\eta/d\lambda$ we obtain another form for the geodesic equation

$$\frac{dP^\mu}{d\eta} + \Gamma^\mu_{\alpha\beta}\frac{P^\alpha P^\beta}{P^0} = 0. \tag{2.36}$$

Using the time component ($\mu = 0$) of the geodesic equation and inserting Eqs. (2.31) and (2.32) in place of P^0 and P^i, the derivative $\frac{dp}{dt} = \frac{dp}{ad\eta}$ can be evaluated. We will not show the details of this derivation here but merely give the final result. The details of this calculation are given in e.g. [24].

$$\frac{dp}{dt} = p\left[-H - \frac{\hat{p}^i}{a}\frac{\partial\Phi}{\partial x^i} + \frac{\partial\Psi}{\partial t}\right]. \tag{2.37}$$

We can now write the collisionless Boltzmann equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\hat{p}^i}{a}\frac{\partial f}{\partial x^i} + p\frac{\partial f}{\partial p}\left[-H - \frac{\hat{p}^i}{a}\frac{\partial\Phi}{\partial x^i} + \frac{\partial\Psi}{\partial t}\right] = 0. \tag{2.38}$$

The first two terms on the right hand side are standard hydrodynamics. The first term in the brackets is the redshift due to expansion of the universe and the remaining terms represent the effects of the perturbations.

There is a collisionless Boltzmann equation in the background universe as well. We can extract it from Eq. (2.38)

$$\frac{d\bar{f}}{dt} = \frac{\partial\bar{f}}{\partial t} - Hp\frac{\partial\bar{f}}{\partial p} = 0. \tag{2.39}$$

It is easy to show, that $\bar{f}$ with $T(t) \propto 1/a$ is a solution of the collisionless background Boltzmann equation. This is the redshift of the temperature of the background universe.

Subtracting Eq. (2.39) from Eq. (2.38) we obtain a collisionless Boltzmann equation for the brightness function ([24])

$$\frac{d\delta f}{dt} = -p\frac{\partial \bar{f}}{\partial p}\left[\frac{\partial \Theta}{\partial t} + \frac{\hat{p}^i}{a}\frac{\partial \Theta}{\partial x^i} + \frac{\hat{p}^i}{a}\frac{\partial \Phi}{\partial x^i} - \frac{\partial \Psi}{\partial t}\right] = 0. \tag{2.40}$$

The term inside the brackets must be zero. It can be expressed in terms of the conformal time

$$\frac{\partial \Theta}{\partial \eta} + \hat{p}^i\frac{\partial \Theta}{\partial x^i} + \hat{p}^i\frac{\partial \Phi}{\partial x^i} - \frac{\partial \Psi}{\partial \eta} = 0. \tag{2.41}$$

This is the *collisionless brightness equation*.

We can Fourier expand both sides of this equation and obtain the collisionless brightness equation for the Fourier modes of the photon brightness function

$$\Theta' + ik\mu\Theta + ik\mu\Phi - \Psi' = 0. \tag{2.42}$$

Here $\mu \equiv \hat{k}^i\hat{p}^i = \hat{\mathbf{k}}\cdot\hat{\mathbf{p}}$ is the cosine of the angle between the Fourier mode wavevector $\mathbf{k}$ and the photon momentum $\mathbf{p}$ and $' \equiv d/d\eta$.

2.3.3 Collision Term

At the time of photon decoupling all baryons are moving with non-relativistic velocities. Scattering of photons from non-relativistic charged baryons is called Thomson scattering. Its differential cross section is (e.g. [24])

$$\frac{d\sigma}{d\Omega} = \frac{3}{16\pi}\sigma_{\mathrm{T}}\left(1 + \cos^2(\theta)\right), \tag{2.43}$$

where $\sigma_{\mathrm{T}} = \frac{8\pi}{3}\frac{\alpha^2}{m^2}$ is the total cross section of the Thomson scattering, α is the fine-structure constant and m is the rest mass of the baryon. The scattering angle between the directions of incoming and outgoing photon (in the baryon rest frame) is θ. Due to the $1/m^2$ dependence Thomson scattering from the electrons is the only relevant scattering here.

Schematically (ignoring the stimulated emission and Pauli blocking) the collision term $C[f]$ (see Eq. (2.24)) can be expressed as ([24])

$$C[f(\mathbf{p})] = \sum_{\mathbf{q},\mathbf{q}',\mathbf{p}'} |A|^2\left[f_e(\mathbf{q}')f(\mathbf{p}') - f_e(\mathbf{q})f(\mathbf{p})\right]. \tag{2.44}$$

Here subscript "e" refers to electron distribution function, $\mathbf{q}$ and $\mathbf{q}'$ are the electron momenta before and after the collision, $\mathbf{p}$ and $\mathbf{p}'$ are the corresponding photon momenta and $|A|^2$ is the scattering magnitude. It is directly proportional to the

differential scattering cross section $\frac{d\sigma}{d\Omega}$. We are interested in the change of distribution of photons with momentum $\mathbf{p}$. Therefore we sum over the other momenta.

The collision term $C[f]$ is zero for the electron and photon background distribution functions ([24]). Thus the collision term is a perturbation and the collisional Boltzmann equation $df/dt = C[f]$ splits in to $d\bar{f}/dt = 0$ and $d\delta f/dt = C[f]$.

The derivation of the collision term is a tedious task that we are not going to do here. We merely give the final result. The details of this calculation are given in e.g. [24]. The collision term for photons is

$$C[f(t,\mathbf{x},\mathbf{p})] = -p\frac{\partial \bar{f}}{\partial p} n_{\mathrm{e}} \sigma_{\mathrm{T}} \left[\Theta_0 - \Theta + \hat{\mathbf{p}} \cdot \mathbf{v}_{\mathrm{b}} + \frac{3}{4} \hat{p}^i \hat{p}^j \Theta_2^{ij} \right]. \tag{2.45}$$

Here n_{e} is the density of free electrons and vector $\mathbf{v}_{\mathrm{b}}$ is the velocity perturbation of the baryons (electrons are tightly coupled to the other baryons due to Coulomb interaction and they all have the same velocity perturbation). The velocity perturbation is a function of $(t, \mathbf{x})$. The term $\hat{\mathbf{p}} \cdot \mathbf{v}_{\mathrm{b}}$ arises from the induced dipole in the rest frame of the scattering electron which is why it is independent of the photon brightness function Θ. The quantities Θ_0 and Θ_2^{ij} are the monopole and quadrupole of the photon brightness function. They are defined as

$$\Theta_0(t,\mathbf{x}) = \frac{1}{4\pi} \int d\Omega_{\hat{\mathbf{p}}} \Theta(t,\mathbf{x},\hat{\mathbf{p}}) \tag{2.46}$$

and

$$\Theta_2^{ij}(t,\mathbf{x}) = \frac{1}{4\pi} \int d\Omega_{\hat{\mathbf{p}}} \left(\hat{p}^i \hat{p}^j - \frac{1}{3}\delta^{ij} \right) \Theta(t,\mathbf{x},\hat{\mathbf{p}}). \tag{2.47}$$

The quadrupole term arises from the angular dependence of Thomson scattering (see Eq. (2.43)). We ignore here the polarization dependence of Thomson scattering.

We finally obtain the collisional Boltzmann equation for the photons

$$p\frac{\partial \bar{f}}{\partial p} \left[\frac{\partial \Theta}{\partial t} + \frac{\hat{p}^i}{a}\frac{\partial \Theta}{\partial x^i} + \frac{\hat{p}^i}{a}\frac{\partial \Phi}{\partial x^i} - \frac{\partial \Psi}{\partial t} \right] = p\frac{\partial \bar{f}}{\partial p} n_{\mathrm{e}} \sigma_{\mathrm{T}} \left[\Theta_0 - \Theta + \hat{\mathbf{p}} \cdot \mathbf{v}_{\mathrm{b}} + \frac{3}{4} \hat{p}^i \hat{p}^j \Theta_2^{ij} \right]. \tag{2.48}$$

The only p dependent (photon energy dependent) term is $p\partial\bar{f}/\partial p$. Because it can be divided off from the both sides of Eq. (2.48), the brightness function remains independent from the photon energy. Therefore the perturbations do not change the frequency spectrum of the CMB photons and their spectrum today is still close to the blackbody spectrum.

We can equate the terms multiplying $p\partial\bar{f}/\partial p$ in both sides of Eq. (2.49) and obtain the collisional brightness equation for photons

$$\frac{\partial \Theta}{\partial t} + \frac{\hat{p}^i}{a}\frac{\partial \Theta}{\partial x^i} + \frac{\hat{p}^i}{a}\frac{\partial \Phi}{\partial x^i} - \frac{\partial \Psi}{\partial t} = n_{\mathrm{e}} \sigma_{\mathrm{T}} \left[\Theta_0 - \Theta + \hat{\mathbf{p}} \cdot \mathbf{v}_{\mathrm{b}} + \frac{3}{4} \hat{p}^i \hat{p}^j \Theta_2^{ij} \right]. \tag{2.49}$$

In terms of conformal time the collisional brightness equation is

$$\frac{\partial \Theta}{\partial \eta} + \hat{p}^i\frac{\partial \Theta}{\partial x^i} + \hat{p}^i\frac{\partial \Phi}{\partial x^i} - \frac{\partial \Psi}{\partial \eta} = a n_{\mathrm{e}} \sigma_{\mathrm{T}} \left[\Theta_0 - \Theta + \hat{\mathbf{p}} \cdot \mathbf{v}_{\mathrm{b}} + \frac{3}{4} \hat{p}^i \hat{p}^j \Theta_2^{ij} \right]. \tag{2.50}$$

We will obtain our final form of the collisional brightness equation after we have applied the following three steps in Eq. (2.50).

1. We have restricted ourselves to scalar perturbations in this chapter. Therefore we assume that the perturbations $\mathbf{v}_\mathrm{b}$ and Θ_2^{ij} are expressed in terms of scalar perturbations v_b and Θ_2: $\mathbf{v}_\mathrm{b} = -\nabla v_\mathrm{b}$ and $\Theta_2^{ij} = (\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2)\Theta_2$. Here ∂_i denotes partial derivation with respect to x^i.

2. We Fourier expand both sides of Eq. (2.50) using Eq. (2.12) (for Θ, Ψ, Φ), Eq. (2.13) (for v_b) and Eq. (2.14) (for Θ_2). Thereafter we can extract the equation for the Fourier modes of the brightness function.

3. We define the *optical depth* $\tau(\eta)$ as an integral of $an_\mathrm{e}\sigma_\mathrm{T}$ from conformal time η to the present time (η_0)

$$\tau(\eta) = \int_\eta^{\eta_0} d\eta a n_\mathrm{e}\sigma_\mathrm{T}. \tag{2.51}$$

   The significance of the optical depth is that the factor $e^{-\tau}$ represents the fraction of photons that have not been scattered between η and η_0. Because $an_\mathrm{e}\sigma_\mathrm{T} = -\tau'$, we can use $-\tau'$ in place of $an_\mathrm{e}\sigma_\mathrm{T}$. For the later use we define here $\tau_\mathrm{r} \equiv \tau(\eta_\mathrm{reion})$, where η_reion is the time when the neutral gas (neutral since photon decoupling) between early stars and galaxies has become ionized because of the radiation from those stars and galaxies. This is called the *reionization* and according to WMAP observations the gas was reionized around $z_\mathrm{reion} \approx 20$ ([7]), which corresponds to $\eta = \eta_\mathrm{reion}$. The parameter τ_r is the optical depth due to reionization.

Now we are ready to write the collisional brightness equation for the Fourier modes of the photon brightness function (cf. Eq. (2.42) for the collisionless equation)

$$\Theta' + ik\mu\Theta + ik\mu\Phi - \Psi' = -\tau'\left[\Theta_0 - \Theta - i\mu v_\mathrm{b} - \frac{1}{2}P_2(\mu)\Theta_2\right]. \tag{2.52}$$

The function $P_2(\mu) = \frac{1}{2}(3\mu^2 - 1)$ is the 2^nd order Legendre polynomial and the parameter $\mu = \hat{\mathbf{k}}\cdot\hat{\mathbf{p}}$ was defined in Eq. (2.42). We remind the reader that here Θ is a function of $(\eta, \mathbf{k}, \hat{\mathbf{p}})$ and Ψ, Φ, v_b, Θ_0 and Θ_2 are functions of $(\eta, \mathbf{k})$.

2.3.4 Boltzmann Hierarchy

Well before the photon decoupling photons and baryons were nearly in a thermal equilibrium. The collision term of the photon brightness equation (right hand side of Eq. (2.52) was small at that time. Because the number density of the free electrons (n_e) was large (leading to large $-\tau'$), the term inside the brackets was nearly zero leading to

$$\Theta(\eta, \mathbf{k}, \hat{\mathbf{p}}) \approx \Theta_0(\eta, \mathbf{k}) - i\mu v_\mathrm{b}(\eta, \mathbf{k}) - \frac{1}{2}P_2(\mu)\Theta_2(\eta, \mathbf{k}) \tag{2.53}$$

We can see that before photon decoupling the brightness function depends on the direction of the photon momentum ($\hat{\mathbf{p}}$) only via the parameter $\mu = \hat{\mathbf{p}} \cdot \hat{\mathbf{k}} = \cos(\theta)$, where θ is the angle between $\hat{\mathbf{p}}$ and $\hat{\mathbf{k}}$.

Assuming that $\hat{\mathbf{k}}$ is aligned with the z-axis of a cartesian coordinate system, θ is the elevation angle of the photon momentum in the corresponding spherical coordinate system (θ, φ). Because the collisional brightness equation involves only η, $\mathbf{k}$ and μ, the brightness function Θ remains a function of $(\eta, \mathbf{k}, \mu)$ and no φ (azimuth) dependence will be developed. This allows us to expand $\Theta(\eta, \mathbf{k}, \mu)$ in terms of Legendre polynomials $P_\ell(\mu)$ (see e.g. [24])

$$\Theta(\eta, \mathbf{k}, \mu) = \sum_{\ell=0}^{\infty} (-i)^\ell (2\ell + 1) \Theta_\ell(\eta, \mathbf{k}) P_\ell(\mu). \tag{2.54}$$

The first three Legendre polynomials are $P_0(\mu) = 1$, $P_1(\mu) = \mu$ and $P_2(\mu) = \frac{1}{2}(3\mu^2 - 1)$. Using the orthogonality property

$$\int_{-1}^{1} d\mu P_\ell(\mu) P_{\ell'}(\mu) = \frac{2}{2\ell + 1} \delta_{\ell\ell'} \tag{2.55}$$

the above equation can be inverted to give the multipoles $\Theta_\ell(\eta, \mathbf{k})$ of the photon brightness function

$$\Theta_\ell(\eta, \mathbf{k}) = i^\ell \int_{-1}^{1} \frac{d\mu}{2} P_\ell(\mu) \Theta(\eta, \mathbf{k}, \mu). \tag{2.56}$$

A straightforward calculation shows that the functions Θ_0 and Θ_2 of Eq. (2.52) are indeed the monopole and quadrupole of Eq. (2.56).

We see from Eq. (2.53) that before the photon decoupling the brightness function contained essentially only monopole and dipole components (quadrupole is zero since it appears on both sides of Eq. (2.53) with different coefficients). The quadrupole and the higher multipoles develop after photon decoupling when the photons are travelling to us.

To obtain an equation for the multipoles we operate with $i^\ell \int_{-1}^{1} d\mu P_\ell(\mu)$ on both sides of Eq. (2.52). Using the orthogonality of the Legendre polynomials (Eq. (2.55)) and the relation $(2\ell + 1)\mu P_\ell(\mu) = (\ell + 1) P_{\ell+1}(\mu) + \ell P_{\ell-1}(\mu)$ we obtain for the multipoles of the photon brightness function

$$\begin{aligned} \Theta'_\ell + \frac{k}{2\ell + 1} \left[(\ell + 1) \Theta_{\ell+1} - \ell \Theta_{\ell-1} \right] - \frac{k}{3} \Phi \delta_{\ell 1} - \Psi' \delta_{\ell 0} = \\ = -\tau' \left[\Theta_0 \delta_{\ell 0} - \Theta_\ell + \frac{1}{3} v_{\mathrm{b}} \delta_{\ell 1} + \frac{1}{10} \Theta_2 \delta_{\ell 2} \right]. \end{aligned} \tag{2.57}$$

Using the photon distribution function the components of the photon energy-momentum tensor in the locally orthonormal frame can be calculated as

$$T_\gamma^{\mu\nu} = \int d^3p f \frac{p^\mu p^\nu}{p}. \tag{2.58}$$

Here $p = |\mathbf{p}|$ is the energy of the photon in the locally orthonormal coordinates of the comoving observer and p^i are the space components of the photon momentum vector $\mathbf{p}$.

We can equate this energy-momentum tensor to the energy-momentum tensor of Eq. (2.11) containing the photon perturbations. This gives the following relations between the photon perturbations and the multipoles of the brightness function

$$\Theta_0(\eta, \mathbf{k}) = \frac{1}{4}\delta_\gamma(\eta, \mathbf{k}), \quad \Theta_1(\eta, \mathbf{k}) = \tfrac{1}{3}v_\gamma(\eta, \mathbf{k}), \quad \Theta_2(\eta, \mathbf{k}) = \frac{1}{12}\Pi_\gamma(\eta, \mathbf{k}). \quad (2.59)$$

Using these relations we can write for the multipoles

$$\delta'_\gamma + \frac{4}{3}kv_\gamma - 4\Psi' = 0 \quad (\ell = 0) \quad (2.60)$$

$$v'_\gamma + \frac{k}{6}\Pi_\gamma - \frac{k}{4}\delta_\gamma - k\Phi = \tau'(v_\gamma - v_\mathrm{b}) \quad (\ell = 1) \quad (2.61)$$

$$\Theta'_2 + \frac{k}{5}\left[3\Theta_3 - 2\Theta_1\right] = \frac{9}{10}\tau'\Theta_2 \quad (\ell = 2) \quad (2.62)$$

$$\Theta'_\ell + \frac{k}{2\ell+1}\left[(\ell+1)\Theta_{\ell+1} - \ell\Theta_{\ell-1}\right] = \tau'\Theta_\ell \quad (\ell \geq 3). \quad (2.63)$$

This system of coupled equations for the multipoles of the photon brightness function is called the *Boltzmann hierarchy*.

2.4 Boltzmann Equations for the Other Particle Species

We will not derive the Boltzmann equations for neutrinos, CDM and baryons in this chapter. We just give them here. The details of their derivations can be found in e.g. [23], [24].

Neutrinos were decoupled from the baryons well before the photon decoupling (neutrino decoupling at $T \approx 1$ MeV). After their decoupling the massless neutrinos behave very much like photons, except that they are fermions and they interact very weakly with other particles. Their perturbations can be described with neutrino brightness function $\Theta^\nu(\eta, \mathbf{x}, \hat{\mathbf{p}})$ that has a similar Boltzmann hierarchy as photons. The neutrino collision term is, however, zero.

The effect of neutrino mass is hardly detected in the present CMB data. Therefore we ignore the neutrino mass in this chapter. The Boltzmann equations for the massive neutrinos are discussed in e.g. [23].

It is assumed that the CDM particles are weakly interacting. Therefore they are described by a collisionless Boltzmann equation. It can be derived in a similar fashion as the collisionless Boltzmann equation for the photons, except that the large non-zero rest mass of the CDM particles needs to be accounted for. Taking the 0^{th} and

1^{st} moments of the Boltzmann equation (operating with $\int d^3p$ and $\int d^3p\hat{p}^i$ on both sides of the Boltzmann equation) we obtain the equations for the density contrast and the velocity perturbation of the CDM. Because p/E of the CDM particles is small (due to large rest mass), the higher order multipoles ($\ell \geq 2$) are insignificant.

The collisionless part of the baryon Boltzmann equation is identical to the Boltzmann equation of CDM. The baryon interactions with other baryons and photons contribute to the collision term.

The Boltzmann equations for the massless neutrinos (subscript "ν"), CDM (subscript "c") and baryons (subscript "b") are

$$\delta'_\nu + \frac{4}{3}kv_\nu - 4\Psi' = 0 \tag{2.64}$$

$$v'_\nu + \frac{k}{6}\Pi_\nu - \frac{k}{4}\delta_\nu - k\Phi = 0 \tag{2.65}$$

$$\delta'_{\rm c} + kv_{\rm c} - 3\Psi' = 0 \tag{2.66}$$

$$v'_{\rm c} + v_{\rm c} - k\Phi = 0 \tag{2.67}$$

$$\delta'_{\rm b} + kv_{\rm b} - 3\Psi' = 0 \tag{2.68}$$

$$v'_{\rm b} + v_{\rm b} - k\Phi = -\tau'\frac{4\rho_\gamma}{3\rho_{\rm b}}(v_\gamma - v_{\rm b}). \tag{2.69}$$

There is a similar hierarchy for neutrinos as for photons, except for the collision term.

2.5 C_ℓ Spectrum

The Boltzmann hierarchy of CMB photons (Eqs. (2.60) - (2.63)), Boltzmann equations for neutrinos, CDM and baryons (Eqs. (2.64) - (2.69)) and Einstein equations for Ψ and Φ (Eqs. (2.15) - (2.18)) constitute a complete set of equations, where the multipoles of the photon brightness function can, in principle, be solved for any time η after the photon decoupling. Due to the large number of intercoupled equations this is, in general, a tedious task. In Sect. 2.6 we will discuss the line-of-sight integration ([26]), which is an approach that makes this problem more tractable and enables more efficient numerical computations.

For the time being let us assume that we have been able to find the values that the multipoles $\Theta_\ell(\eta, \mathbf{k})$ have today (at $\eta = \eta_0$). The CMB temperature anisotropy $\delta T(\eta_0, \mathbf{x} = 0, \hat{\mathbf{p}})$, that we detect today, depends on the photon brightness function: $\delta T(\eta_0, \mathbf{x} = 0, \hat{\mathbf{p}}) = T_0\Theta(\eta_0, \mathbf{x} = 0, \hat{\mathbf{p}})$. Here T_0 is the background CMB temperature ($T_0 = 2.725$ K) and $\mathbf{x} = 0$ is the location of the observer.

The CMB temperature anisotropy can be expressed in terms of the Fourier mode expansion of the photon brightness function (cf. Eq. (2.12))

$$\delta T(\eta_0, \hat{\mathbf{p}}) = \delta T(\eta_0, \mathbf{x} = 0, \hat{\mathbf{p}}) = T_0\Theta(\eta_0, \mathbf{x} = 0, \hat{\mathbf{p}}) = \frac{T_0}{(2\pi)^{3/2}}\int d^3k\Theta(\eta_0, \mathbf{k}, \mu). \tag{2.70}$$

We can insert the multipole expansion (Eq. (2.54)) in place of $\Theta(\eta_0, \mathbf{k}, \mu)$. Before doing that we use the identity

$$(2\ell+1)P_\ell(\mu) = (2\ell+1)P_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{p}}) = 4\pi\sum_{m=-\ell}^{\ell} Y^*_{\ell m}(\hat{\mathbf{k}})Y_{\ell m}(\hat{\mathbf{p}}) \tag{2.71}$$

and substitute $(2\ell+1)P_\ell(\mu)$ with $4\pi\sum_{m=-\ell}^{\ell} Y^*_{\ell m}(\hat{\mathbf{k}})Y_{\ell m}(\hat{\mathbf{p}})$ in the expansion. We obtain now for the anisotropy

$$\delta T(\eta_0, \hat{\mathbf{p}}) = \sum_{\ell m}\left[\frac{4\pi T_0}{(2\pi)^{3/2}}(-i)^\ell\int d^3k\Theta_\ell(\eta_0,\mathbf{k})Y^*_{\ell m}(\hat{\mathbf{k}})\right]Y_{\ell m}(\hat{\mathbf{p}}). \tag{2.72}$$

Here the ℓ-sum goes from 0 to ∞ and the m-sum from $-\ell$ to ℓ (at each ℓ). The CMB anisotropy $\delta T(\eta_0, \hat{\mathbf{p}})$ is a function on the celestial sphere and we have above its expansion in terms of the spherical harmonics $Y_{\ell m}(\hat{\mathbf{p}})$. We can extract the $a_{\ell m}$ expansion coefficients from Eq. (2.72)

$$a_{\ell m} = \frac{4\pi T_0}{(2\pi)^{3/2}}(-i)^\ell\int d^3k\Theta_\ell(\eta_0,\mathbf{k})Y^*_{\ell m}(\hat{\mathbf{k}}). \tag{2.73}$$

Because we assume first order perturbation theory, the coefficients $\Theta_\ell(\eta_0, \mathbf{k})$ depend linearly on the primordial curvature perturbations (see Sect. 2.2)

$$\Theta_\ell(\eta_0,\mathbf{k}) \equiv T_\ell(\eta_0, k)\mathcal{R}_{\mathbf{k}}(\mathrm{rad}). \tag{2.74}$$

This equation defines the *transfer function* $T_\ell(\eta_0, k)$. It can be calculated using the equations we have discussed above. Note that it depends on the magnitude of the wavevector and the direction dependence is in the primordial perturbations only. Because we assumed that the primordial curvature perturbations are complex zero mean Gaussian distributed random variables, the $a_{\ell m}$ are complex, zero mean and Gaussian distributed as well.

Under the assumption of a Gaussian distribution the covariance $\langle a_{\ell m}a^*_{\ell' m'}\rangle$ contains the full statistical description of the CMB temperature anisotropy. The covariance is

$$\langle a_{\ell m}a^*_{\ell' m'}\rangle = \frac{2T_0^2}{\pi}(-i)^\ell(i)^{\ell'}\int d^3k d^3k' Y^*_{\ell m}(\hat{\mathbf{k}})Y_{\ell' m'}(\hat{\mathbf{k}'})\langle\Theta_\ell(\eta_0,\mathbf{k})\Theta^*_{\ell'}(\eta_0,\mathbf{k}')\rangle, \tag{2.75}$$

where the expectation value $\langle\Theta_\ell\Theta^*_{\ell'}\rangle$ can be expressed in terms of the primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ (see Eq. (2.23))

$$\begin{aligned}\langle\Theta_\ell(\eta_0,\mathbf{k})\Theta^*_{\ell'}(\eta_0,\mathbf{k}')\rangle &= T_\ell(\eta_0,k)T^*_{\ell'}(\eta_0,k')\langle\mathcal{R}_{\mathbf{k}}(\mathrm{rad})\mathcal{R}^*_{\mathbf{k}'}(\mathrm{rad})\rangle = \\ &= \frac{2\pi^2}{k^3}T_\ell(\eta_0,k)T^*_{\ell'}(\eta_0,k')\mathcal{P}_\mathcal{R}(k)\delta(\mathbf{k}-\mathbf{k}').\end{aligned} \tag{2.76}$$

We can insert the latter form back to Eq. (2.75), perform d^3k' integration (trivial due to delta function $\delta(\mathbf{k}-\mathbf{k}')$) and split the remaining integral in $d^3k = k^2 dk d\Omega_{\hat{\mathbf{k}}}$

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = 4\pi T_0^2 (-i)^\ell (i)^{\ell'} \int_0^\infty \frac{dk}{k} T_\ell(\eta_0, k) T^*_{\ell'}(\eta_0, k) \mathcal{P}_\mathcal{R}(k) \int d\Omega_{\hat{\mathbf{k}}} Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell' m'}(\hat{\mathbf{k}}). \tag{2.77}$$

Due to the completeness of the spherical harmonics the $d\Omega_{\hat{\mathbf{k}}}$ integral gives a product of delta functions $\delta_{\ell\ell'}\delta_{mm'}$. We finally obtain

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = 4\pi T_0^2 \delta_{\ell\ell'} \delta_{mm'} \int_0^\infty \frac{dk}{k} |T_\ell(\eta_0, k)|^2 \mathcal{P}_\mathcal{R}(k). \tag{2.78}$$

The quantity multiplying the delta functions is the expectation value C_ℓ of the *angular power spectrum* of the CMB temperature anisotropy

$$C_\ell = 4\pi T_0^2 \int_0^\infty \frac{dk}{k} |T_\ell(\eta_0, k)|^2 \mathcal{P}_\mathcal{R}(k). \tag{2.79}$$

For a given ℓ, each $a_{\ell m}$ has the same variance. Because $\delta T(\eta_0, \hat{\mathbf{p}})$ is real, $a_{\ell,-m} = (-1)^m a^*_{\ell m}$. This means that $a_{\ell 0}$ is real and there are $2\ell + 1$ degrees of freedom for a given ℓ.

The fact, that the covariance $\langle a_{\ell m} a^*_{\ell' m'} \rangle$ is diagonal, is a reflection of the statistical isotropy of the CMB field. From a point in the celestial sphere the CMB anisotropy field looks the same in a statistical sense in every direction in the sphere. There are no preferred directions. The correlation $\langle \delta T(\eta_0, \hat{\mathbf{p}}) \delta T(\eta_0, \hat{\mathbf{q}}) \rangle$ between two points of the CMB anisotropy field depends only on the angular spectrum C_ℓ and the angle between $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$.

The CMB angular power spectrum can be both predicted by cosmological theories and observed in the CMB experiments. Therefore it is in a central role when we develop cosmological models that describe the history of the early universe.

As a rule-of-thumb, C_ℓ describes CMB temperature anisotropies whose minima and maxima have a typical angular separation $\sim \pi/\ell$ in the sky. Therefore a multipole ℓ reflects the angular scale of the anisotropies. In a PLANCK-like experiment, where the angular resolution of the observations is ∼5 arcmin ([8], [9]), the interesting multipoles of the CMB angular power spectrum extend up to $\ell = 2000 \ldots 3000$.

In Chapter 5 of this thesis we call C_ℓ the underlying "theoretical" angular power spectrum of the CMB sky. We use a symbol C_ℓ^{th} for it there.

2.6 Line-of-Sight Integration

Line-of-sight (LOS) integration is an efficient technique to solve the present time values of the multipoles $\Theta_\ell(\eta, \mathbf{k})$ ([26]). We will not show the complete derivation of the LOS integral here, but we describe the main steps to reach it. The detailed derivation is given in e.g. [24].

We start from the photon brightness equation (Eq. (2.52)) and put it in the following form

$$\Theta' + (ik\mu\ - \tau')\Theta = -ik\mu\Phi + \Psi' - \tau'\left[\Theta_0 - i\mu v_{\rm b} - \frac{1}{2}P_2(\mu)\Theta_2\right] \equiv \widetilde{S}(\eta, \mathbf{k}, \mu). \quad (2.80)$$

This can be further modified to

$$\frac{d}{d\eta}\left(\Theta e^{ik\mu\eta - \tau(\eta)}\right) = \widetilde{S}(\eta, \mathbf{k}, \mu)e^{ik\mu\eta - \tau(\eta)}, \quad (2.81)$$

where $\widetilde{S}(\eta, \mathbf{k}, \mu)$ equals the right hand side of Eq. (2.80). We can now integrate both sides of Eq. (2.81) from an initial time ($\eta = \eta_{\rm init}$) to the present time ($\eta = \eta_0$). The initial time is some time well before the photon decoupling. The integration of the left hand side is simple. The term at η_0 is $\Theta(\eta_0, \mathbf{k}, \mu)e^{ik\mu\eta_0}$, because $\tau(\eta_0) = 0$. The term at $\eta_{\rm init}$ is zero because $\tau(\eta_{\rm init}) \gg 1$, which leads to $e^{-\tau(\eta_{\rm init})} \ll 1$. Due to large τ at early times ($\eta < \eta_{\rm dec}$) we can start the integration from $\eta = 0$ in the right hand side. We have now an expression for $\Theta(\eta_0, \mathbf{k}, \mu)$

$$\Theta(\eta_0, \mathbf{k}, \mu) = \int_0^{\eta_0} d\eta e^{-\tau(\eta)}\widetilde{S}(\eta, \mathbf{k}, \mu)e^{ik\mu(\eta - \eta_0)}. \quad (2.82)$$

We can see from Eq. (2.80) that $\widetilde{S}(\eta, \mathbf{k}, \mu)$ is a second order polynomial in μ. Terms $\mu e^{ik\mu(\eta-\eta_0)}$ and $\mu^2 e^{ik\mu(\eta-\eta_0)}$ can be replaced with derivatives $\frac{1}{ik}\frac{d}{d\eta}\left(e^{ik\mu(\eta-\eta_0)}\right)$ and $-\frac{1}{k^2}\frac{d^2}{d\eta^2}\left(e^{ik\mu(\eta-\eta_0)}\right)$ and after integration by parts we obtain

$$\Theta(\eta_0, \mathbf{k}, \mu) = \int_0^{\eta_0} d\eta S(\eta, \mathbf{k})e^{ik\mu(\eta-\eta_0)}. \quad (2.83)$$

Here we have dropped some monopole and dipole terms at the observer that arise from the integration by parts ([24]). The function $S(\eta, \mathbf{k})$ is

$$S(\eta, \mathbf{k}) = g(\eta)\left(\Theta_0 + \Phi + \frac{v_{\rm b}'}{k} + \frac{1}{4}\Theta_2 + \frac{3}{4k^2}\Theta_2''\right) +$$
$$+g'(\eta)\left(\frac{v_{\rm b}}{k} + \frac{3}{2k^2}\Theta_2'\right) + g''(\eta)\frac{3}{4k^2}\Theta_2 + e^{-\tau}\left(\Psi' + \Phi'\right). \quad (2.84)$$

The quantity $g(\eta) \equiv -\tau' e^{-\tau}$ is called the visibility function. For a given CMB photon, $g(\eta)d\eta$ gives the probability that it last scattered during $(\eta, \eta + d\eta)$. The visibility function peaks strongly at the time of photon decoupling ($\eta = \eta_{\rm dec}$) and drops rapidly to zero at other times (see e.g. [24]). Its time integral has the value $\int_0^{\eta_0} d\eta g(\eta) = 1$.

Eq. (2.83) is a convenient result because μ appears only in the exponential of the right hand side and we can now determine the multipole coefficients $\Theta_\ell(\eta_0, \mathbf{k})$,

that was our original task. After applying $i^\ell \int_{-1}^{1} \frac{d\mu}{2} P_\ell(\mu)$ to both sides of Eq. (2.83) we obtain

$$\Theta_\ell(\eta_0, \mathbf{k}) = \int_0^{\eta_0} d\eta S(\eta, \mathbf{k}) j_\ell(k\eta_0 - k\eta). \tag{2.85}$$

Note that Eq. (2.85) applies to multipoles $\ell \geq 2$ only, because some monopole and dipole terms were dropped when deriving Eq. (2.83). In Eq. (2.85) $j_\ell(x)$ is a spherical Bessel function and it arises from the Legendre transform of the exponential term

$$i^\ell \int_{-1}^{1} \frac{d\mu}{2} P_\ell(\mu) e^{ik\mu(\eta - \eta_0)} = j_\ell(k\eta_0 - k\eta). \tag{2.86}$$

Eq. (2.85) is the LOS integral for $\Theta_\ell(\eta_0, \mathbf{k})$. The function $S(\eta, \mathbf{k})$ used in the integral is shown in Eq. (2.84). The first three terms of $S(\eta, \mathbf{k})$ are significant only around $\eta = \eta_{\rm dec}$, because the visibility function has a strong peak at photon decoupling. The last term of $S(\eta, \mathbf{k})$ ($e^{-\tau}(\Psi' + \Phi')$) is the integrated Sachs-Wolfe (ISW) effect, which is zero before the photon decoupling (because $e^{-\tau} \ll 1$ then) and non-zero whenever Ψ or Φ evolve in time after the photon decoupling.

Because at early times ($\eta < \eta_{\rm dec}$) only the lowest multipoles Θ_ℓ are significant, the Boltzmann hierarchy can be truncated at fairly low ℓ to obtain the quantities required in $S(\eta, \mathbf{k})$. The advantage of the LOS integral is that we need only a small number of the lowest multipoles at early times to obtain the multipoles $\Theta_\ell(\eta_0, \mathbf{k})$ at the present time.

The LOS integration method is applied in two widely used software codes CMBFAST ([26], [28]) and CAMB (Code for Anisotropies in the Microwave Background, [29]) that calculate the CMB angular power spectrum from theory.

2.7 C_ℓ Spectrum Today

We use some approximations to further simplify the LOS integral that we derived in the previous section (see Eqs. (2.84) and (2.85)). During photon decoupling $v_{\rm b} \approx 3\Theta_1$ (see Eqs. (2.59) and (2.61)) and the higher multipoles of the Boltzmann hierarchy are small due to photon-electron interactions. Therefore we drop the terms with Θ_2 in $S(\eta, \mathbf{k})$. After these approximations $S(\eta, \mathbf{k})$ can be written as (we have combined Θ_1 terms into a single derivative)

$$S(\eta, \mathbf{k}) \approx g\,(\Theta_0 + \Phi) + \frac{3}{k}\frac{d}{d\eta}\,(g\Theta_1) + e^{-\tau}\,(\Psi' + \Phi')\,. \tag{2.87}$$

Because the visibility function $g(\eta)$ peaks strongly at $\eta = \eta_{\rm dec}$, we approximate it with a delta function $g(\eta) \approx \delta(\eta - \eta_{\rm dec})$. Inserting the $S(\eta, \mathbf{k})$ approximation to the LOS integral Eq. (2.85), using the delta function approximation in place of the visibility function and carrying out an integration by parts for the $\frac{d}{d\eta}$ term, we

obtain an approximation for the multipoles

$$\Theta_\ell(\eta_0, \mathbf{k}) \approx [\Theta_0(\eta_{\rm dec}, \mathbf{k}) + \Phi(\eta_{\rm dec}, \mathbf{k})]\, j_\ell(k\eta_0 - k\eta_{\rm dec}) -$$

$$+3\Theta_1(\eta_{\rm dec}, \mathbf{k}) \left(j_{\ell-1}(k\eta_0 - k\eta_{\rm dec}) - \frac{\ell+1}{k(\eta_0 - \eta_{\rm dec})} j_\ell(k\eta_0 - k\eta_{\rm dec}) \right) +$$

$$+ \int_0^{\eta_0} d\eta e^{-\tau} \left(\Psi' + \Phi'\right) j_\ell(k\eta_0 - k\eta). \qquad (2.88)$$

The spherical Bessel term multiplying Θ_1 is equal to $\frac{d}{d\eta}[j_\ell(k\eta_0 - k\eta)]$ at $\eta = \eta_{\rm dec}$.

The above approximation gives the CMB anisotropy as a sum of three terms. The first term is the monopole shifted by the metric perturbation Φ at the photon decoupling. We call this the monopole term. Φ represents here the redshift that the CMB photon experiences when climbing up from a gravitational well. The second term is the dipole at photon decoupling (dipole term) and the third term is the ISW effect.

Because the spherical Bessel function $j_\ell(x)$ has its largest magnitude around $x \approx \ell$, the monopole at scale k contributes mainly to those Θ_ℓ, where $\ell \approx k(\eta_0 - \eta_{\rm dec}) \approx k\eta_0$. It is of interest to find out the angular scale $\ell = \ell_{\rm dec}$ that corresponds to the scale $k_{\rm dec}$ that entered the horizon at the photon decoupling. We assume a flat background universe, which today contains mainly matter and a cosmological constant: $\Omega_{\rm m} + \Omega_\Lambda = 1$. Here $\Omega_{\rm m}$ and Ω_Λ are the density parameters for the matter and cosmological constant. They are their present time energy densities as a fraction of the critical energy density $\rho_{\rm cr} = 3H_0^2/(8\pi G)$, where H_0 is the present value of the Hubble constant. We can integrate η_0 from the Friedmann equation (Eq. (2.7)). The approximate result is

$$\eta_0 \approx \frac{2H_0^{-1}}{\sqrt{\Omega_{\rm m}}}. \qquad (2.89)$$

The value of $k_{\rm dec}$ is equal to the comoving Hubble parameter $\mathcal{H}$ evaluated at $a = a_{\rm dec} = 1/(1 + z_{\rm dec})$. Because the universe is not yet fully matter dominated at $\eta = \eta_{\rm dec}$, we need to include both radiation and matter (contribution from the cosmological constant is insignificant at $\eta = \eta_{\rm dec}$). Using the Friedmann equation again we obtain

$$k_{\rm dec} = \mathcal{H}(a_{\rm dec}) = H_0 \sqrt{\Omega_{\rm m}} \left[\frac{1}{a_{\rm dec}} + \frac{\Omega_{\rm r}/\Omega_{\rm m}}{a_{\rm dec}^2} \right]. \qquad (2.90)$$

Here $\Omega_{\rm r}$ is the present value of the energy density of radiation (normalized with the critical density). The energy density ratio is $\Omega_{\rm r}/\Omega_{\rm m} = a_{\rm eq}$, where $a_{\rm eq}$ is the scale factor at matter-radiation equality. We use experimental values $z_{\rm dec} \approx 1089$ and $z_{\rm eq} \approx 3233$ (WMAP first year results, [7]) and obtain $\ell_{\rm dec} \approx k_{\rm dec}\eta_0 \approx 76$. Roughly speaking we can say that the CMB anisotropy with $\ell \lesssim \ell_{\rm dec}$ emanates from perturbations that are outside the horizon during the photon decoupling and the rest of the multipoles arise from perturbations that have entered the horizon before the photon decoupling.

2.7.1 Large Scales - Sachs-Wolfe Effect

We consider here scales that are outside horizon during the photon decoupling. At these scales $\Theta_\ell(\eta_0, \mathbf{k})$ is dominated by the monopole term $\Theta_0 + \Phi$ at photon decoupling ([24]). This allows us to write an approximation for the multipoles (cf. Eq. (2.88))

$$\Theta_\ell^{\mathrm{SW}}(\eta_0, \mathbf{k}) = \left[\Theta_0(\eta_{\mathrm{dec}}, \mathbf{k}) + \Phi(\eta_{\mathrm{dec}}, \mathbf{k})\right] j_\ell(k\eta_0 - k\eta_{\mathrm{dec}}). \tag{2.91}$$

This large scale anisotropy is called the Sachs-Wolfe effect ([15]).

For the scales outside horizon the curvature perturbation is constant and it has its primordial value during the photon decoupling $\mathcal{R}(\eta_{\mathrm{dec}}, \mathbf{k}) = \mathcal{R}_{\mathbf{k}}(\mathrm{rad})$. Assuming that the anisotropic stress is insignificant (leading to $\Psi \approx \Phi$) and making the approximation that the universe is matter dominated during the photon decoupling, Einstein equations (Eqs. (2.15) - (2.18)) give a constant (no time dependence) $\Phi(\eta, \mathbf{k})$ for scales outside horizon. This constant is $\Phi(\eta, \mathbf{k}) = \Phi(\eta_{\mathrm{dec}}, \mathbf{k}) = -\frac{3}{5}\mathcal{R}_{\mathbf{k}}(\mathrm{rad})$ (from Eq. (2.21)). The matter density contrast $\delta_{\mathrm{m}}(\eta_{\mathrm{dec}}, \mathbf{k})$ can be obtained from Eq. (2.15). It is $\delta_{\mathrm{m}}(\eta_{\mathrm{dec}}, \mathbf{k}) = -2\Phi(\eta_{\mathrm{dec}}, \mathbf{k}) = \frac{6}{5}\mathcal{R}_{\mathbf{k}}(\mathrm{rad})$. Because we assume adiabatic perturbations, the last term on the right hand side of Eq. (2.22) is zero leading to a relation between the matter and photon density contrasts: $\frac{3}{4}\delta_\gamma = \delta_{\mathrm{m}}$. Because $\Theta_0 = \frac{1}{4}\delta_\gamma$, we finally obtain $\Theta_0(\eta_{\mathrm{dec}}, \mathbf{k}) = \frac{1}{3}\delta_{\mathrm{m}}(\eta_{\mathrm{dec}}, \mathbf{k}) = \frac{2}{5}\mathcal{R}_{\mathbf{k}}(\eta_{\mathrm{dec}}, \mathbf{k})$. The multipole can now be written as

$$\Theta_\ell^{\mathrm{SW}}(\eta_0, \mathbf{k}) = -\frac{1}{5} j_\ell(k\eta_0 - k\eta_{\mathrm{dec}}) \mathcal{R}_{\mathbf{k}}(\mathrm{rad}). \tag{2.92}$$

A comparison to Eq. (2.74) shows that the transfer function is here $T_\ell^{\mathrm{SW}}(\eta_0, k) = -\frac{1}{5} j_\ell(k\eta_0 - k\eta_{\mathrm{dec}})$. Inserting it to Eq. (2.79) and assuming a scale invariant primordial power spectrum ($\mathcal{P}_{\mathcal{R}}(k) = A = const.$) we obtain the angular power spectrum of the Sachs-Wolfe effect

$$C_\ell^{\mathrm{SW}} = \frac{4\pi A T_0^2}{25} \int_0^\infty \frac{dk}{k} j_\ell^2(k\eta_0 - k\eta_{\mathrm{dec}}) = \frac{A T_0^2}{25} \frac{2\pi}{\ell(\ell+1)}. \tag{2.93}$$

We can see that $\frac{\ell(\ell+1)}{2\pi} C_\ell^{\mathrm{SW}}$ has a constant value. This is the reason, why the angular power spectrum is often displayed as $\frac{\ell(\ell+1)}{2\pi} C_\ell$. If the spectral index n (in $\mathcal{P}_{\mathcal{R}}(k) \propto k^{n-1}$) is larger than 1, $\frac{\ell(\ell+1)}{2\pi} C_\ell^{\mathrm{SW}}$ will increase with increasing ℓ. If $n < 1$ the opposite will occur.

2.7.2 Small Scales - Acoustic Oscillations

We consider the monopole and dipole terms of $\Theta_\ell(\eta_0, \mathbf{k})$ (see Eq. (2.88)) first and ignore the ISW term for the time being. We will discuss the ISW effect at the end of this section.

The quantities $\Theta_0(\eta_{\mathrm{dec}}, \mathbf{k})$, $\Theta_1(\eta_{\mathrm{dec}}, \mathbf{k})$ and $\Phi(\eta_{\mathrm{dec}}, \mathbf{k})$ required in the calculation of $\Theta_\ell(\eta_0, \mathbf{k})$ are the growing mode solution of the Boltzmann hierarchy (Eqs. (2.60)

- (2.69)) and Einstein equations of Ψ and Φ (Eqs. (2.15) - (2.18)). We assume that $-\tau' = an_{\rm e}\sigma_{\rm T}$ is large at $\eta \leq \eta_{\rm dec}$ (tight-coupling limit), which leads to a photon brightness function whose main multipoles are monopole and dipole ([24]). In this approximation we can ignore photon multipoles $\Theta_\ell(\eta, \mathbf{k})$ at $\ell \geq 2$ in the Boltzmann hierarchy when solving $\Theta_0(\eta, \mathbf{k})$ and $\Theta_1(\eta, \mathbf{k})$.

We consider here adiabatic initial conditions, which are specified at an early radiation-dominated era ($\eta = \eta_{\rm rad}$, $\eta_{\rm rad} \ll \eta_{\rm dec}$), when all relevant scales were well outside horizon (see Sect. 2.2). At that time the Fourier modes of the curvature perturbation had values $\mathcal{R}_{\mathbf{k}}({\rm rad})$. They stay in these values as long as the corresponding scales are outside horizon. In linear perturbation theory the initial values of the other perturbations (Θ_0, Θ_1, Ψ, Φ, $v_{\rm b}$ etc.) at $\eta = \eta_{\rm rad}$ are obtained from the Boltzmann hierarchy and Einstein equations and they depend linearly on $\mathcal{R}_{\mathbf{k}}({\rm rad})$. At later times ($\eta > \eta_{\rm rad}$) these perturbations evolve differently in time but the linear dependence on $\mathcal{R}_{\mathbf{k}}({\rm rad})$ will remain.

The adiabatic initial values of the perturbations can be expressed in terms of $\mathcal{R}_{\mathbf{k}}({\rm rad})$ ([20], [27])

$$\begin{aligned}
\Phi(\eta, \mathbf{k}) &= -\frac{2}{3}\left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
\Psi(\eta, \mathbf{k}) &= -\frac{2}{3}\left(1 + \frac{2}{5}f_\nu\right)\left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
\delta_\nu(\eta, \mathbf{k}) = \delta_\gamma(\eta, \mathbf{k}) = \delta(\eta, \mathbf{k}) &= \frac{4}{3}\left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
\delta_{\rm c}(\eta, \mathbf{k}) = \delta_{\rm b}(\eta, \mathbf{k}) = \frac{3}{4}\delta(\eta, \mathbf{k}) &= \left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
v_\nu(\eta, \mathbf{k}) = v_\gamma(\eta, \mathbf{k}) = v(\eta, \mathbf{k}) &= -\frac{1}{3}(k\eta)\left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
v_{\rm c}(\eta, \mathbf{k}) = v_{\rm b}(\eta, \mathbf{k}) = v(\eta, \mathbf{k}) &= -\frac{1}{3}(k\eta)\left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
\Theta_\ell^\nu &= -\frac{(k\eta)^\ell}{(2\ell+1)!!}\frac{1}{3}\left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
\Pi_\nu(\eta, \mathbf{k}) &= -\frac{4}{15}(k\eta)^2\left(1 + \frac{4}{15}f_\nu\right)^{-1} \mathcal{R}_{\mathbf{k}}({\rm rad}) \\
\Pi_\gamma &= 0. \qquad (2.94)
\end{aligned}$$

These equations are valid for the times η ($\eta \geq \eta_{\rm rad}$), when the universe was radiation dominated and for scales k that were well outside horizon at η. Here f_ν is the energy density of neutrinos as a fraction of the total energy density of radiation (its value is $f_\nu \sim 0.405$). The quantities Θ_ℓ^ν are the multipoles of the neutrino brightness function with the following relations to the neutrino perturbations (cf. corresponding photon

relations in Eq. (2.59)): $\Theta_0^\nu = \delta_\nu/4$, $\Theta_1^\nu = v_\nu/3$, and $\Theta_2^\nu = \Pi_\nu/12$. The symbol $(2\ell+1)!!$ is defined as $(2\ell+1)!! \equiv 1 \cdot 3 \cdot 5 \cdots (2\ell+1)$.

As the initial conditions of the perturbations depend linearly on $\mathcal{R}_\mathbf{k}(\mathrm{rad})$ (see Eq. (2.94)) and as we assume linear perturbation theory, $\Theta_\ell(\eta_0, \mathbf{k})$ will depend linearly on $\mathcal{R}_\mathbf{k}(\mathrm{rad})$ as well: $\Theta_\ell(\eta_0, \mathbf{k}) = T_\ell(\eta_0, k)\mathcal{R}_\mathbf{k}(\mathrm{rad})$, where $T_\ell(\eta_0, k)$ is the transfer function. If we choose an initial value $\mathcal{R}_\mathbf{k}(\mathrm{rad}) = 1$, we can equate the right hand side of Eq. (2.88) with the transfer function $T_\ell(\eta_0, k)$ (ignore ISW term and use $k\eta_0 - k\eta_\mathrm{dec} \approx k\eta_0$)

$$T_\ell(\eta_0, k) = \left[\Theta_0(\eta_\mathrm{dec}, k) + \Phi(\eta_\mathrm{dec}, k)\right] j_\ell(k\eta_0) + 3\Theta_1(\eta_\mathrm{dec}, k) j_\ell'(k\eta_0). \tag{2.95}$$

Inserting this transfer function to Eq. (2.79) we obtain the angular power spectrum (assuming a scale invariant primordial power spectrum)

$$C_\ell = 4\pi A T_0^2 \int_0^\infty \frac{dk}{k} \left[(\Theta_0 + \Phi)^2 j_\ell^2(k\eta_0) + 6(\Theta_0 + \Phi)\Theta_1 j_\ell(k\eta_0) j_\ell'(k\eta_0) + 9\Theta_1^2 j_\ell'^2(k\eta_0)\right] \tag{2.96}$$

The cross term $j_\ell j_\ell'$ will oscillate with positive and negative values and will thus integrate to a small value ([24]). Non-mixed monopole and dipole contributions remain significant.

When a scale k enters the horizon in the radiation dominated era (before photon decoupling), the corresponding monopole and dipole terms start to oscillate and they reach some values $[\Theta_0(\eta_\mathrm{dec}, k) + \Phi(\eta_\mathrm{dec}, k)]$ and $\Theta_1(\eta_\mathrm{dec}, k)$ at the photon decoupling. Before photon decoupling the dipole term is proportional to the time derivative of the monopole. Therefore the monopole and dipole terms are out of phase with one another. The amplitude of the dipole term is smaller than the amplitude of the monopole term. Schematic plots of the squared monopole ($[\Theta_0(\eta_\mathrm{dec}, k) + \Phi(\eta_\mathrm{dec}, k)]^2$) and dipole terms ($[3\Theta_1(\eta_\mathrm{dec}, k)]^2$) are shown in Fig. 2.1. Because the magnitude k of the wavevector and the multipole ℓ have the relation $\ell \approx k\eta_0$, we can use ℓ in the x-axis instead of k. Note that the curves in Fig. 2.1 have not been derived from any particular cosmological model, but they just show how the monopole and dipole terms behave in principle. The x-axis scaling and the relative magnitudes between the terms have been chosen arbitrarily.

The monopole term $(\Theta_0 + \Phi)^2$ dominates over the dipole term $(3\Theta_1)^2$ in the angular power spectrum. Therefore it will have the same oscillations as the monopole term (see Fig. 2.1). The first peak of the angular power spectrum is where $[\Theta_0(\eta_\mathrm{dec}, k) + \Phi(\eta_\mathrm{dec}, k)]^2$ has its first maximum (at $\ell \approx 200$ in Fig. 2.1). This is called the first *acoustic peak*. The corresponding multipole is $\ell \approx \pi\eta_0/r_\mathrm{s}(\eta_\mathrm{dec})$. Here $r_\mathrm{s}(\eta)$ is the sound horizon

$$r_\mathrm{s}(\eta) = \int_0^\eta d\eta' c_\mathrm{s}(\eta'). \tag{2.97}$$

Actually the first acoustic peak occurs at $\sim$25% lower ℓ ([24]), because the k integration with j_ℓ^2 collects contribution also from other k than $k = \ell/\eta_0$ and because we ignored the ISW term. Since the dipole term is out of phase with the monopole term, it fills the troughs between the acoustic peaks.

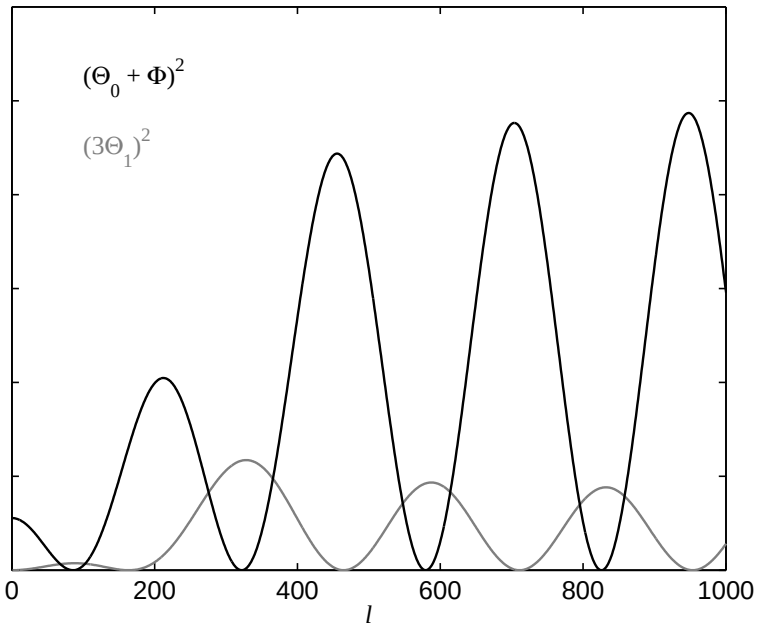

Figure 2.1: Schematic plot of the squared monopole and dipole terms ($(\Theta_0 + \Phi)^2$ and $(3\Theta_1)^2$) at the photon decoupling ($\eta = \eta_{\rm dec}$). The curves do not represent any cosmological model. They just show the principal behavior of the terms. The x-axis scale is arbitrary. ℓ can be used there instead of k because they are related as $\ell \approx k\eta_0$.

The ISW term (third term of Eq. (2.88)) gives no contribution before the photon decoupling, because $e^{-\tau} \ll 1$ then. After the photon decoupling it contributes when the potentials Ψ or Φ are changing. They are constant when the universe is well in the matter dominated era. After the photon decoupling the universe is still in a transition phase from the radiation domination to the matter domination and the potentials are changing. This is called the early ISW. Its influence is largest to the scales that are entering the horizon at that time ([24]). The corresponding ℓ are typically around the first acoustic peak and to the left (smaller ℓ) of it. Therefore the early ISW boosts the magnitude of the first acoustic peak and shifts it to smaller ℓ. The next time when the potentials change is when the cosmological constant becomes significant relative to matter. This happens late and it impacts mainly the scales that are entering the horizon at that time. These are large scales which influence mainly the lowest multipoles of the angular power spectrum. This effect is called the late ISW.

Perturbations at small scales are reduced by diffusion of the photons. Therefore the amplitudes of the acoustic oscillations decrease with increasing k. This pushes down the higher acoustic peaks of the angular power spectrum. This effect is called the diffusion damping. It is not included in Fig. 2.1.

2.8 C_ℓ and Cosmological Parameters

The CMB anisotropy spectrum depends on cosmological parameters. The estimation of these parameters from the CMB observations is an important final step in a major CMB experiment. We will not discuss here the methods used for this estimation but we give a brief overview how the CMB angular power spectrum depends on these parameters. The cosmological parameters that we consider here are listed in Table 2.1. Additionally we assume that all neutrinos are massless, dark energy equation-of-state parameter is w = -1 (i.e., dark energy is equivalent to the cosmological constant), perturbations are scalar only (no tensor perturbations) and the perturbations are adiabatic.

Table 2.1: Cosmological parameters used in this study. Note that these parameters are not independent. Some of them can be expressed in terms of the other. Reference values chosen for the parameters are shown in the third column.

Symbol	Description	Ref. value
Ω	Total density parameter (a)	1.0
A	Amplitude of the primordial power spectrum	(b)
n	Spectral index of the primordial curvature perturbation	1.0
$\tau_{\rm r}$	Optical depth due to reionization	0
h	Hubble constant	0.71 (c,d)
Ω_Λ	Cosmological constant density parameter	0.73 (c)
$\Omega_{\rm b}$	Baryon density parameter	0.044 (c)
$\Omega_{\rm m}$	Matter density parameter	0.27 (c,e)

(a) Total energy density as a fraction of the critical density. $\Omega = \Omega_{\rm m} + \Omega_\Lambda$.
(b) Refers to the power spectrum of the primordial curvature perturbations. We use COBE normalization implemented in CMBFAST ([28],[30]), which fixes the power at large scales corresponding to $\ell \sim 10$.
(c) Taken from the WMAP first year data (Table 3 of [7]).
(d) In units of 100 kms^{-1}Mpc^{-1}. Ref. value corresponds to H_0 = 71 kms^{-1}Mpc^{-1}.
(e) $\Omega_{\rm m} = \Omega_{\rm b} + \Omega_{\rm c}$, where $\Omega_{\rm c}$ is the CDM density parameter.

We used the CMBFAST code ([26], [28]) to produce the angular power spectra from the cosmological parameters. The spectrum of the reference values (reference spectrum) is shown in Fig. 2.2 (black curve). It shows many of those features that we discussed in the previous sections. The first acoustic peak is clearly visible at $\ell \approx 220$. The remaining peaks are significantly lower and their amplitudes decrease with increasing ℓ (diffusion damping).

The geometry of the universe influences strongly the location of the acoustic peaks. The comoving angular diameter distance to $\eta = \eta_{\rm dec}$ is larger in an open universe and smaller in a closed universe than in the flat universe ([24]). In a flat

universe this comoving angular diameter distance is $\eta_0 - \eta_{\rm dec}$. An identical pattern of inhomogeneity at $\eta = \eta_{\rm dec}$ (like an acoustic peak) gets projected to smaller angular scales in an open universe than in a closed universe (due to larger angular diameter distance in the open universe). Therefore an acoustic peak will shift towards higher ℓ in an open universe and towards lower ℓ in a closed universe (from its position in a flat universe).

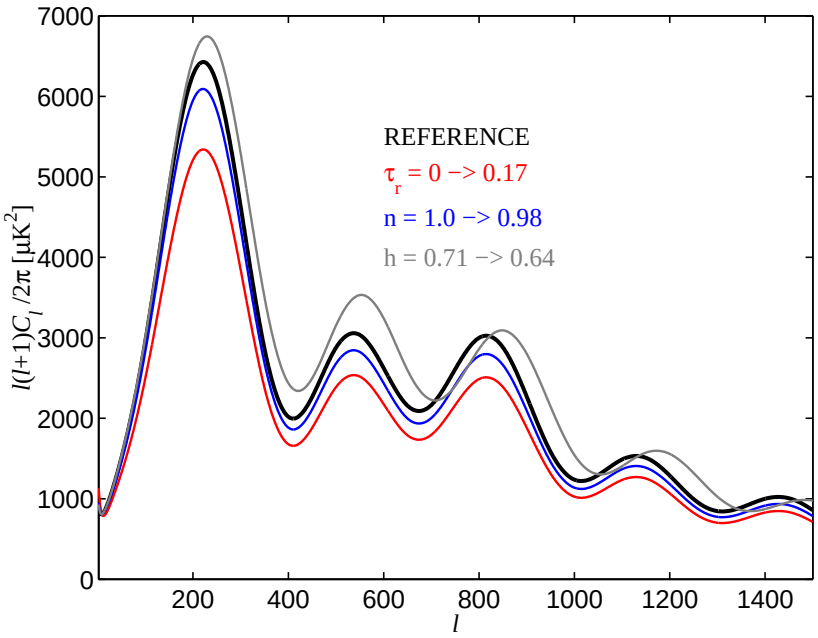

Figure 2.2: Angular power spectra from different sets of cosmological parameters. The reference spectrum (black curve) is from the reference parameters (see Table 2.1). The red curve is for the optical depth $\tau = 0.17$, the blue curve is for a tilted primordial curvature spectrum ($n = 0.98$) and the gray curve is for $h = 0.64$ Hubble constant. The other parameters had their reference values in these spectra.

For the remaining parts of this section we assume a flat universe ($\Omega = 1$). Fig. 2.2 shows three angular power spectra where one cosmological parameter has been varied, while the rest of the parameters had their reference values. Increasing the optical depth $\tau_{\rm r}$ from 0 to 0.17 ([7]) increases the number of photons that have rescattered from the electrons since the photon decoupling. In the observed CMB the rescattered photons come from the different locations of the $\eta = \eta_{\rm dec}$ sky than the unscattered photons (no scatterings since $\eta = \eta_{\rm dec}$). This tends to blurr the CMB anisotropy and the power of the spectrum is reduced. This can be noted in the "$\tau_{\rm r}$" curve of Fig. 2.2.

The "n" curve shows what happens when the spectral index of the primordial curvature perturbations is decreased from 1.0 to 0.98. Due to this spectral tilt there is less power in small scales and the angular spectrum will be correspondingly lower

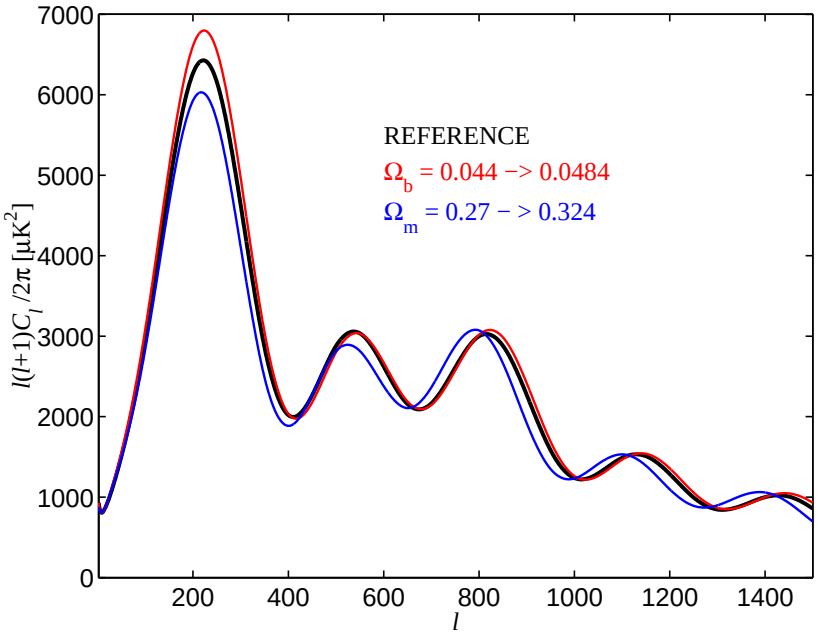

Figure 2.3: Influence of the matter and baryon density parameters on the CMB angular power spectrum. The black curve is the reference spectrum (same as in Fig. 2.2).

at high ℓ. This can be seen in the "n" curve of Fig. 2.2. If the spectral tilt would be opposite ($n > 1$), the power at high ℓ would be increased from its reference value.

For the "h" curve of Fig. 2.2 the Hubble constant was reduced by 10% (from 0.71 to 0.64). Due to smaller H_0, η_0 will be increased (see Eq. (2.89)) and a perturbation at scale k will be shifted towards higher ℓ. This shift can be detected in the "h" curve. The effect of the reduced h on the height of the first acoustic peak is discussed at the end of this section.

The influence of the baryon and matter density parameters is shown in Fig. 2.3. For the "$\Omega_{\rm b}$" curve we increased $\Omega_{\rm b}$ by 10% (from 0.044 to 0.0484) but kept the other parameters in their reference values (see Table 2.1). $\Omega_{\rm m}$ and Ω_Λ were kept in their reference values by reducing the CDM density parameter correspondingly. Larger density of baryons slows down the sound speed of the baryon-photon fluid leading to a smaller sound horizon during photon decoupling. This shifts the acoustic peaks towards higher ℓ, which can be seen in the "$\Omega_{\rm b}$" curve. Increased baryon density also brings a positive offset in $\Theta_0(\eta_{\rm dec}, k) + \Phi(\eta_{\rm dec}, k)$. Squaring it enhances the odd peaks and reduces the even peaks.

The increased baryon (electron) density means a shorter mean free path of photons before photon decoupling. This reduces the scale where the diffusion damping becomes significant. Therefore the diffusion damping of the angular power spectrum shifts to higher ℓ and the height of the high ℓ acoustic peaks will correspondingly

increase. This phenomenon is typicaly visible at high ℓ ($\ell \gtrsim 1500$) and is therefore barely detectable in Fig. 2.3.

For the "Ω_{m}" curve of Fig. 2.3 we increased Ω_{m} by 20% (from 0.27 to 0.324). We kept Ω_{b} in its reference value but reduced Ω_Λ (from 0.73 to 0.676) to maintain $\Omega = 1$. The other parameters were kept in their reference values. In this case the influence to the angular power spectrum is more complex. The value of η_0 will decrease (see Eq. (2.89)) which pushes the acoustic peaks towards lower ℓ. On the other hand, increased Ω_{m} will lead to an earlier matter-radiation equality which tends to reduce the sound horizon. This pushes the acoustic peaks towards higher ℓ. In this case the total effect is towards lower ℓ as can be seen in the "Ω_{m}" curve. The earlier matter-radiation equality reduces the change of the potentials Ψ and Φ after photon decoupling. This reduces the magnitude of the early ISW effect which decreases the height of the first acoustic peak.

Finally, let us return back to the "h" curve of Fig. 2.2. Reducing h while keeping the other parameters in their reference values means that the present matter energy density is correspondingly reduced. Because the present radiation energy density is not affected, the matter-radiation equality shifts towards later times. This enhances the early ISW effect which increases the first acoustic peak. This increase can be detected in the "h" curve of Fig. 2.2.

Chapter 3

Overview of a Satellite CMB Experiment

A successful on-ground data processing (e.g. map-making and angular power spectrum estimation) requires a good understanding of the relevant aspects of the setup of the CMB experiment. This chapter gives an overview of those aspects in a satellite based CMB experiment using the PLANCK mission ([8], [9]) as an example.

3.1 Instrumentation

The PLANCK satellite is shown in Fig. 3.1 ([8]). The received photons hitting the primary mirror (item A in Fig. 3.1) of the telescope are directed to the focal plane unit (FPU, item C in Fig. 3.1) that contains the horn antennas and the front end electronics of the radiometers. The diameter of the primary mirror is about 1.5 m. For radiometers operating at the lower frequency channels the angular resolution is set by the diameter of the primary mirror and the wavelength of the observed photons. The beam widths of these radiometers are proportional to the ratio between the wavelength and the diameter of the primary mirror (diffraction limited beams). The horn antennas of the higher frequency radiometers are designed to under-illuminate the telescope leading to the same nominal beam widths for these channels. The beams are diffraction limited for the frequency channels lower than $\sim$150 GHz.

The PLANCK satellite contains 74 detectors in total. They are split in two instruments. The Low Frequency Instrument (LFI) contains 22 radiometers covering three frequency bands with band centers at 30 GHz, 44 GHz and 70 GHz ([31]). All radiometers are sensitive to one linear polarisation and two such radiometers (with orthogonal polarisation sensitive directions) are connected to one horn antenna. The active microwave electronics of the LFI is implemented with semiconductor components.

The High Frequency Instrument (HFI) contains 52 detectors covering six frequency bands with band centers at 100 GHz, 143 GHz, 217 GHz, 353 GHz, 545 GHz and 857 GHz ([33]). All HFI detectors are bolometers. HFI contains both polarisa-

tion sensitive bolometers (PSB) and polarisation insensitive bolometers (spider web bolometers).

A detailed view of the FPU is shown in Fig. 3.2 ([8]). The LFI and HFI horn antennas are clearly visible. The set of horns as projected in the sky is shown in Fig. 3.3 ([8]). When in orbit the satellite will spin around its spin axis (vertical line in Fig. 3.3) at a nominal rate of 1 round per minute (rpm). The satellite line of sight (LOS) indicates the direction where the incoming photons will hit the center of the FPU. The LOS and the spin axis are nearly perpendicular thus allowing the beams to draw circles on the sky while the satellite spins. Each detector will sample the intensity of the received photons with constant sampling intervals in time.

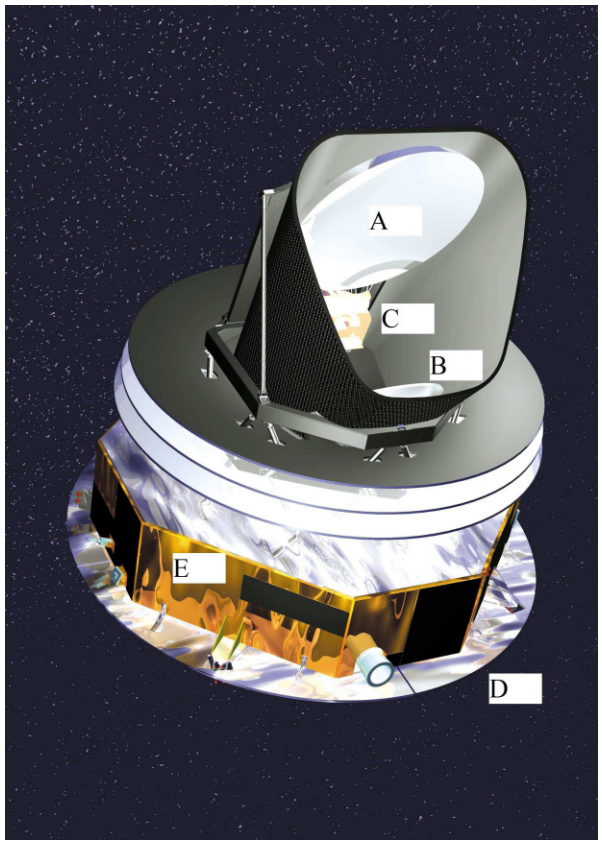

Figure 3.1: Schematic view of the Planck satellite. The primary (A) and the secondary (B) mirrors of the telescope are shown. The received photons hit the radiometer horns located in the focal plane (C). The solar panels (D) and the service module (E) containing most of the electronics and the cooling systems are shown as well. The figure is adapted from [8].

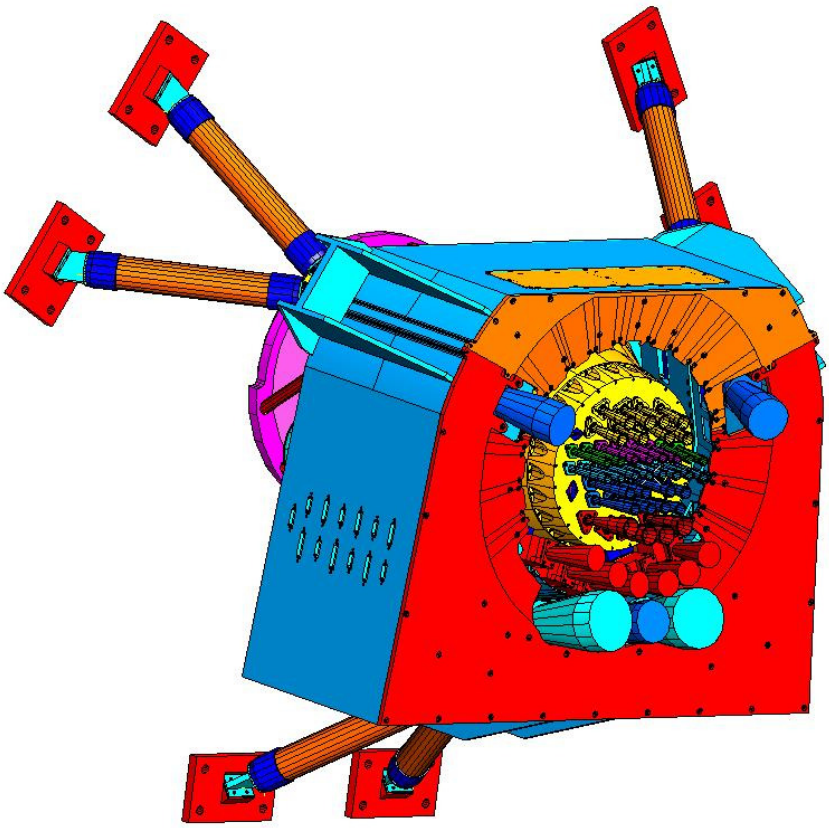

Figure 3.2: A detailed view of the Planck focal plane. The LFI horns are located on the outer circles of the FPU and the HFI horns (the smaller ones) are placed in the center. FPU also contains the HFI bolometers and the front end sections of the LFI radiometers. The main body of the FPU is cooled to 20 K which is the ambient temperature of the LFI front ends. HFI horns in the middle are cooled to 4 K and the bolometer sensors (at the rear parts of the FPU) are cooled to 0.1 K. The amplified microwave signals from the LFI and the post-detection signals from the HFI are coupled to the service module of the satellite (item E in Fig. 3.1). The figure is from [8].

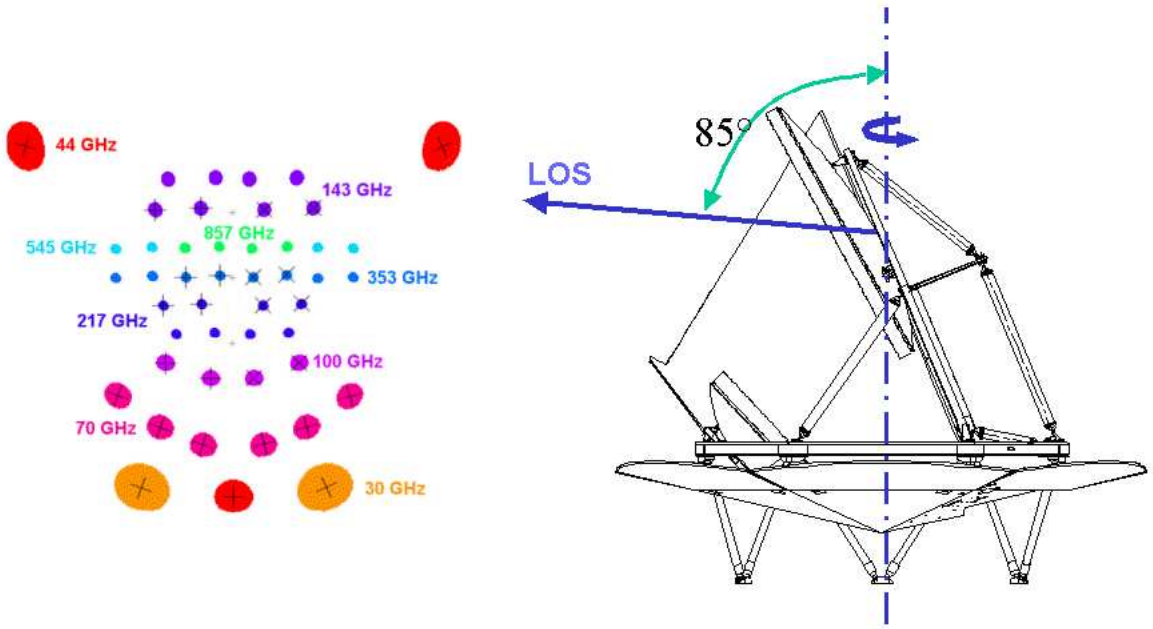

Figure 3.3: The foot-print of the FPU horn antennas projected in the sky. The line-of-sight (LOS) points approximately to the center of the foot-print. The angle (85 deg) between the LOS and the satellite spin axis is shown. The scanning of the foot-print is from left to right (in the horizontal direction). The crosses indicate the polarisation sensitive directions of the pair of detectors in a horn. Horns with no cross contain a polarisation insensitive detector. The radiation patterns are indicated by the spots. Their diameters represent their relative beam widths, and their shapes are relative to their ellipticities. The field of view of the foot-print is about 8 deg in its maximum width. The figure is from [8].

3.1.1 Low Frequency Instrument

The LFI is built from 22 pseudo-correlation radiometers ([32]). Two such radiometers are connected to a single horn antenna. The front end section is shown in Fig. 3.4 (redrawn from [32]). It operates in 20 K ambient temperature. The stream of photons from the horn antenna is split to two radiometers in the orthomode transducer (OMT). OMT separates the incoming photons to two streams with orthogonal linear polarisations. The radiometers following the OMT are identical and are not polarisation sensitive themselves. Referring to two orthogonal linear polarisations they are often called X and Y radiometers.

The details of the front-end unit of one of the radiometers is shown in Fig. 3.4. The radiometer receives photons simultaneously from the sky and from the reference blackbody load at 4 K temperature. The sky and the reference signals are amplified

and further directed to the back-end unit (see Fig. 3.5, redrawn from [32]) that is located in the service module of the satellite (item E in Fig. 3.1). The back-end unit operates in 300 K ambient temperature. In the back-end unit the signal is amplified, filtered and detected. The diode detectors and the post-detection low pass filters (not shown in the figure) act as square-and-integrate devices producing a quantity that is directly proportional to the mean microwave power at the diode input.

The gains and the noise temperatures of the semiconductor amplifiers fluctuate slowly in time ([32]) which induce correlated ($1/f$ type) low-frequency noise in the output of the radiometer chain. This noise would be unacceptably high in a simple total power receiver ([32]). The pseudo-correlation receiver can efficiently suppress the power of the $1/f$ noise. During the radiometer operation the relative phase shift between the radiometer legs (Fig. 3.4) is periodically modulated between 0 and 180 deg with the phase switches. As a result of this modulation the post-detection signal alternates between the sky and the reference signals. The phase of the alternation is opposite in the other leg of the radiometer. Due to the different temperatures of the sky and the load the magnitudes of their post-detection signals are different. In the LFI the modulation period is 1/4096 s.

The data acquisition section (Fig. 3.5) takes one sky and one reference sample from each modulation period. A number of sky and reference samples are averaged to produce sky and reference sample streams with the final output sampling rate. Each detector has its own output sampling rate and their values for the LFI are given in Table 3.2. Sky and reference sample streams are produced in both radiometer legs leading to 4 parallel streams per one pseudo-correlation radiometer.

Because the sky and the reference signals are passing through the same chain of amplifiers they are exposed to the same gain and noise temperature fluctuations. Subtracting the reference signal from the sky signal will reduce the low-frequency fluctuations considerably. Because the sky temperature differs from the reference temperature the subtraction need to be done with unequal weights: SKY - $r\times$REF, where r is the gain modulation factor ([32]). It is determined from the constraint that the mean of the difference should be zero. The way how the differencing reduces the magnitude of the $1/f$ noise is clarified in Fig. 3.6. The sky and the reference streams with high $1/f$ noise (large gradients) are shown. In the differencing the level of the $1/f$ noise is reduced significantly. The reduction would be perfect (assuming perfect match between the legs) if the sky and the reference temperatures would be equal. Because they are not some residual $1/f$ noise always remains. The stream of difference samples (sampling rate according to Table 3.2) is the final output of the LFI radiometer and the difference samples constitute the raw time-ordered data (TOD) stream of the radiometer.

The response times of the LFI radiometers are determined by the bandwidths of their post-detection filters. Their response times to the changes of the photon intensities are clearly shorter than the response times of the HFI bolometers (cf. Sect. 3.1.2).

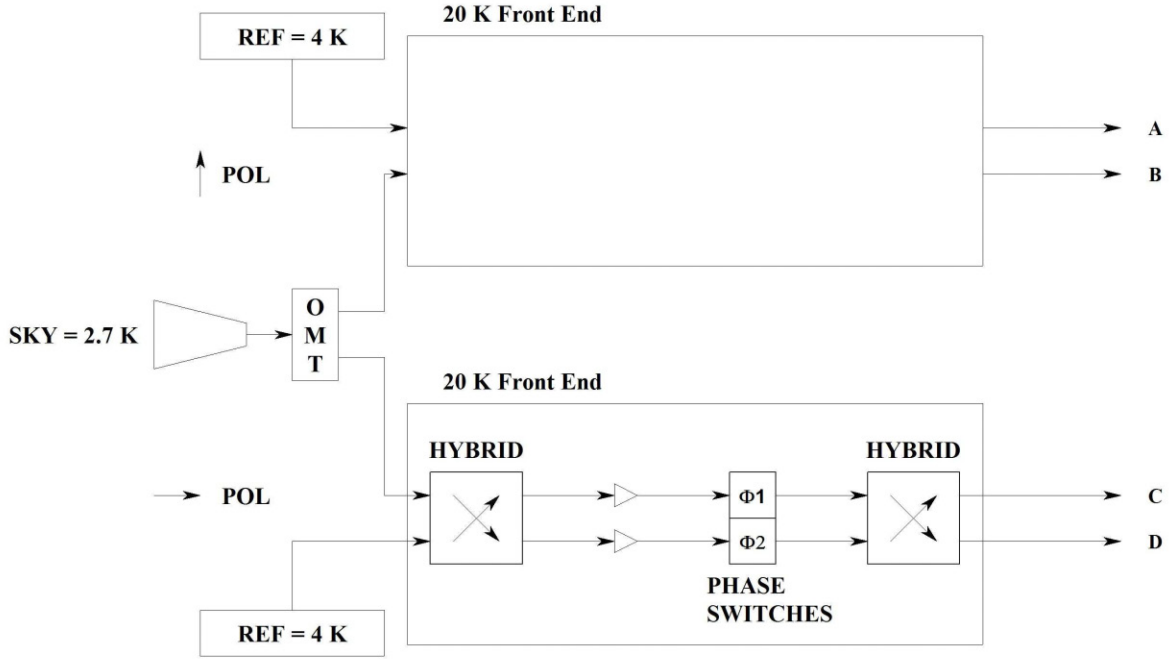

Figure 3.4: Front-end unit of an LFI radiometer.

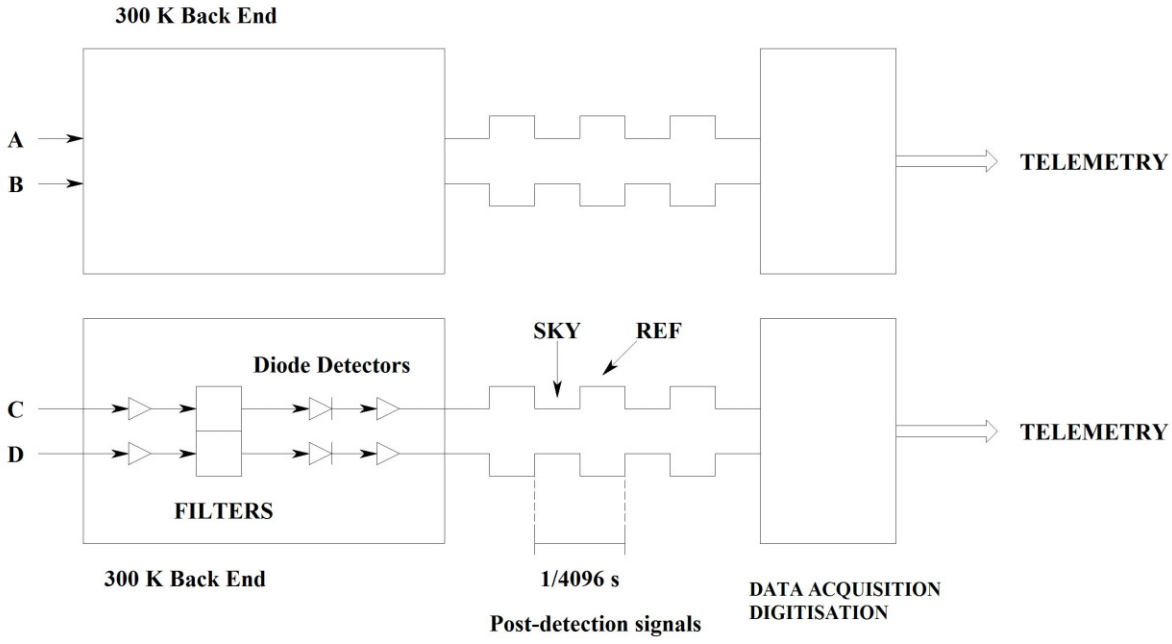

Figure 3.5: Back-end unit of an LFI radiometer.

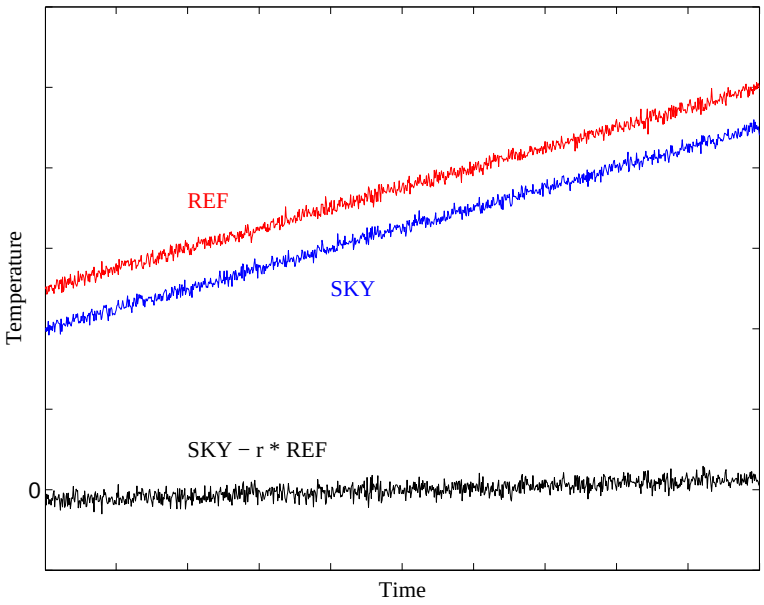

Figure 3.6: Principle of the reduction of the $1/f$ noise in the LFI radiometer.

3.1.2 High Frequency Instrument

A cut view of the HFI is shown in Fig. 3.7 ([8]). The photons from the telescope are delivered to the bolometer sensors by the back-to-back horn antennas located in the focal plane unit. The heart of the bolometer sensor is a resistor (thermistor) whose resistance value depends on its temperature in a known way ([33]). The fluctuations in the intensity of the photons heating the thermistor will cause fluctuations in its temperature. This leads to the fluctuations in the resistance value which are detected. The detected signal is low pass filtered to the Nyquist critical frequency (cut off at the half of the sampling frequency) and then sampled. The sampling rate values are given in Table 3.2. For a review of the bolometer techniques, see [34].

The bolometer resistance follows the changes of the photon intensity with some delay. This delay is characterized by a bolometer response time ([35]). The bolometer response performs a low pass filtering in the time stream of data. This response needs to be removed from the received data in the data processing phase. This can be carried out by deconvolving the data with an inverted bolometer response. These techniques are beyond the scope of this study. Typical time constants for the HFI bolometers are in the order of some milliseconds ([35]).

3.1.3 Telescope Beams

We consider the responses of the telescope main beams to the total intensity sky signal. Telescope sidelobes are not considered in this study. The spots in Fig. 3.3 indicated the relative angular widths and the shapes of the main beam responses

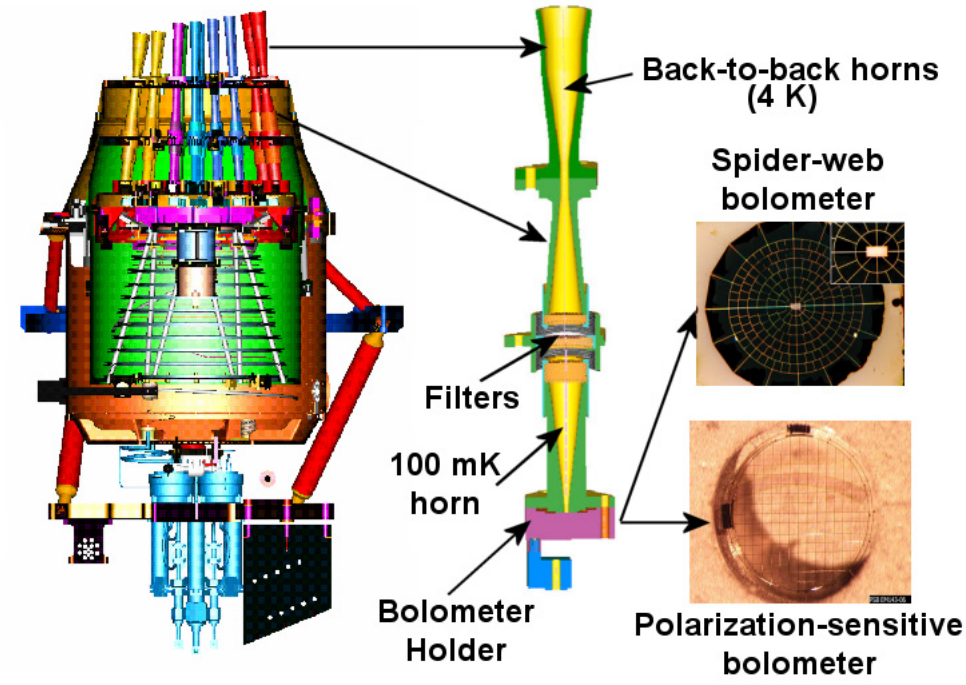

Figure 3.7: A cut view of the HFI focal plane unit. The back-to-back horns collect the incoming photons from the telescope and deliver them to the bolometer sensor. The sensors are either polarisation sensitive or insensitive (spider web). Blow-up views of both are shown. The figure is from [8].

of the PLANCK detectors. The angular width decreases (leading to better angular resolution) with increasing frequency. The shapes of the responses tend to become more asymmetric when the detector horn is farther away from the LOS.

The main beam response $B(\mathbf{n}, \mathbf{n}_0, \psi)$ is a field in the sky depending on three variables: $\mathbf{n} = (\theta, \varphi)$ are the spherical coordinates of the field point, $\mathbf{n}_0 = (\theta_0, \varphi_0)$ is the pointing of the maximum intensity of the beam response (or some other suitably defined point in the beam) and ψ is the orientation angle of the beam (rotation around $\mathbf{n}_0$). The radiation pattern is normalized to $\int_{4\pi} d\Omega_{\mathbf{n}} B(\mathbf{n}, \mathbf{n}_0, \psi) = 1$.

One can imagine that $B(\mathbf{n}, \mathbf{n}_0, \psi)$ is obtained by rotating the beam from its reference pointing and orientation to the actual pointing $\mathbf{n}_0$ and orientation ψ. The selection of the reference position is arbitrary but $\mathbf{n}_0 = \hat{\mathbf{z}}$ (pointing to the north pole) and $\psi = 0$ is assumed in this study. This leads for the beam response

$$B(\mathbf{n}, \mathbf{n}_0, \psi) = \widehat{D}(\mathbf{n}_0, \psi) B(\mathbf{n}, \hat{\mathbf{z}}, 0), \tag{3.1}$$

where $B(\mathbf{n}, \hat{\mathbf{z}}, 0)$ is the beam response in its reference pointing and orientation and

$\widehat{D}(\mathbf{n}_0, \psi) = \widehat{D}(\varphi_0, \theta_0, \psi)$ is the active right-handed rotation operator corresponding to the Euler angles $\alpha = \varphi_0$, $\beta = \theta_0$ and $\gamma = \psi$ ([36]).

The beam in its reference pointing and orientation can be expanded in spherical harmonics

$$B(\mathbf{n}, \hat{\mathbf{z}}, 0) = \sum_{\ell m} b_{\ell m} Y_{\ell m}(\mathbf{n}), \tag{3.2}$$

where $b_{\ell m}$ are the expansion coefficients. The normalization of the response leads to $b_{00} = 1/\sqrt{4\pi}$. After inserting Eq. (3.2) in to Eq. (3.1) and utilizing the relation ([37])

$$\widehat{D}(\mathbf{n}_0, \psi) Y_{\ell m}(\mathbf{n}) = \sum_{m'=-\ell}^{\ell} D^{\ell}_{m'm}(\varphi_0, \theta_0, \psi) Y_{\ell m'}(\mathbf{n}), \tag{3.3}$$

where $D^{\ell}_{m'm}(\varphi_0, \theta_0, \psi)$ are the Wigner D-functions (the matrix elements of the rotation operator, [36]), one obtains for the beam response

$$B(\mathbf{n}, \mathbf{n}_0, \psi) = \sum_{\ell m} b_{\ell m}(\mathbf{n}_0, \psi) Y_{\ell m}(\mathbf{n}), \tag{3.4}$$

where the pointing and orientation dependent expansion coefficients are

$$b_{\ell m}(\mathbf{n}_0, \psi) = \sum_{m'=-\ell}^{\ell} b_{\ell m'} D^{\ell}_{mm'}(\varphi_0, \theta_0, \psi). \tag{3.5}$$

Due to their small angular extent the main beam responses of the PLANCK telescope can be approximated with flat sky elliptic Gaussian beams ([38], [39]). We consider an elliptic Gaussian beam with its center pointing towards north pole ($\mathbf{n}_0 = \hat{\mathbf{z}}$) and its major axis aligned along the x-axis of an (x,y) coordinate system placed on the surface of the sky sphere (Fig. 3.8). The beam response is ([39])

$$B(\mathbf{n}, \hat{\mathbf{z}}, 0) = B(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}}. \tag{3.6}$$

The relation between the (x,y) coordinates and the spherical coordinates (θ, φ) is given in Fig. 3.8. The quantities σ_x and σ_y give the angular extents of the major and minor axes of the elliptic response. Instead of these two quantities the full width half maxima (FWHM) of the major and minor axes are usually used. $\mathrm{FWHM_x}$ is the quantity $\mathrm{x_0}$ solved from the equation $B(x = \frac{x_0}{2}, y = 0) = \frac{1}{2}B(x = 0, y = 0)$. $\mathrm{FWHM_y}$ is defined similarly. One obtains $\mathrm{FWHM_x} = \sqrt{8\ln 2}\sigma_\mathrm{x}$ and $\mathrm{FWHM_y} = \sqrt{8\ln 2}\sigma_\mathrm{y}$. The mean FWHM values for the PLANCK beams are shown in Table 3.2.

The coefficients ($b_{\ell m}$) of the spherical harmonic expansion of the elliptic Gaussian beam response are ([40])

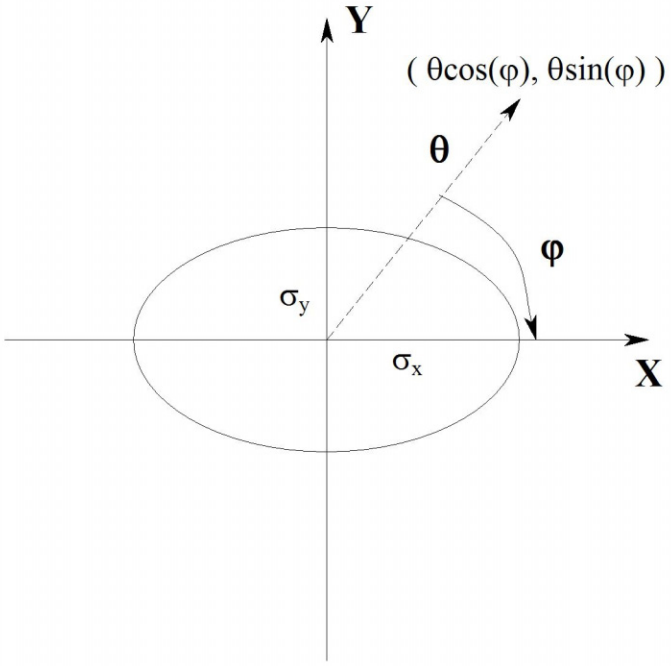

Figure 3.8: Elliptic Gaussian beam pointing to the north pole ($\mathbf{n}_0 = \hat{\mathbf{z}}$). A small patch of the sky is approximated as a flat plane and (x,y) coordinate system is placed on it. The z-axis points towards the reader. The major axis of the ellipse is aligned along the x-axis. The relation between the (x,y) coordinates and the spherical coordinates (θ, φ) is shown.

$$b_{\ell m} = \left[\frac{2\ell+1}{4\pi}\frac{(\ell+|m|)!}{(\ell-|m|)!}\right]^{\frac{1}{2}} \ell^{-|m|} I_{|m|/2}\left[\frac{\ell(\ell+1)\sigma^2(\epsilon-1/\epsilon)}{4}\right] \cdot \\ \cdot \exp\left[-\frac{\ell(\ell+1)\sigma^2(\epsilon+1/\epsilon)}{4}\right], \tag{3.7}$$

where σ is the geometric mean of σ_x and σ_y ($\sigma = \sqrt{\sigma_x\sigma_y}$) and $\epsilon = \sigma_x/\sigma_y$ ($\epsilon \geq 1$). $I_\nu(x)$ is the modified Bessel function of the first kind. Due to the 180 deg rotational symmetry the expansion coefficients with odd m are zero. The expansion coefficients are shown at various ℓ and m in Fig. 3.9.

If the beam response is approximated with a symmetric Gaussian response ($\sigma_x = \sigma_y$, leading to $\epsilon = 1$) the expansion coefficients with $m \neq 0$ will all be zero due to the rotational symmetry and only the coefficients with $m = 0$ will survive

$$b_{\ell 0} = \left[\frac{2\ell+1}{4\pi}\right]^{\frac{1}{2}} \exp\left[-\frac{\ell(\ell+1)\sigma^2}{2}\right]. \tag{3.8}$$

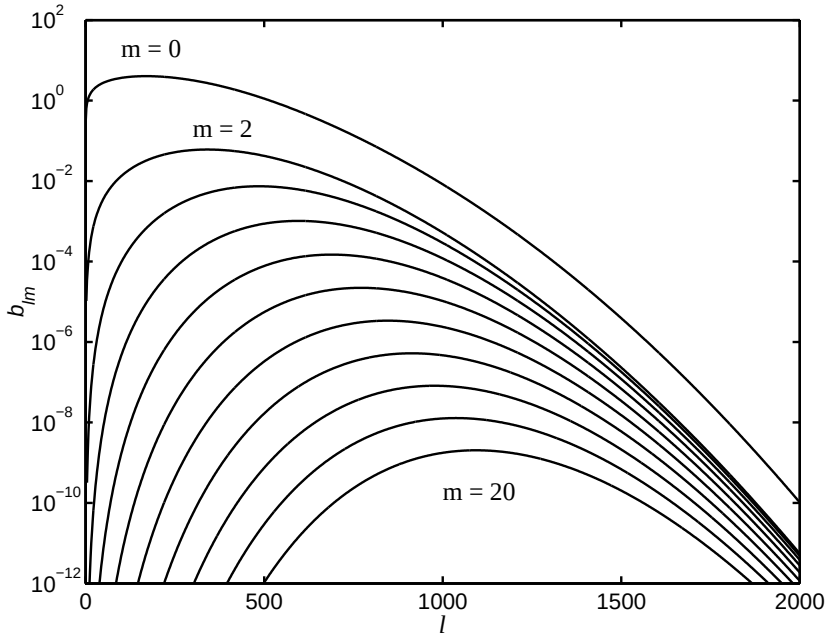

Figure 3.9: The coefficients ($b_{\ell m}$) of the spherical harmonic expansion of an elliptic Gaussian beam pointing to the north pole ($\mathbf{n}_0 = \hat{\mathbf{z}}$). Curves are for different values of m ($m = 0, 2, 4, \ldots 20$). The geometric mean of the FWHMs of the major and minor axes is 33 arcmin and the ellipticity (ϵ) is 1.4. These are characteristic values for an LFI 30 GHz detector beam, which has the largest ellipticity of the PLANCK beams.

In the CMB experiment the telescope beam sweeps the sky due to the spinning of the spacecraft. The low pass filtering of the bolometer response (in HFI) and the sampling integration (in HFI and in LFI) elongate the effective beam in the direction of the sweep. Further treatment of this phenomenon is beyond the scope of this study.

3.1.4 Detected Signal

Consider an observer sitting with a radiometer in the origin of a three dimensional coordinate system. The CMB photons hitting the telescope of the radiometer from the direction $\mathbf{n} = (\theta, \varphi)$ obey the Planck radiation law with a (thermodynamic) temperature $T(\mathbf{n}) = T_0 + \delta T(\mathbf{n})$ ([24]). $T_0 = 2.725$ K is the mean temperature (monopole) of the CMB photons today ([4]). $\delta T(\mathbf{n})$ is the CMB anisotropy signal. The *brightness* $B_\nu(\mathbf{n}, \nu)$ is the power of the flux of the photons coming from the direction $\mathbf{n}$ and measured in a unit frequency band, in a unit solid angle and in a unit collecting area ([41]). The frequency of the photons is ν. The unit of the brightness is Jy sr^{-1}, where 1 Jy = 10^{-26} Wm^{-2}Hz^{-1}. The term "specific intensity"

is sometimes used instead of "brightness" ([24]), but the term "brightness" is used in this study.

The brightness of the CMB photons (including two polarisation degrees of freedom) is ([41])

$$B_\nu(\mathbf{n},\nu) = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{k_\mathrm{B}T(\mathbf{n})}}-1}, \tag{3.9}$$

where c is the speed of light in vacuum, k_B is the Boltzmann constant and h is the Planck constant. In the case $h\nu \ll k_\mathrm{B}T(\mathbf{n})$ the brightness can be approximated with the Rayleigh-Jeans brightness

$$B_\nu(\mathbf{n},\nu) = \frac{2\nu^2 k_\mathrm{B}T(\mathbf{n})}{c^2}. \tag{3.10}$$

Equating this with Eq. (3.9) at all frequencies introduces a new temperature quantity $T_\mathrm{A}(\mathbf{n})$

$$\frac{2\nu^2 k_\mathrm{B}T_\mathrm{A}(\mathbf{n})}{c^2} = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{k_\mathrm{B}T(\mathbf{n})}}-1}. \tag{3.11}$$

$T_\mathrm{A}(\mathbf{n})$ is called *antenna temperature*. The right side of Eq. (3.11) can be expanded in a Taylor series around $T = T_0$. Because $|\delta T(\mathbf{n})| \ll T_0$ we can drop all except the first two terms of the series. This leads for the antenna temperature

$$T_\mathrm{A}(\mathbf{n}) = T_\mathrm{A}^0 + \delta T_\mathrm{A}(\mathbf{n}) = \frac{x}{e^x-1}T_0 + \left(\frac{x}{e^{x/2}-e^{-x/2}}\right)^2 \delta T(\mathbf{n}), \tag{3.12}$$

where $x = h\nu/(k_\mathrm{B}T_0)$. For CMB photons $x = \nu[\mathrm{GHz}]/56.77985$. The conversion factors from the thermodynamic temperatures to the antenna temperatures at the center frequencies of the Planck detectors are given in Table 3.1.

Note that the above treatments apply to photons obeying the Planck radiation law (like CMB photons). The antenna temperature of any radiation source can be defined by equating

$$\frac{2\nu^2 k_\mathrm{B}T_\mathrm{A}}{(1+p)c^2}$$

to its measured brightness. The quantity p depends on the type of detector used in measuring the radiation: p is zero for the polarisation insensitive detectors and one for the polarisation sensitive detectors.

For CMB photons the measured brightness is (cf. Eq. (3.9))

$$B_\nu(\mathbf{n},\nu) = \frac{2\nu^2 k_\mathrm{B}T_\mathrm{A}(\mathbf{n})}{(1+p)c^2} = \frac{2h\nu^3}{(1+p)c^2}\frac{1}{e^{\frac{h\nu}{k_\mathrm{B}T(\mathbf{n})}}-1}. \tag{3.13}$$

For the polarisation sensitive detectors the magnitude of the observed CMB anisotropy ($\delta T(\mathbf{n})$) will depend on the orientation of the polarisation sensitive direction of the detector[1].

Table 3.1: The conversion factors from the thermodynamic temperatures to the antenna temperatures at the center frequencies of the PLANCK detectors. The factors are introduced in Eq. (3.12), and their values given here apply to the CMB photons only ($T_0 = 2.725$ K).

Frequency[GHz]	x	$T_{\rm A}^0 = \frac{xT_0}{e^x-1}$ [K]	$\delta T_{\rm A}/\delta T_{\rm CMB} = \left(\frac{x}{e^{x/2}-e^{-x/2}}\right)^2$
30	0.53	2.068	0.977058
44	0.77	1.804	0.951425
70	1.23	1.382	0.882418
100	1.76	0.996	0.777158
143	2.52	0.601	0.604625
217	3.82	0.233	0.334178
353	6.22	3.387×10^{-2}	7.74265×10^{-2}
545	9.60	1.774×10^{-3}	6.25027×10^{-3}
857	15.09	1.146×10^{-5}	6.34747×10^{-5}

The power of the CMB anisotropy radiation received by a detector is ([41])

$$\delta P(\mathbf{n}_0,\psi) = \int_{4\pi} d\Omega_{\mathbf{n}} \int_0^\infty d\nu \eta F(\nu) \delta B_\nu(\mathbf{n},\nu) A_{\rm e}(\mathbf{n},\mathbf{n}_0,\psi). \tag{3.14}$$

The pointing of the center of the telescope beam and its orientation are $\mathbf{n}_0$ and ψ (see Sect. 3.1.3). $F(\nu)$ is the frequency response of the detector. The detector bandwidth ($\Delta\nu$) is given by

$$\Delta\nu = \int_0^\infty d\nu F(\nu). \tag{3.15}$$

The nominal bandwidths of the PLANCK detectors are quoted in Table 3.2. The brightness of the observed CMB anisotropy radiation is

$$\delta B_\nu(\mathbf{n},\nu) = \frac{2\nu^2 k_{\rm B} \delta T_{\rm A}(\mathbf{n})}{(1+p)c^2}. \tag{3.16}$$

The optical efficiency is η (the fraction of the incoming CMB anisotropy photons that becomes detected).

[1]For polarisation sensitive detectors $\delta T(\mathbf{n}) = I(\mathbf{n}) + Q(\mathbf{n})\cos(2\phi) + U(\mathbf{n})\sin(2\phi)$, where I, Q and U are the Stokes parameters of the CMB anisotropy ([42]) and ϕ is the angle between the polarisation sensitive direction and the x-axis of the coordinate system (located on the celestial sphere), where the Stokes parameters are given. For polarisation insensitive detectors $\delta T(\mathbf{n}) = I(\mathbf{n})$.

$A_{\rm e}(\mathbf{n}, \mathbf{n}_0, \psi)$ is the collecting area (effective aperture) of the telescope beam for the photons hitting the telescope from direction $\mathbf{n}$. The pointing and the orientation of the beam are $\mathbf{n}_0$ and ψ. The collecting area is given by ([41])

$$A_{\rm e}(\mathbf{n}, \mathbf{n}_0, \psi) = n\lambda^2 B(\mathbf{n}, \mathbf{n}_0, \psi), \tag{3.17}$$

where $B(\mathbf{n}, \mathbf{n}_0, \psi)$ is the response of the telescope (Sect. 3.1.3) and λ is the wavelength of the photons ($\lambda = c/\nu$). The number of electromagnetic wave guide modes excited by the CMB photons in the detector horn antenna is n. For the PLANCK detectors up to (and including) 353 GHz $n = 1$, $n = 4$ for 545 GHz detectors and $n = 8$ for 857 GHz detectors.

Inserting the above definitions in Eq. (3.14) the received power (detected signal) is

$$\delta P(\mathbf{n}_0, \psi) = \frac{2k_{\rm B} n \eta \Delta\nu}{1+p} \delta T_{\rm A}(\mathbf{n}_0, \psi), \tag{3.18}$$

where $\delta T_{\rm A}(\mathbf{n}_0, \psi)$ is the convolution of the CMB anisotropy field with the telescope response

$$\delta T_{\rm A}(\mathbf{n}_0, \psi) = \int_{4\pi} d\Omega_{\mathbf{n}} B(\mathbf{n}, \mathbf{n}_0, \psi) \delta T_{\rm A}(\mathbf{n}). \tag{3.19}$$

This beam smoothed CMB anisotropy signal ($\delta T_{\rm A}(\mathbf{n}_0, \psi)$) is the input signal in the map-making. With proper calibration it can be extracted simply (in principle) from the detected signal $\delta P(\mathbf{n}_0, \psi)$.

3.1.5 Radiometer Noise Model

A widely used model for the radiometer noise (noise in the detected signal) is a sum of white (non-correlated) and $1/f$ (correlated) noises ([70], [71]). The non-correlated part is the thermal noise of the electronics and the correlated part emanates mainly from the gain and noise temperature fluctuations of the amplifiers. The noise samples have zero mean, Gaussian distribution and their power spectral density (PSD) is

$$P(f) = \frac{\sigma^2}{f_{\rm s}} \left[1 + \left(\frac{f_{\rm k}}{f} \right)^{\alpha} \right] (f > f_{\rm min}). \tag{3.20}$$

By definition PSD ($P(f)$) is the Fourier transform of the autocorrelation of the noise samples. Frequency $f_{\rm s}$ is the sampling rate of the detected signal (values for the PLANCK detectors are given in Table 3.2), σ is the standard deviation (std) of the non-correlated part of the noise, $f_{\rm k}$ is the knee frequency indicating the frequency where the spectral densities of the non-correlated and the correlated noise are equal and α gives the spectral slope of the correlated noise. It is generally believed that the coherence of the noise processes in the amplifiers will not last for indefinitely long periods of time. This feature is modelled by the minimum frequency ($f_{\rm min}$) parameter. The inverse of the minimum frequency sets the time scale over which the noise is correlated. Below the minimum frequency ($f_{\rm min}$) the noise spectrum

becomes flat. The PSD of the noise model for parameter values representing an HFI 217 GHz detector is shown in Fig. 3.10.

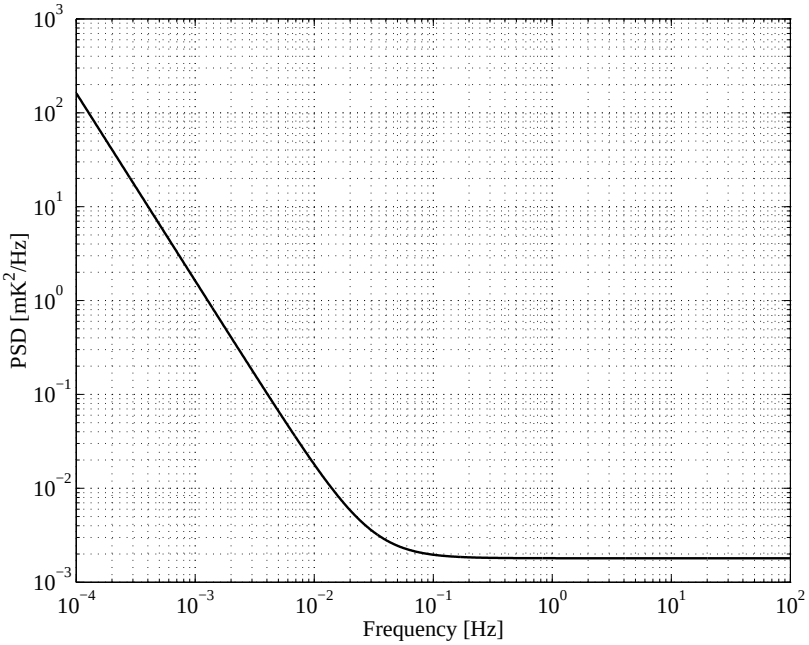

Figure 3.10: The PSD of the radiometer noise. The curve was determined from the model (Eq. (3.20)) with the parameter values representing an HFI 217 GHz detector: $\sigma = 0.6$ mK, $f_\mathrm{k} = 0.03$ Hz, $f_\mathrm{min} = 1.15 \times 10^{-5}$ Hz, $\alpha = 2.0$ and $f_\mathrm{s} = 200$ Hz (Table 3.2).

With an appropriate set of parameter values this model can describe the noise properties of the PLANCK HFI and LFI radiometers. For PLANCK the typical ranges of the parameters are: $f_\mathrm{k} = 0.03 \ldots 0.1$ Hz and $\alpha = 1.0 \ldots 2.0$. As far as the author knows, the actual minimum frequencies have not been measured from the PLANCK radiometers but it is generally believed that they will lie in the range $10^{-6} \ldots 10^{-4}$ Hz.

The noise model is flexible in a sense that the noise can be stationary or the model parameters can be varied over time (piecewise stationary or non-stationary noise).

For an LFI radiometer the standard deviation (std) of the non-correlated noise (in the antenna temperature domain) is obtained from ([32])

$$\sigma = \sqrt{\frac{2 f_\mathrm{s}}{\Delta\nu}} (T_\mathrm{sys} + T_\mathrm{A}^0), \tag{3.21}$$

where T_sys is the system noise temperature of the radiometer ([32]) and T_A^0 is the CMB monopole in the antenna temperature units (Eq. (3.12)). The goals of the

system noise temperatures are $7.5 \ldots 21.5$ K for different LFI radiometer channels ([31]).

For an HFI bolometer the std of the non-correlated noise (in the δP domain, see Eq. (3.18)) is ([34])

$$\sigma_{\delta P} = \mathrm{NEP}\sqrt{\mathrm{BW_{PD}}}, \tag{3.22}$$

where the noise equivalent power (NEP) is a detected power that produces a unit signal-to-noise ratio (S/N = 1) in the post-detection circuitry. NEP is determined experimentally for each bolometer. Its unit is $\mathrm{WHz}^{-1/2}$. The post-detection band $0 \ldots f_\mathrm{s}/2$ (Nyquist critical frequency) is assumed leading to a post-detection bandwidth $\mathrm{BW_{PD}} = f_\mathrm{s}/2$. Dividing both sides of Eq. (3.22) with the calibration factor (given in Eq. (3.18)) one obtains σ in the antenna temperature domain

$$\sigma = (1+p)\frac{\mathrm{NEP}\sqrt{f_\mathrm{s}}}{2\sqrt{2}k_\mathrm{B}n\eta\Delta\nu}. \tag{3.23}$$

Computer software (like Level S developed in the Planck collaboration, [43]) exists that generate a stream of noise samples according to the PSD of Eq. (3.20). As an example, a sequence of noise samples representing an HFI 217 GHz detector is shown in Fig 3.11. The corresponding PSD (from the model in Eq. (3.20)) is shown in Fig 3.10. The non-correlated (rapid fluctuations) and the correlated (slow offset drift) noise components can be clearly recognized in the time domain plot.

3.1.6 Radiometer Performance Goals

The performance goals of the Planck payload have been published in [8] and [9]. Table 3.2 presents a summary of the performance goals extracted from these sources.

3.2 Scanning Strategy

The Planck satellite will be placed in an orbit around the 2nd Lagrangian point (L2) of the Sun-Earth system ([9], [44], [45]). A satellite placed in L2 will follow the Earth when the Earth is orbiting the Sun (Fig. 3.12). To maintain the desired pointing of the satellite spin axis (nearly anti-Sun pointing, see Fig. 3.12), the spin axis is repointed at regular intervals. The pointing of the beam center is nearly perpendicular (85 deg) to the spin axis. While the satellite spins (nominal spin rate 1 rpm) the beam will draw nearly great circles in the sky. In 14 months the sky will be covered twice with observations.

The LFI and HFI detectors will average the intensity of the radiation with constant time intervals. The averages are the output samples of the detectors. The inverse of the time interval is the detector sampling rate. The sampling rates and the corresponding nominal angular spacings between the samples are given in Table 3.2.

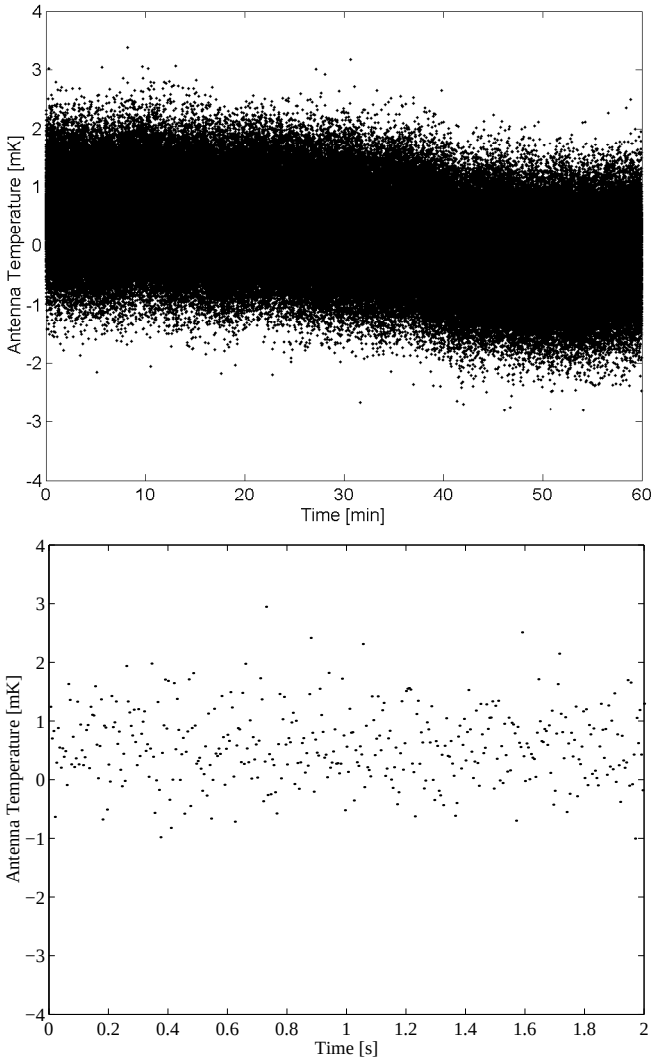

Figure 3.11: **Top:** A 60 minute sequence of simulated radiometer noise samples representing a single HFI 217 GHz detector. The stream was generated with the Level S software ([43]). The noise model parameter values were as in Fig. 3.10. The displayed sequence of noise contains 720 000 samples. **Bottom:** A blow-up plot of the first two seconds of the noise stream. This sequence contains 400 samples.

The pattern how the projection of the satellite focal plane sweeps in the sky is called scanning strategy. The scanning strategy has a significant impact on the scientific output of the mission (e.g. maps and angular power spectra). Ideally, the scanning strategy should fulfill the following requirements ([44]).

- Full sky coverage.
- Integration time spread homogeneously across the sky (leading to low map

Table 3.2: Goals of the PLANCK payload ([8], [9]).

	LFI			HFI					
Center frequency (GHz)	30	44	70	100	143	217	353	545	857
Detector bandwidth (GHz)	6	8.8	14	33	47.2	71.6	116.5	179.9	282.8
Polarised detectors	4	6	12	8	8	8	8		
Unpolarized detectors					4	4	4	4	4
Angular resolution (FWHM arcmin)	33	24	14	9.5	7.1	5	5	5	5
Sampling rate (Hz)	32.5	45	76.8	200	200	200	200	200	200
Angular spacing between samples (arcmin) (a)	11.1	8	4.7	1.8	1.8	1.8	1.8	1.8	1.8
$\Delta T/T$ per pixel ($\mu K/K$) (b)	2.0	2.7	4.7	2.5	2.2	4.8	14.7	147	6700

(a) 1 rpm spin rate
(b) 14 months of integration, 1σ, square pixels with size = FWHM, temperature map, equal number of hits in every pixel

noise).

- Redundant observations at many time scales (stable map solution, control of systematic errors of the instrument).
- Small angular deviation of the spin axis from the anti-Sun pointing (leading to minimum straylight from the Sun).
- Fixed deviation of the spin axis from the anti-Sun pointing (leading to minimum thermal modulation).
- Simply connected regions of sky with large integration times (leading to better detection of the systematics of the instrumentation).
- No sharp gradients in the integration time (leading to low systematics in the data processing).
- In every point of the sky the scanning directions evenly spread around the clock (leading to the azimuth symmetry of the effective beam pattern in the sky).

The satellite design sets several limitations on the feasible scanning strategies. A selected scanning strategy can not usually fulfill all the requirements. The options considered for the PLANCK scanning strategies are the *nominal* and the *slow precession* strategy ([44]). In the nominal scanning the satellite spin axis stays always

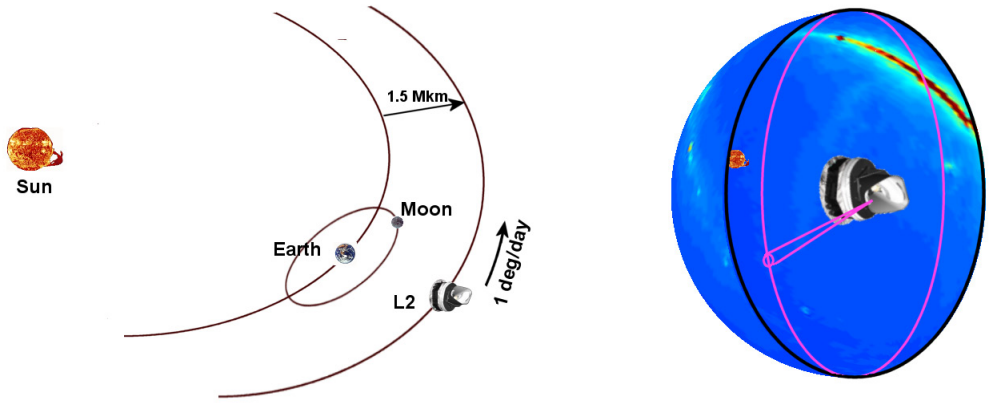

Figure 3.12: PLANCK satellite orbiting the 2nd Lagrangian point of the Sun-Earth system. The figure is from [8].

in the ecliptic plane. In order to maintain the anti-Sun pointing the spin axis is repointed once per hour (2.46 arcmin azimuth repointing). Because the angle between the pointing of the beam center and the spin axis deviates from 90 deg this scanning strategy will leave small unobserved areas in the ecliptic poles. As an example, the number of hits in the 7×7 arcmin^2 areas (pixels) for one 217 GHz HFI detector are shown in the top panel of Fig. 3.13. The areas with long integration times are concentrated near the poles. In the ecliptic equator the integration times are the lowest and the spread of the scan directions is small (most of the scans in the pole-to-pole directions). The spread in the pole regions is more favorable.

In the slow precession strategy the satellite spin axis is periodically deflected above and below the ecliptic plane. While the satellite advances in the ecliptic plane, the tip of the spin axis vector draws a sinusoidal curve perpendicular to the ecliptic plane (sinusoidal precession) or a circular path around the anti-solar axis (cycloidal precession). A typical period of the precession could be couple of months and a typical amplitude of the precession around $5 \ldots 10$ deg. Note that at least 7.5 deg amplitude is needed for all PLANCK detectors to scan the whole sky. In the slow precession strategy the pointing of the satellite spin axis is fixed between the repointings and it is repointed once per hour but the repointing will accur both in azimuth and in elevation. The spread of the integration time (in 7×7 arcmin^2 pixels) for a single 217 GHz HFI detector and for the cycloidal scanning is shown in Fig. 3.13. In this case the sky is fully covered with observations and relatively large areas with high integration times are located at the ecliptic poles. Due to the spin axis precession the spread of the scanning directions in the ecliptic equator is more favorable than in the nominal scanning strategy.

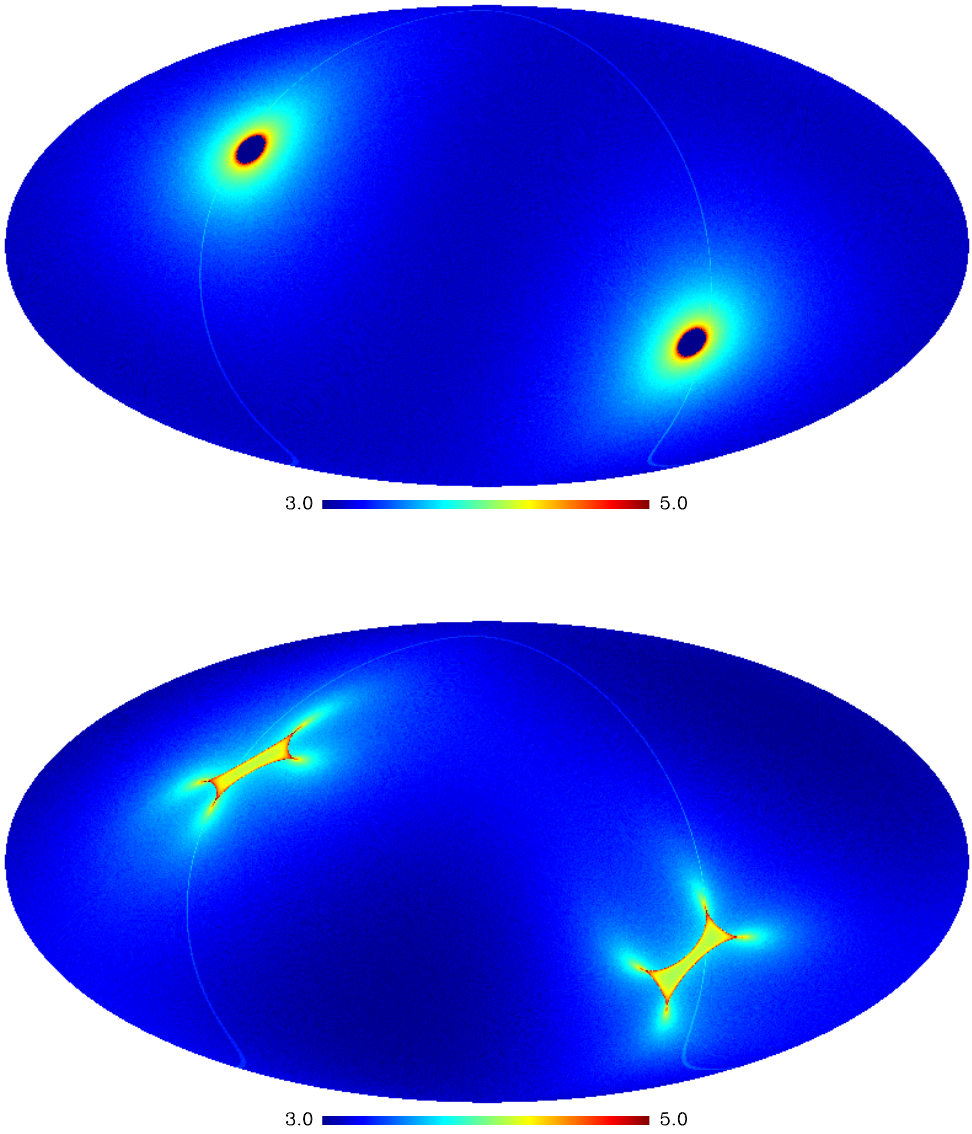

Figure 3.13: Number of hits in 7×7 arcmin^2 areas (pixels) for two scanning strategies. **Top:** Nominal scanning strategy. **Bottom:** Cycloidal precession scanning strategy. For the descriptions of the strategies see the text. The maps are in the galactic coordinate system and they were produced for a single 217 GHz HFI detector observing the sky for 12 months. The period of the cycloidal precession was 6 months and its amplitude was 7 deg. The diameter of the unobserved areas (ecliptic poles) of the nominal scanning strategy are about 8.5 deg for this detector (the size of the unobserved area depends on the location of the detector horn antenna in the focal plane). The scale is $\log_{10}(n_{hit})$, where n_{hit} is the number of hits in a pixel. The projection is Mollweide ([46]). The faint circles in the maps result from the overlap of the end of the mission with the beginning.

3.3 Foregrounds

The microwave signal received from the sky is a sum of emissions from several sources. The CMB photons that we see today emanated from the recombination. A number of sources of microwave photons and noncosmological sources of CMB anisotropy exist between the last scattering surface and today's observer. These signals are treated here under the title "Foregrounds". In the CMB data analysis these signals are contaminants that typically degrade the quality of the CMB detections.

3.3.1 Satellite Motion Induced Signal

The motion of an observer relative to the (nearly) isotropic CMB radiation field produces a Doppler shift temperature field in the sky ([47])

$$T(\mathbf{n}) = T_0 \frac{(1-\beta^2)^{1/2}}{1-\beta\hat{\mathbf{v}}\cdot\mathbf{n}} = T_0[1+\beta\cos(\alpha)+\beta^2\cos(2\alpha)+O(\beta^3)], \tag{3.24}$$

where $\beta = \mathrm{v/c}$, $\mathbf{v} = \mathrm{v}\hat{\mathbf{v}}$ is the velocity of the observer relative to the CMB radiation field, $\hat{\mathbf{v}}$ is the unit vector in the direction of the velocity and c is the speed of light in vacuum. $\mathbf{n}$ is the direction of the observation and $\cos(\alpha) = \hat{\mathbf{v}}\cdot\mathbf{n}$. The mean temperature of the CMB photons is T_0 (Sect. 3.1.4).

The velocity of the observer (satellite) can be split in a sum of two parts: the velocity ($\mathbf{v}_1$) of the Solar system relative to the CMB radiation field (last scattering surface) and the velocity ($\mathbf{v}_2$) of the satellite relative to the Sun. Because $\beta \ll 1$, the sum of the temperature fields of these two velocity components is a good approximation of the total temperature field. The velocity component $\mathbf{v}_1$ produces a constant dipole type temperature field in the sky. The magnitude of the dipole ($T_0\beta$) is 3.346 mK in the thermodynamic temperature scale ([7]). As an example, the dipole field seen by one of the 217 GHz HFI detectors is shown in Fig. 3.14. A sequence of observations scanned from the map of Fig. 3.14 is shown in Fig. 3.15. The dipole signal is the sinusoidal fluctuation.

The satellite motion relative to the Sun ($\mathbf{v}_2$) leads to a modulation of the dipole signal. The amount of modulation is known and it can be used in the amplitude calibration of the observed signal. The magnitude of the modulation is of the order of 10% of the dipole signal.

3.3.2 Galactic Emissions

The main microwave emissions emanating from our own galaxy are the thermal radiation from the interstellar dust, synchrotron and free-free radiations ([24]). Synchrotron radiation is generated when the trajectories of electrons are bent in the magnetic fields. The free-free radiation is created when free electrons decelerate in the electromagnetic fields of other free electrons.

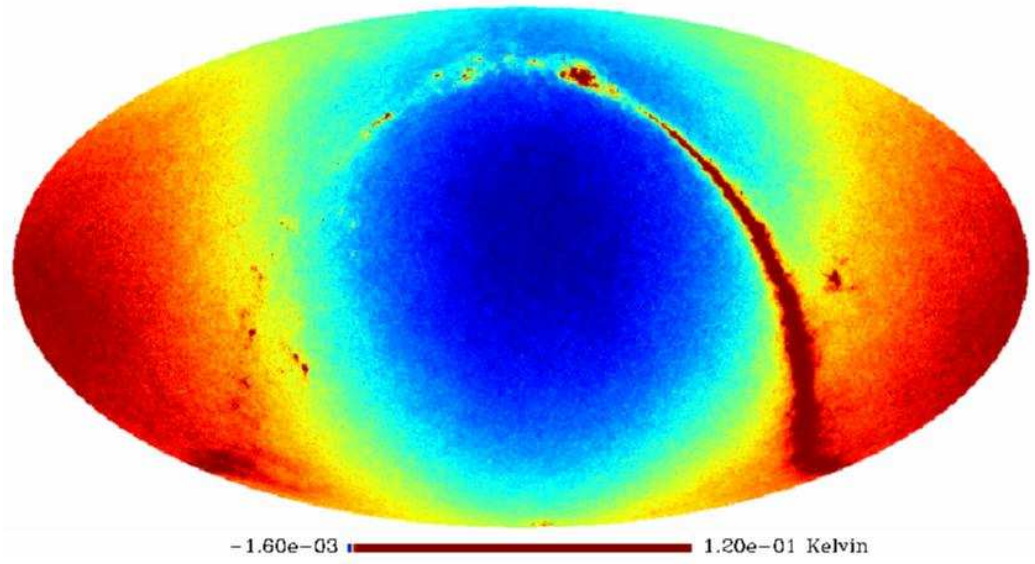

Figure 3.14: The CMB dipole and the galaxy seen by one of the HFI 217 GHz detectors. No radiometer noise is assumed. The map contains the CMB signal too, but it is buried under the galactic and the dipole signals. The temperatures are in the antenna temperature scale. The map is in the ecliptic coordinates. The data for this simulation was created with the Level S software ([43]).

The frequency dependence of the radiation intensity of these sources is significantly different than the frequency dependence of the CMB anisotropy signal. The radiation intensities of these sources are shown in Fig. 3.16. The best window for the CMB anisotropy detection is the frequency range $60 \ldots 100$ GHz. The lowest and the highest frequency channels of the PLANCK satellite will mainly be used to observe the galactic signals. The galactic emissions seen by one of the HFI 217 GHz detectors are demonstrated in Fig. 3.14. In the scanned time stream the galactic signal shows up as spikes occurring with regular intervals (Fig. 3.15). The galactic emissions have most of their power at the large angular scales.

Making observations at multiple frequencies allow us to separate the components and to effectively reduce the galactic contamination in the CMB anisotropy maps ([24]). The details of the component separation methods are beyond the scope of this study.

3.3.3 Extragalactic Emissions

Point sources and the cosmic infrared background (CIB) are important extragalactic microwave sources ([24], [49]). Point sources are the galaxies excluding our own. CIB

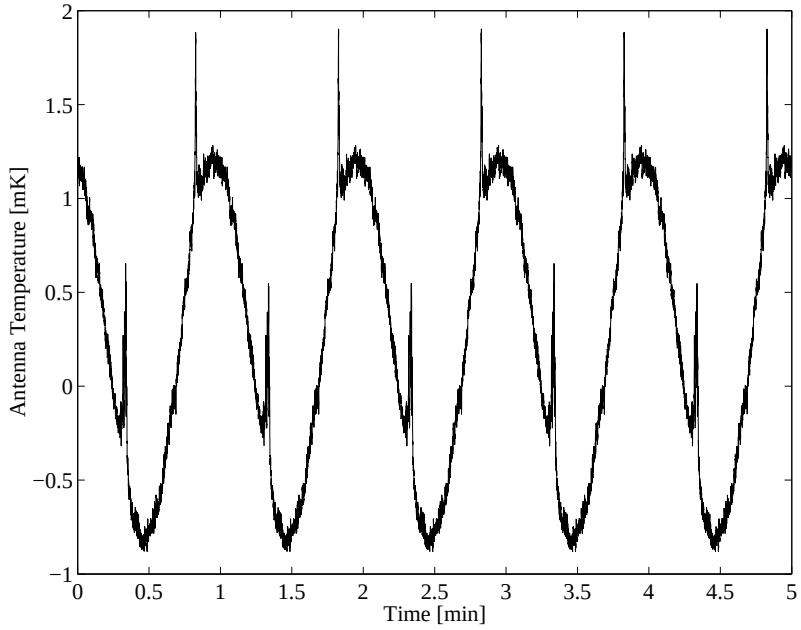

Figure 3.15: A sequence of observations scanned from the map of Fig. 3.14. The sinusoidal fluctuation is the satellite motion induced signal (dipole signal), the rapid fluctuation on top of the dipole signal is the CMB anisotropy signal and the spikes are from the scans across the galaxy. The observations in the plot contain no instrument noise.

radiation is produced by the stars in the galaxies (excluding the stars of our own galaxy) and partly absorbed/re-emitted by the interstellar dust.

At large angular scales the extragalactic contaminations are insignificant compared to the contaminations from the galactic sources. Point sources become more significant at smaller angular scales ([47]).

3.4 Time-ordered Data

The intensity recordings of the radiometers are collected in a time-ordered data (TOD) vector where the observations of the entire mission are organized in a chronological order. One such vector is produced for each detector.

A TOD vector $\mathbf{y}$ can be represented as

$$\mathbf{y} = \mathbf{s} + \mathbf{n}, \tag{3.25}$$

where the vector $\mathbf{s}$ is the signal received from the sky and the vector $\mathbf{n}$ is the instrument noise, which is a sum of the correlated $(1/f)$ and non-correlated (white) noise. The signal $\mathbf{s}$ is a sum of the CMB anisotropy and foreground emissions (Sect. 3.3). The TOD samples are usually given in the antenna temperature units.

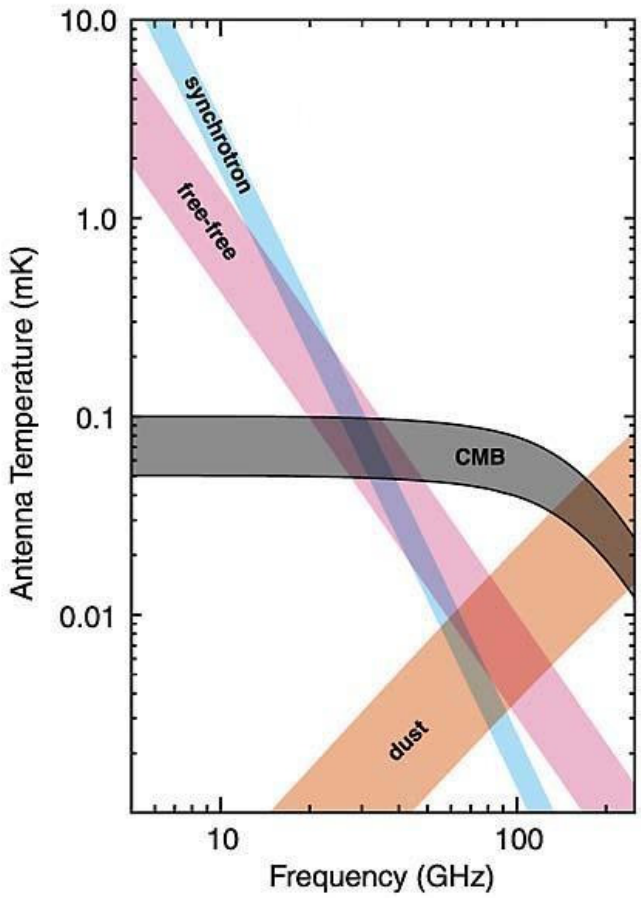

Figure 3.16: Radiation intensities of the CMB anisotropy and some galactic emissions. The figure is adapted from [48].

For the data processing purposes three angles need to be specified for each TOD sample. The pointing of the center of the beam is described by the spherical coordinates (θ_0, φ_0) of the celestial sphere. The third angle describes the orientation of the beam. It is given by a rotation angle ψ around $\mathbf{n}_0 = (\theta_0, \varphi_0)$. These angles were described in more detail in Sect. 3.1.3.

Systematic errors are contaminations in the TOD that decrease our capability to detect the wanted signal $\mathbf{s}$. Systematic errors have both astrophysical and instrumental origins. The correlated part of the instrumental noise is an example of a systematic error. Other types of systematic errors occurring in real experiments are ([31])

- Straylight from the Sun, solar system planets and galaxy entering the detectors through the telescope sidelobes.

- Incomplete knowledge of the actual beam pointing.
- Incomplete knowledge of the true main beam radiation pattern.
- Error in the amplitude calibration of the TOD.
- The ambient temperature of the radiometers may vary in time due to variations of the satellite orientation relative to the Sun and due to fluctuations of the satellite cooling systems. Fluctuations in the ambient temperature introduce extra noise in the TOD.
- Cosmic ray particles hitting the detector lead to the glitches in the TOD.

The detection and removal of the systematic errors (except the correlated noise) are beyond the scope of this study.

Examples of the signal and noise TODs ($\mathbf{s}$ and $\mathbf{n}$) for a 217 GHz PLANCK detector were shown in Figs. 3.11 and 3.15. The observed TOD ($\mathbf{y} = \mathbf{s} + \mathbf{n}$) is shown in Fig. 3.17. The noise and the satellite motion induced signal (dipole) are the dominating components.

The amount of TOD and the beam pointing and orientation data that PLANCK produces during its nominal 14 months mission is significant. The size of the data is summarized in Table 3.3.

Table 3.3: Amount of TOD and beam pointing and orientation data produced by the PLANCK detectors during the 14 months mission. The numbers assume that each TOD sample is given with 4 B (bytes) and each angle by 8 B (1 B = 8 bits and 1 GB = 10^9 B). The total amount of data is 12 040 GB (TOD + pointing and orientation, all detectors combined).

	LFI			HFI					
Center frequency (GHz)	30	44	70	100	143	217	353	545	857
Detectors (pol+non-pol)	4	6	12	8	12	12	12	4	4
Observations (10^9)(a)	5	10	34	59	88	88	88	29	29
TOD data (GB)	20	40	136	236	352	352	352	116	116
Pointing & orientation (GB)	120	240	816	1416	2112	2112	2112	696	696

(a) Sampling rates given in Table 3.2, 426 days mission time, all detectors for a frequency channel.

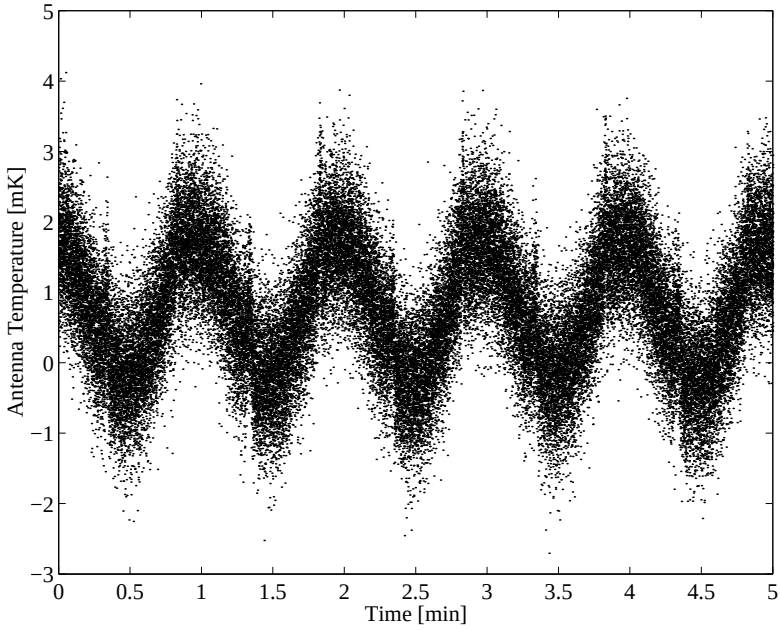

Figure 3.17: A sequence of the observed TOD. The plot was obtained by summing the curves from Figs. 3.11 and 3.15.

3.5 Data Processing

3.5.1 Data Processing Pipeline

A typical data processing pipeline of a satellite CMB experiment is shown in Fig. 3.18. The telemetry data obtained from the satellite contains both the satellite housekeeping data and the science data. They are split in step A and the instrument health checks are carried out using both data streams. The raw TOD (science data of the telemetry signal) is passed on to step B, where the TOD is calibrated and cleaned (as much as possible) from the systematics. The calibrated and cleaned TOD is the input to the map-making (step C). A single map for each frequency channel (frequency map) is produced in this step. In step D all frequency maps are jointly processed to perform the component separation. As a result of this step maps for the CMB anisotropy and foreground emissions will be made. Component separation is not perfect and the CMB maps will contain some foreground residues. The angular power spectrum estimates and their associated covariance matrices are produced in step E. Finally they are used in step F to produce the estimates of the cosmological parameters (e.g. $\Omega, A, n, \tau_{\mathrm{r}}, h, \Omega_{\Lambda}, \Omega_{\mathrm{b}}, \Omega_{\mathrm{m}}$, see Table 2.1 of Sect. 2.8).

The emphasis of this thesis lies in the methods required in steps C (map-making) and E (angular power spectrum estimation).

In the Planck terminology the data processing pipeline is split in Level 1, Level 2 and Level 3. Level 1 is equal to step A, Level 2 contains steps B and C and Level

3 contains the functionalities of step D.

The science output of the PLANCK mission is going to be ([9])

- Frequency maps for 9 frequency channels
- Component maps (CMB, galactic foregrounds)
- Catalogs of extragalactic sources
- Calibrated and cleaned TODs
- Angular power spectrum estimates (temperature and polarisation) and their covariance matrices
- Estimates of the cosmological parameters

3.5.2 Sky Pixelization

For the CMB data analysis (e.g. map-making and angular power spectrum estimation) the sky is divided in pixels. The map is a set of temperature values, where a value represents the temperature in a pixel. The Hierarchical Equal Area isoLatitude Pixelization (HEALPix, [50], [51], [52]) is applied throughout this study. In HEALPix the resolution of the map is given by the parameter $N_{\rm side}$. Its values are integer powers of 2. The number of pixels covering the entire sky is $N_{\rm pix} = 12 \times N_{\rm side}^2$. The angular sizes of the pixels (in solid angle) are equal: $\Omega_{\rm pix} = 4\pi/N_{\rm pix}$. The centers of the pixels are organized in iso-latitude rings with an equal azimuthal spacing between the centers. This allows the use of Fast Fourier Techniques (FFT) in the spherical harmonic transformations. In terms of the speed of the calculation this gives clear benefits. HEALPix is the baseline pixelization scheme in the PLANCK project.

The parameter $N_{\rm side}$, the number of pixels covering the sky and the angular size of a pixel for a set of some often used HEALPix resolutions are shown in Table 3.4.

3.5.3 Level S

The data processing software needs to be developed during the pre-launch period of the CMB experiment. The sofware development requires realistic TOD streams including all relevant signals (CMB anisotropy and foregrounds signals convolved with the realistic telescope beams) and systematic errors (e.g. instrument noise, glitches, pointing errors and thermal modulations). Additionally, the TOD streams shall have the length representing the entire duration of the mission. Level S is a collection of software modules developed in the PLANCK collaboration to generate the TODs ([43]). TODs generated by Level S were used in some of the studies of this thesis (paper **III** [3]).

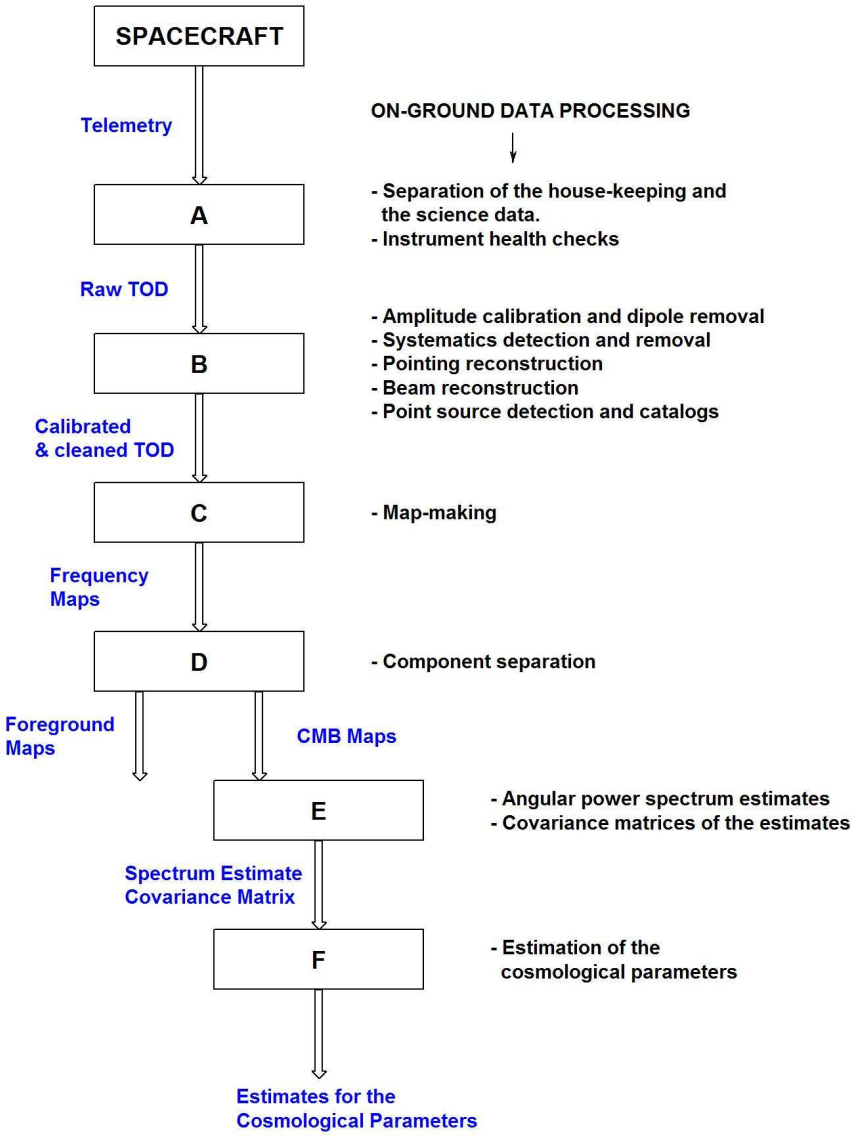

Figure 3.18: The on-ground data processing pipeline of a satellite CMB experiment.

The focal plane database file and sky templates are the inputs to Level S. The database file gives the locations of the horn antennas in the focal plane, describes the main beam radiation patterns in terms of elliptic Gaussian models and gives the parameter values of the instrument noise model. The noise model applied in Level S is identical to the model described in Sect. 3.1.5. Sky templates (maps and spherical harmonic expansion coefficients) of the CMB anisotropy, galactic and extragalactic emissions have been produced in the PLANCK collaboration. The main outputs of the Level S tools are the TODs and the beam pointings & orientations for a selected set of LFI or HFI detectors.

Table 3.4: The relation between $N_{\rm side}$, the number of pixels and the typical angular size of a pixel for some HEALPix resolutions. The shapes of the HEALPix pixels are different in different parts of the sky.

$N_{\rm side}$	$N_{\rm pix}$	$\Omega_{\rm pix}^{1/2}$ (arcmin)
64	49 152	54.97
128	196 608	27.48
256	786 432	13.74
512	3 145 728	6.87
1024	12 582 912	3.44
2048	50 331 648	1.72
4096	201 326 592	0.86
8192	805 306 368	0.43

Level S package has software tools for the following tasks

- Movement of the satellite in its orbit at L2. The pointings of the satellite axes (x,y,z) as a function of time. The z-axis is aligned with the satellite spin axis and the x- and y-axes are perpendicular to it.
- Creation of the spherical harmonic coefficients ($b_{\ell m}$) for the detector beams.
- Convolution of the sky template maps with the beam to produce the samples for the signal part of the TOD (vector $\mathbf{s}$ of Eq. (3.25)). The filtering with the frequency responses of the bolometer time constant and the sampling integration is included.
- Conversion of the pointings of the satellite axes to the pointings and the orientations of the detector main beams.
- Generation of the samples of the instrument noise (white and $1/f$ noise in vector $\mathbf{n}$ of Eq. (3.25)).
- Generation of the noise emanating from the thermal fluctuations of the satellite cooling systems.
- Synfast and anafast codes (originally from the HEALPix package) to perform forward and backward spherical harmonic transforms of a map.

Level S is capable of delivering both temperature-only and polarized TODs.

Chapter 4

Map-making

4.1 Map-making Problem

Suppose we have N_t samples of the observed temperature TOD in a column vector $\mathbf{y}$. The TOD is assumed to be a sum of signal and instrument noise

$$\mathbf{y} = \mathbf{s} + \mathbf{n}. \tag{4.1}$$

The signal vector $\mathbf{s}$ is the underlying true sky temperature field convolved with the telescope response while the center of the telescope beam follows the scan path. In the map-making we model the underlying true sky temperature field with a pixelized *sky map*. The sky map is represented by a column vector $\mathbf{m}$. The length of the map vector is N_pix, the number of pixels in the map. A pixel temperature of the sky map represents the mean of the true sky field across the pixel area.

We assume that the signal depends linearly on the sky map and write $\mathbf{Am}$ instead of $\mathbf{s}$. Matrix $\mathbf{A}$ is the observation matrix whose elements encode the telescope response and the scanning strategy. Each row of $\mathbf{A}$ contains a rotated map of the telescope response. The size of $\mathbf{A}$ is $N_\mathrm{t} \times N_\mathrm{pix}$. In the Planck experiment the typical values are $N_\mathrm{t} = 10^9 \ldots 10^{11}$ (see Table 3.3) and $N_\mathrm{pix} = 10^6 \ldots 10^8$.

It should be noted that vector $\mathbf{s}$ contains more information on the underlying true sky than vector $\mathbf{Am}$, because the sky signal structure at all angular scales contribute to $\mathbf{s}$ but only those scales that are larger than the pixel size of the sky map contribute to $\mathbf{Am}$. If the size of the sky map pixel is decreased towards zero, the vector $\mathbf{Am}$ will approach asymptotically the vector $\mathbf{s}$.

We suppose that the noise vector $\mathbf{n}$ is a collection of zero mean, Gaussian distributed random variables. The power spectral density (PSD) of the noise vector was introduced in Eq. (3.20). The noise samples are strongly correlated in time which, when coupled to the scanning strategy, will lead to stripes in the output maps that would distort the signal statistics. One of the map-making goals is to minimize the magnitude of these stripes.

Although a TOD is often associated with the observations of a single detector the above formalism can readily be generalized to multiple detectors. Suppose we

have $N_{\rm det}$ detectors with $\mathbf{y}_1, \mathbf{y}_2, \ldots \mathbf{y}_{N_{\rm det}}$ being the their TODs and $\mathbf{A}_1, \mathbf{A}_2, \ldots \mathbf{A}_{N_{\rm det}}$ being their observation matrices. We can now join the TODs and the observation matrices in a single vector and matrix

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{N_{\rm det}} \end{pmatrix} \tag{4.2}$$

and

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_{N_{\rm det}} \end{pmatrix}. \tag{4.3}$$

The formalism introduced in this chapter applies to these multidetector quantities and we will no longer make a distinction between single detector and multidetector observations.

The map-making problem can now be stated: Given the observed TOD (vector $\mathbf{y}$) we need to find the map estimate $\widetilde{\mathbf{m}}$ that minimizes the variance of the residual $\widetilde{\mathbf{m}} - \mathbf{m}$? The solution to this problem is called the *minimum variance map*. The residual $\widetilde{\mathbf{m}} - \mathbf{m}$ is also called *reconstruction error* in the literature ([54]).

4.2 Minimum Variance Map

Under the assumption of a Gaussian distribution the probability density function (PDF) of the noise vector $\mathbf{n}$ is

$$\mathrm{P}(\mathbf{n}) = (2\pi)^{-\mathrm{N_t}/2} \exp\left(-\frac{1}{2}(\mathbf{n}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{n} + \mathrm{Tr}[\ln \mathbf{N}])\right), \tag{4.4}$$

where $\mathbf{N} \equiv \langle \mathbf{n}\mathbf{n}^{\mathrm{T}} \rangle$ is the time domain noise covariance matrix. Superscript T indicates a transpose and $\langle \cdot \rangle$ is the ensemble mean. $\mathrm{Tr}[\cdot]$ is the trace of a matrix. Inserting $\mathbf{n} = \mathbf{y} - \mathbf{A}\mathbf{m}$ in Eq. (4.4) and assuming no prior knowledge on the underlying sky map $\mathbf{m}$, the probability that a certain sky will lead to our observed TOD is

$$\mathrm{P}(\mathbf{y}|\mathbf{m}) = (2\pi)^{-\mathrm{N_t}/2} \exp\left(-\frac{1}{2}((\mathbf{y} - \mathbf{A}\mathbf{m})^{\mathrm{T}}\mathbf{N}^{-1}(\mathbf{y} - \mathbf{A}\mathbf{m}) + \mathrm{Tr}[\ln \mathbf{N}])\right). \tag{4.5}$$

This is the likelihood function of the problem and the map that maximizes it is the minimum variance map $\widetilde{\mathbf{m}}$. The log-likelihood function is the logarithm of the inverted likelihood function. Setting the derivative of the log-likelihood function (with respect to $\mathbf{m}$) to zero leads to a linear matrix equation where the minimum variance map can be solved

$$\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{A}\widetilde{\mathbf{m}} = \mathbf{A}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{y}. \tag{4.6}$$

The covariance matrix of the residual map is

$$\langle(\widetilde{\mathbf{m}} - \mathbf{m})(\widetilde{\mathbf{m}} - \mathbf{m})^{\mathrm{T}}\rangle = (\mathbf{A}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{A})^{-1}. \tag{4.7}$$

It has been shown ([54]) that the minimum variance map is lossless in a sense that it contains the same cosmological information as $\mathbf{Am} + \mathbf{n}$ (under the assumptions of Gaussian distributed $\mathbf{n}$ and that we know the noise covariance $\mathbf{N}$).

In a hypothetical noiseless experiment the minimum variance map minimizes $|\mathbf{s} - \mathbf{A}\widetilde{\mathbf{m}}|$, the magnitude of the difference between the observed and the model data sets ([54]). We obtain for the minimum variance map of the noiseless experiment

$$\widetilde{\mathbf{m}} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{s}. \tag{4.8}$$

This map is also called the *binned noiseless map*. A pixel temperature of the binned noiseless map is a weighted sum of the observations that fall in that pixel. The observed (noiseless) TOD ($\mathbf{s}$) contains usually several observations of a pixel (spread across its area), whereas each pixel is represented by a single temperature value in the binned noiseless map. Thus the vector $\mathbf{s}$ contains information on the sub-pixel structure of the signal, whereas that information is lost in the binned noiseless map. If the binned noiseless map is scanned back to a TOD, this loss of the sub-pixel signal structure will lead to a difference between the scanned TOD ($\mathbf{A}\widetilde{\mathbf{m}}$) and the original TOD ($\mathbf{s}$). The difference ($\mathbf{A}\widetilde{\mathbf{m}} - \mathbf{s}$) is called *pixelization noise* ([55]).

In the Planck-like experiments the size of the observation matrix can become so large that it may set unrealistic requirements on the amount of computer memory required in the data processing. Additionally, the inversion of the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ (size $N_{\mathrm{pix}} \times N_{\mathrm{pix}}$) required in the data processing will become a major computational burden. For most of the work presented in this study, a simpler approach has been adopted, where each TOD sample represents the temperature of the pixel where the beam center is pointing. Under this assumption the observation matrix of a total power experiment (like Planck) contains simply ones and zeros. Each row of the matrix has a single non-zero element and the rest are zeros. We denote this simplified observation matrix by symbol $\mathbf{P}$ and call it *pointing matrix*.

When applying the pointing matrix (instead of the observation matrix) the memory requirements to store $\mathbf{P}$ are considerably relaxed. Additionally, the matrix $\mathbf{P}^{\mathrm{T}}\mathbf{P}$ is diagonal, the diagonal elements giving the number of hits in a pixel. However, there is a drawback too. The temperature of an output map pixel will not be the temperature of the underlying true sky in that point, but it will be the true sky convolved with a response that is the mean of the telescope beam responses in different orientations. The number of orientations equals the number of hits in that pixel and the orientations are given by the scan directions across the pixel. In general each pixel will have its own smoothing that depends on the telescope beam and the scanning strategy. If the symmetric beam approximation is valid each pixel will simply

be convolved with that beam and no scanning dependent smoothing exists. The map smoothing needs to be deconvolved in some later stage of the data processing (e.g. in angular power spectrum estimation).

The pointing matrix is considered for the most part of this study. Some discussion on possibilities to use the full observation matrix in the map-making are given in Sect. 4.6.

4.3 GLS Map-making Algorithms

To find the map that minimizes the log-likelihood function (logarithm of the inverse of Eq. (4.5)) is a general least squares (GLS) fitting problem ([53]). The methods that aim at finding the optimal minimum variance map ($\widetilde{\mathbf{m}}$) from the equation

$$\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{P}\widetilde{\mathbf{m}} = \mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{y}. \tag{4.9}$$

are called *GLS map-making algorithms*. A number of such algorithms have been proposed ([56] - [61]).

Usually GLS map-making assumes the noise to be stationary throughout the mission. It is further assumed that the elements of the covariance matrix (N_{ij}) vanish when $|i-j|$ is larger than some N_η and $N_\eta \ll N_\mathrm{t}$ ([60]). This means that the correlation is significant only across a number of samples that is a tiny fraction of the total length of the TOD. Thus the noise correlation matrix $\mathbf{N}$ can be approximated by a circulant matrix ([60]). Note that the matrix $\mathbf{N}^{-1}$ is approximately circulant as well. The multiplication $\mathbf{N}^{-1}\mathbf{y}$ can be carried out more easily in the frequency domain where $\mathbf{N}^{-1}$ is diagonal ([60]). In most of the proposed GLS map-making algorithms the output map is solved from Eq. (4.9) with an iterative preconditioned conjugate gradient (CG) method ([53]). The residual error after the k^{th} CG iteration is measured by the fractional difference

$$\varepsilon^{(k)} = \frac{\parallel \mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{P}\widetilde{\mathbf{m}}^{(k)} - \mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{y} \parallel}{\parallel \mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{y} \parallel}. \tag{4.10}$$

The norm $\parallel \cdot \parallel$ is the Euclidean norm. The iterations are repeated until the fractional difference has reached a low enough value. This limit is typically on the order of 10^{-6} ([60]).

Due to the circulant matrix approximation each row of the matrix $\mathbf{N}^{-1}$ contains the same element values in a different cyclic permutation. Only the elements of a row with $|i-j| \leq N_\xi$ ($N_\xi \ll N_\mathrm{t}$) have non-zero values. The rest of the elements are zero. The collection of the non-zero elements (of a row) is called the *noise filter*. The lag of an element is the difference of its indices $(i-j)$. The choice of the value N_ξ is a significant decision for the quality of the output maps and for the computation time of the algorithm ([60]).

Some examples of the noise filters derived from the radiometer noise PSD (Eq. (3.20)) are shown in Fig. 4.1. The noise filter amplitudes at the lags of some tens of

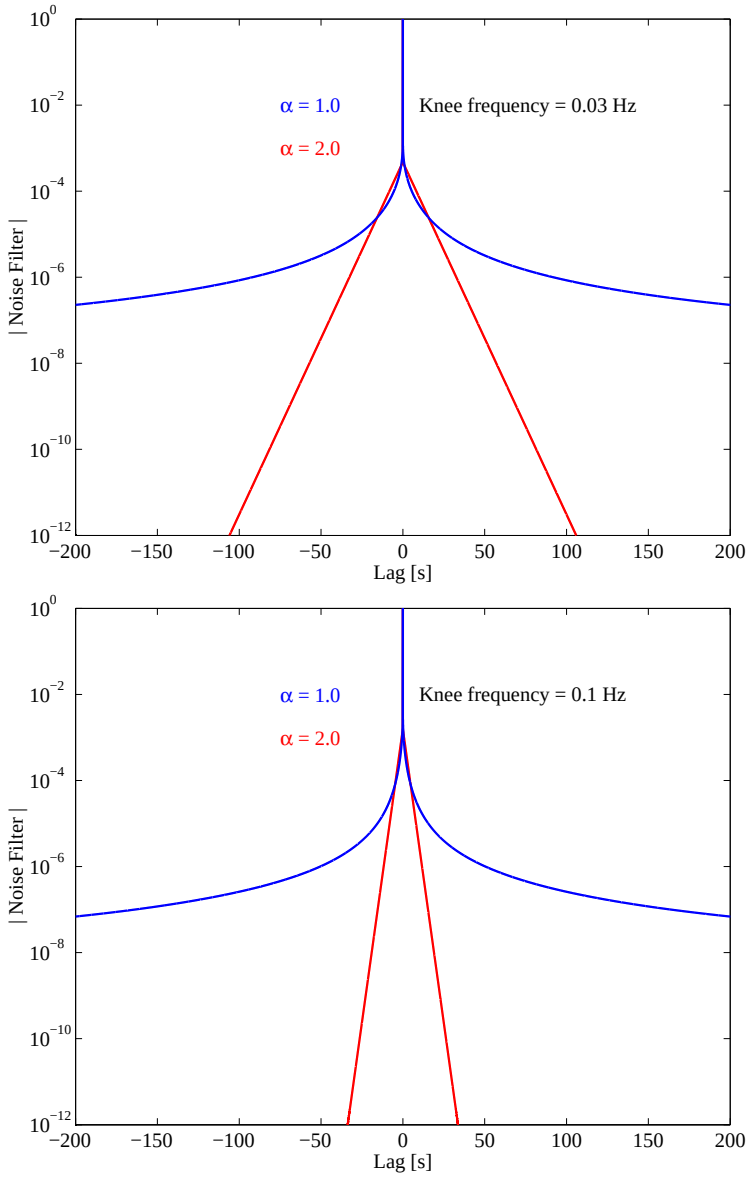

Figure 4.1: The absolute values of the noise filters determined from the PSD of the radiometer noise model (Eq. (3.20)). The lag of 200 s corresponds to 40 000 HFI samples and between 6 500 and 15 360 LFI samples. All noise filters were normalized to one at zero lag. The filter is positive at zero lag and negative at all other lags. The minimum frequency was 1.15×10^{-5} Hz. **Top:** Noise filters at knee frequency 0.03 Hz and at $1/f$ and $1/f^2$ slopes. **Bottom:** Same as the top panel but now the knee frequency is 0.1 Hz.

thousands of samples are suppressed by several orders of magnitudes relative to the amplitude at zero lag. This demonstrates the validity of the assumption $N_\xi \ll N_\mathrm{t}$.

For a given knee frequency, the amplitude gets more suppressed with increasing spectral slope (α). The amplitude (at a given lag) will be more suppressed at higher knee frequencies. The difference between the noise filters of knee frequencies 0.03 Hz and 0.1 Hz is significant at $\alpha = 2.0$ but relatively modest at $\alpha = 1.0$. It can be noted that as long as $f_{\rm min} \ll f_{\rm k}$, the noise filter is nearly independent from the minimum frequency. The diagonal elements of the frequency domain representation of the noise filter are samples (frequency bin values) taken from $1/P(f)$, where $P(f)$ is the noise PSD given in Eq. (3.20). For the instrument noise obeying the $1/f$ + white noise model, the noise filter is a high-pass filter.

In order for the GLS output map to fall close to the optimal minimum variance map, the algorithms require knowledge on the statistical properties of the instrument noise. In a real CMB experiment these properties must be estimated from the observed data. Both iterative ([55], [62], [63]) and non-iterative ([64], [65]) methods have been proposed for this task. Many schemes start by binning a map from the observed TOD, rescanning it back to a TOD and subtracting this TOD from the original observed TOD ($\widetilde{\mathbf{n}} = \mathbf{y} - \mathbf{P}(\mathbf{P}^{\rm T}\mathbf{P})^{-1}\mathbf{P}^{\rm T}\mathbf{y}$). The resulting time stream ($\widetilde{\mathbf{n}}$) is a crude approximation of the noise part of the observed TOD.

In the non-iterative methods either the noise power spectrum (binned in frequency) of the crude noise time stream is estimated or a PSD model (like Eq. (3.20)) is fitted to its power spectrum. The power spectra (estimated or fitted) of the crude noise time stream are then used as an estimate of the power spectrum of the noise of the observed TOD ([64]).

In the iterative methods the noise filter is first determined from the power spectrum of the crude noise time stream and then the map is solved from Eq. (4.9) using CG iterations. With the map as an estimate of the signal, a better noise time stream is obtained by spreading the map in to a TOD and subtracting it from the observed TOD ($\widehat{\mathbf{n}} = \mathbf{y} - \mathbf{P}\widetilde{\mathbf{m}}$). A better noise filter can now be determined, which leads to a better map and which can again be used to produce a better estimate of the noise stream and so on. If the parameters of the algorithm (like the frequency bins of the noise power spectrum) have been selected properly, this iteration will converge in a couple of cycles ([62], [63]).

The iterative noise estimation has been applied in BOOMERanG and MAXIMA experiments ([66], [67]). No studies of the impact of an inaccurate noise filter on the GLS maps have been published for the Planck-like experiments but it is shown in [64] that in the BOOMERanG type of experiments the iterative and non-iterative noise estimation can lead to GLS maps that are close to the maps where a perfect knowledge of the statistical properties of the noise has been available.

It is shown in [60] how GLS map-making can be applied when the instrument noise is only piece-wise stationary.

4.4 Destriping

4.4.1 Destriping Principle

The destriping method was introduced for the map-making in the early planning phases of the PLANCK project ([68]). The early software implementations ([69] - [73]) of the destriping made use of the fact that PLANCK is a spinning spacecraft whose detector beams draw nearly great circles in the sky. Each circle is observed 60 times before the spin axis is repointed (see Sect. 3.2). The time between two consecutive repointings is called a *repointing period*. It was assumed that the pointings of the samples of a circle will repeat themselves exactly in all 60 circles. In order to reduce the noise and the amount of data, the samples of those 60 circles were averaged. This process is called *co-adding*. We call one co-added circle a *ring*.

The early implementations of destriping ([69] - [73]) assumed that the $1/f$ part of the instrument noise in a ring can be approximated by an offset or a *uniform baseline*. This approximation was first introduced in [74], which stated that a uniform baseline is a good approximation if the knee frequency of the $1/f$ noise is smaller than the spin frequency of the satellite.

We call the length of a baseline an *elementary scanning period*. If the amount of data is reduced by co-addition, a ring is a natural choice for an elementary scanning period. Other methods proposed for the TOD reduction are pixel binning and phase binning. In the pixel binning the TOD samples corresponding to a repointing period are binned to map pixels. This is repeated leading to a specific set of pixels for every repointing period. Typically the number of pixels is smaller by a factor $20 \ldots 40$ than the number of samples in a repointing period. In the phase binning the scanning circle is split in angular bins and the TOD samples of a repointing period are collected in these bins and averaged ([75]). Co-addition is not applicable when the pointings of the samples of the consecutive circles do not repeat themselves. This may happen due to e.g. satellite spin axis nutation or due to fluctuations in the satellite spin rate. In these cases pixel or phase binning can be applied. In the pixel and phase binnings a natural choice for the elementary scanning period is the number of pixels or phase bins in a repointing period. If no TOD reduction is applied the elementary scanning period becomes a freely chosen parameter whose value does not have to be related to the scanning circles.

The uniform baseline approximation of the $1/f$ noise is demonstrated in Fig. 4.2. The top panel shows a simulated sequence of instrument noise representing one LFI 30 GHz detector. Note that the noise here was not co-added. The elementary scanning period was one minute and uniform baselines were determined by taking the means of the noise for those periods. The baselines determined this way are shown in the top panel as well. The noise after the subtraction of the baselines is shown in the bottom panel. As can be seen, uniform baselines provide a good approximation of the slow fluctuations ($1/f$ fluctuations) of the noise. This demonstrates that the instrument noise can be accurately modelled with a set of baselines added to the

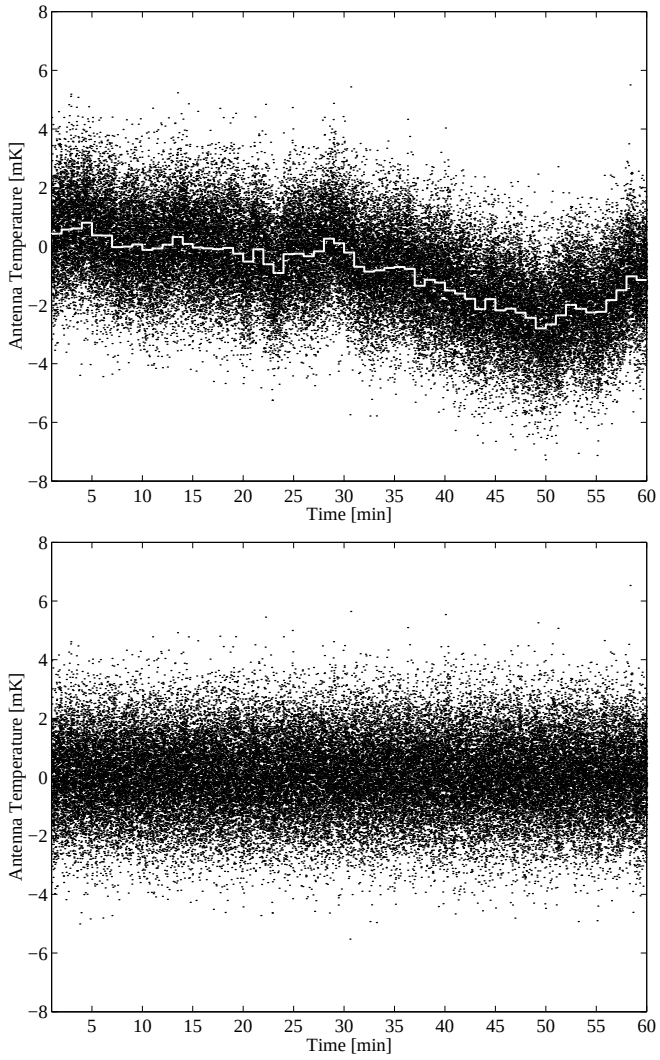

Figure 4.2: **Top:** A one hour sequence of simulated radiometer noise samples representing a single LFI 30 GHz detector. The noise is a sum of $1/f$ and white noise. The stream was generated with the Level S software ([43]). The uniform baselines (means of the noise for a one minute elementary scanning period) are shown too (white steppy line). **Bottom:** The noise after the uniform baselines have been subtracted. The parameter values of the noise PSD were (Eq. (3.20)): $\sigma = 1.3$ mK, $f_{\mathrm{k}} = 0.05$ Hz, $f_{\mathrm{min}} = 1.15 \times 10^{-5}$ Hz, $\alpha = 1.7$ and $f_{\mathrm{s}} = 32.5$ Hz. Note that the figure shows the noise for each scan circle (no co-addition).

non-correlated (white) noise.

For the rest of this section we will assume a co-added TOD and use a ring as an elementary scanning period. The discussion, however, is also valid for the TODs with no data reduction and with an arbitrary elementary scanning period.

In destriping one solves the estimates of the noise baselines, produces a cleaned TOD by subtracting the estimates (fitted baselines) from the original observed TOD and bins the cleaned TOD in the final output map. In a real experiment the observed TOD contains signal and noise. Because one cannot assume that the mean of the signal over a ring is zero, the baselines cannot be determined by simply taking means of the observed TOD (as we did for the noise-only TOD in Fig. 4.2). However, the signal should be the same in the crossing points of two different rings. Destriping adds a baseline to each data ring and adjusts the relative values of these baselines to minimize the data differences in the crossing points. A crossing point is identified when the samples of two different rings fall on the same sky pixel. The adjustment of the pixel size will influence the number of crossing points that we detect. It should be noted that, according to this definition, two closely spaced parallel rings may have a crossing point (or a row of them) without actually crossing each other if a pair of samples, with members of the pair from different rings, fall on the same pixel. Because only the data differences are considered, the same constant can be added to every baseline without affecting the differences. This degeneracy needs to be removed by e.g. requiring that the sum of the fitted baseline magnitudes is zero.

The general function to be minimized in destriping is ([70])

$$S = \sum_{p \in \text{sky}} \sum_{\text{pairs}} w(p, ik, jl)(y_{ik} - y_{jl} - a_i + a_j)^2, \tag{4.11}$$

where index p labels pixels, i, j rings and k, l samples on a given ring. Combined indeces ik and jl identify observations y_{ik} and y_{jl}. The magnitudes of the unknown uniform baselines are a_i and a_j for rings i and j, respectively. The second sum refers to all pairs (ik, jl) of measurements that fall on the pixel p and w is a weight function. Different weight functions were used in the early implementations of the destriping. Because pixel p contributes $n_p(n_p - 1)/2$ pairs (n_p is the number of hits in the pixel p), $w \propto 1/(n_p - 1)$ was suggested in [70]. Weight $w = \text{const.}$ was used in [71] and [72].

Minimizing S with respect to the unknown baseline magnitudes a_i and taking into account the constraint that the sum of a_i is zero, lead to a group of linear equations for the magnitudes

$$\mathbf{Ca} = \mathbf{b}. \tag{4.12}$$

Vector $\mathbf{a}$ contains the baseline magnitudes (a_i). The number of crossing points and their distribution in the sky determine the elements of the square matrix $\mathbf{C}$. Matrix $\mathbf{C}$ is symmetric and non-negative definite and its dimension equals the number of baselines. The elements of the vector $\mathbf{b}$ are weighted sums of the temperature differences of the crossing points $(y_{ik} - y_{jl})$. The length of $\mathbf{b}$ equals the number of baselines.

To solve the baseline magnitudes from Eq. (4.12) the matrix $\mathbf{C}$ needs to be inverted. For rings the number of uniform baselines is $\sim$9000 for a one year mission. This leads to a symmetric matrix $\mathbf{C}$ that can be directly inverted in today's supercomputers using e.g. Cholesky decomposition ([53]). However, if we have no TOD

reduction and e.g. one minute uniform baselines (leading to ∼500 000 baselines in a year), the direct inversion is too demanding a task for the current computers and iterative methods (e.g. conjugate gradient iteration, [53]) needs to be applied to solve the matrix equation. Solving the baseline magnitudes is a less complex problem than solving the map in the GLS method (cf. Eq. (4.9)), because the number of pixels in the output map is typically $10^6 \ldots 10^8$ and no multiplications of type $\mathbf{N}^{-1}\mathbf{y}$ are required in the destriping.

In a real experiment the samples of a crossing point do not necessarily measure the same point in the sky but they may have different pointings within the pixel. This means that the vector $\mathbf{b}$ will have a (small) non-zero contribution from the sub-pixel structure of the sky signal. This structure gives rise to the pixelization noise too (as described in Sect. 4.2). The sub-pixel structure leads to non-zero baseline magnitudes even in a noiseless experiment. The output map of the destriping would deviate from the binned noiseless map in this experiment. In a real experiment, however, the instrument noise is typically the main contributor in the baseline magnitudes and the contributions from the sky signal are mostly small ([73]).

Destriping is not an optimal map-making algorithm in a sense that it does not produce a minimum variance map. The difference is, however, small in many cases. The great virtues of destriping are that the algorithm is significantly simpler than the GLS map-making and it requires no prior information on the statistical properties of the instrument noise. We have assumed until now that the variance of the white noise (σ^2 in Eq. (3.20)) is the same for all TOD samples. Destriping with no prior knowledge on noise can also be applied to TODs with non-uniform white noise. The performance of destriping can, however, be enhanced if the levels of white noise are known and used in solving the baselines and the output map. This issue is discussed in more details in Sect. 4.4.2.

In addition to removing the $1/f$ noise from the TOD, destriping has been applied to clean the TOD from the drifts caused by other systematic effects, e.g. the thermal fluctuations of the satellite cooling systems ([76]).

4.4.2 Maximum Likelihood Approach to Destriping

The function S (Eq. (4.11)) that is minimized in the destriping was derived from practical considerations and it contained elements (e.g. the weight function w) that were given without more fundamental justifications ([70], [71]). In this thesis a systematic maximum likelihood approach to derive the function to be minimized was developed (paper **I** [1]). Additionally, destriping codes based on these principles were implemented (described in paper **I** [1]).

We continue to consider uniform baselines. As before, we assume that the signal part of the observed TOD is scanned from a sky map $\mathbf{m}$ (see Sect. 4.2). This discussion applies equally well to TODs with the data reduction (co-addition, pixel binning or phase binning) or to TODs with no reduction. Thus we will make no distinction between them. To model the correlated $(1/f)$ part of the instrument noise we define

a uniform baseline for each elementary scanning period. The observed TOD ($\mathbf{y}$) can be written

$$\mathbf{y} = \mathbf{P}\mathbf{m} + \mathbf{n}. \tag{4.13}$$

The underlying sky map is $\mathbf{m}$, $\mathbf{P}$ is the pointing matrix and the instrument noise is the vector $\mathbf{n}$. It is modelled as a sum of white noise and the baselines

$$\mathbf{n} = \mathbf{F}\mathbf{a} + \mathbf{n}'. \tag{4.14}$$

Vector $\mathbf{a}$ contains the unknown baseline amplitudes (one for each elementary scanning period) and the matrix $\mathbf{F}$ spreads the baselines to a TOD. Vector $\mathbf{n}'$ is the non-correlated (white) part of the instrument noise. The covariance matrix of the white noise is $\mathbf{C}_n \equiv \langle(\mathbf{n}')(\mathbf{n}')^{\mathrm{T}}\rangle$. It is diagonal but not necessarily uniform.

We assign no prior knowledge on the map and the baseline amplitudes. With these assumptions, we obtain the log-likelihood function

$$\chi^2 = (\mathbf{n}')^{\mathrm{T}}\mathbf{C}_n^{-1}(\mathbf{n}') = (\mathbf{y} - \mathbf{P}\mathbf{m} - \mathbf{F}\mathbf{a})^{\mathrm{T}}\mathbf{C}_n^{-1}(\mathbf{y} - \mathbf{P}\mathbf{m} - \mathbf{F}\mathbf{a}). \tag{4.15}$$

We need to minimize the log-likelihood function with respect to both of the unknown variables $\mathbf{m}$ and $\mathbf{a}$. After taking the derivative with respect to $\mathbf{m}$ and setting it zero, the map can be solved

$$\mathbf{m} = (\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}(\mathbf{y} - \mathbf{F}\mathbf{a}). \tag{4.16}$$

Inserting the map back to the log-likelihood function and minimizing it with respect to the baseline amplitudes, an equation for the vector $\mathbf{a}$ is obtained

$$\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}\mathbf{a} = \mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{y}, \tag{4.17}$$

where

$$\mathbf{Z} = \mathbf{I} - \mathbf{P}(\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}. \tag{4.18}$$

Matrix $\mathbf{I}$ is a unit matrix with dimension equal to the length of the TOD. Although not apparent, the matrix $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ is symmetric and non-negative definite (paper **I** [1]). When the matrix $\mathbf{Z}$ is acting on a TOD it subtracts from each sample the inverse (white) noise variance weighted average of the samples hitting the same pixel. The matrix $\mathbf{F}^{\mathrm{T}}\mathbf{F}$ is diagonal and it is normalized to $\mathbf{F}^{\mathrm{T}}\mathbf{F} = \mathrm{diag}(n_b)$, where n_b is the number of samples in an elementary scanning period. Eqs. (4.16) and (4.17) were the main theoretical results of paper **I** ([1]).

The baseline amplitudes cannot be solved from Eq. (4.17) by directly inverting the matrix $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$, because it is singular. It was shown in Sect. 4.4.1 that a constant can be added to every baseline amplitude without affecting the log-likelihood function. Every amplitude vector $\mathbf{a}'$ that fulfills $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}\mathbf{a}' = 0$ belongs to the nullspace of the matrix $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$. If vector $\mathbf{a}$ is a solution to Eq. (4.17) then $\mathbf{a} + \mathbf{a}'$ is a solution too. It is easy to show that an amplitude vector $\mathbf{a}'$ containing the same

constant in each of its elements (corresponding to an addition of a constant to every baseline amplitude) belongs to the nullspace. To remove this ambiguity we required that the sum of the amplitudes is zero (Sect. 4.4.1). This is equivalent to replacing the matrix $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ in Eq. (4.17) with a matrix that is the original matrix with number 1 added to each of its element. Provided that there are enough crossing points between the elementary scanning periods the modified matrix will be well-behaving. Note that this simple modification of $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ forces the sum of the amplitudes to zero, because the right hand side of Eq. (4.17) is orthogonal to the nullspace of $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$.

Because the matrix $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ is symmetric and non-negative definite, its eigenvectors form a complete orthogonal basis and its eigenvalues are real and ≥ 0. The eigenvectors with zero eigenvalue belong to the nullspace. The nullspace is not a problem when the amplitudes are solved using iterative conjugate gradient method. Provided that the initial guess of the amplitude vector is orthogonal to the nullspace the amplitude vector given by the conjugate gradient iterations will remain orthogonal to it (paper **I** [1]). For the uniform baselines this means that the sum of the fitted amplitudes will be zero.

When the amplitudes have been solved in Eq. (4.17) they shall be inserted in Eq. (4.16) to obtain the final output map.

Inserting Eq. (4.16) in Eq. (4.15) and assuming a uniform white noise covariance ($\mathbf{C}_n = \mathrm{diag}(\sigma^2)$), the log-likelihood function can be put in a similar form as the function S in Eq. (4.11). The weight will now be $w = 1/n_p$ (paper **I** [1]). The maps of this weight were compared to the maps of the earlier weights ($w = 1/(n_p - 1)$ of [70] and $w =$ const. of [71]). The comparison was made in paper **I** ([1]). The noise level in the maps of $w =$ const. was higher than in the maps of $w = 1/n_p$ but, the noise levels of $w = 1/n_p$ and $w = 1/(n_p - 1)$ were nearly equal.

To demonstrate the maximum likelihood destriping principle, one year noise-only TODs ($1/f$ + white noise) were generated for four LFI 30 GHz detectors (cf. Fig. 4.2). Level S software package was used in this task ([43]). The parameter values of the noise PSD were as in Fig. 4.2. One minute uniform baselines were fitted in the TODs of all four detectors and a single temperature map was binned from the cleaned TODs. A cycloidal precession scanning strategy with 7.5 deg amplitude (maximum deflection of the satellite spin axis above and below the ecliptic plane) and 6 months period was assumed (see Sect. 3.2). For comparison, another map was binned directly from the original TODs (no baseline subtraction). Both maps are shown in Fig. 4.3. The stripes are clearly visible in the binned map (top panel) but hardly visible in the output map of the destriping (bottom panel). The standard deviation (std) of the pixel temperatures of the destriped map is only 1.01 times higher than the pixel std of the theoretical white noise map, whereas for the binned map the corresponding factor is 9.3.

The angular power spectra of the temperature maps are shown in Fig. 4.4. The blob structure in the spectrum of the binned map is due to the stripes. Blobs dominate the spectral power and their level is dramatically reduced in the destriping.

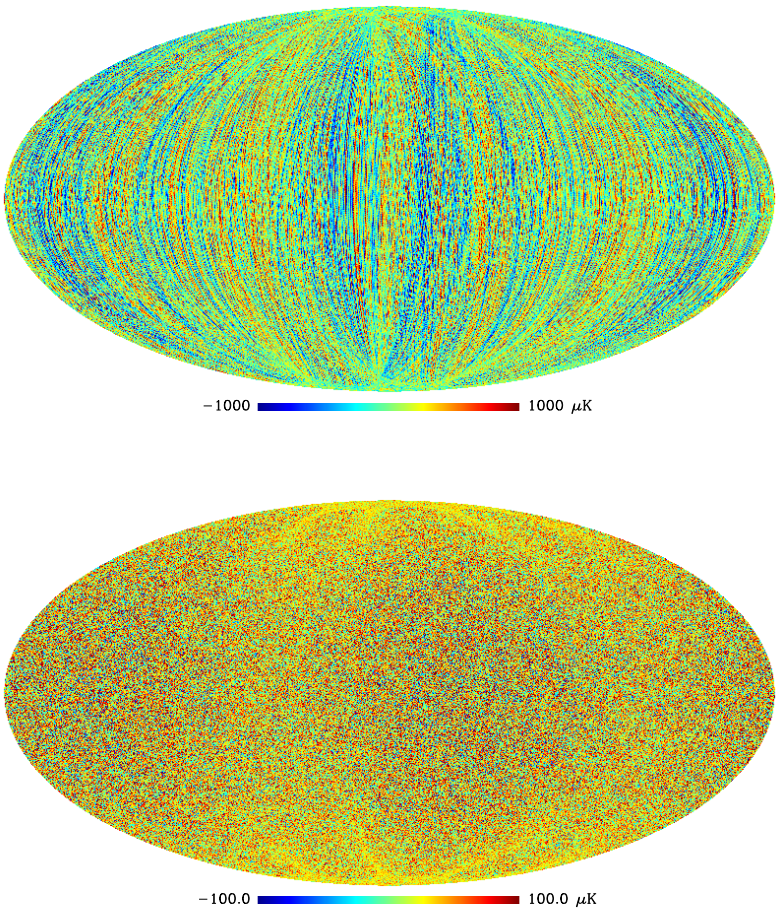

Figure 4.3: **Top:** A temperature map binned from the simulated one year noise TODs of four LFI 30 GHz detectors. The pixel std of the map is 388.0 μK. **Bottom:** A temperature map made from the same TODs but now the maximum likelihood destriping has been applied (one minute uniform baselines, no co-addition). The pixel std of this map is 42.03 μK. The pixel std of the theoretical white noise map is 41.66 μK. It was calculated as $\sqrt{\langle 1/n \rangle}\sigma$, where $\langle 1/n \rangle$ is the mean (taken over the hit pixels, 100% sky coverage in this case) of the inverse of the number of hits in a pixel and σ is the std of a sample of the TOD. The value of σ and the other noise parameters were as in Fig. 4.2. The map resolution is $N_{\rm side} = 512$ and the maps are in ecliptic coordinates. The scanning strategy was the cycloidal precession strategy and the units of temperature are antenna μKelvins. Note the different color scales of the maps.

Fig. 4.4 shows that the wanted CMB signal would be nearly buried under the noise spectrum unless the stripes were removed.

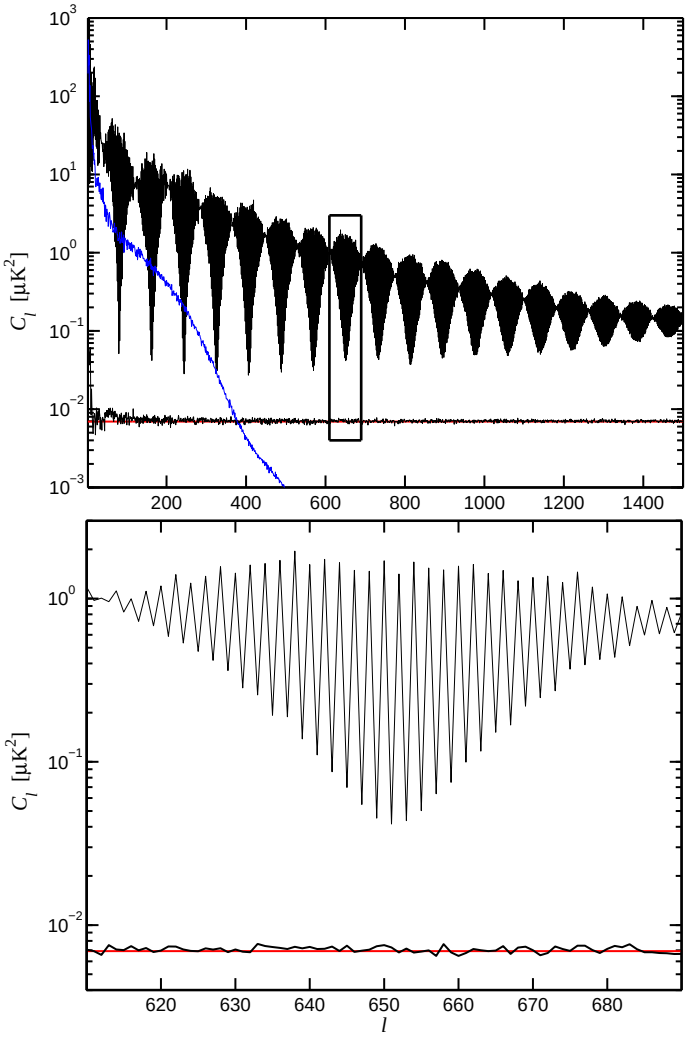

Figure 4.4: Angular power spectra of the noise maps of Fig. 4.3. **Top:** Power spectrum of the binned map (upper black curve with blobs), power spectrum of the destriped map (lower black curve) and the expectation value of the power spectrum of the white noise map (horizontal red line). The level of the white noise spectrum is $4\pi\sigma^2_{\rm map}/N_{\rm pix}$, where $\sigma_{\rm map}$ is the pixel std of the white noise map (see the caption of Fig. 4.3). A realization of a theoretical CMB power spectrum convolved with the beam of an LFI 30 GHz detector is shown for comparison (blue curve). The ΛCDM (cosmological constant + Cold Dark Matter) model in a flat universe was considered. **Bottom:** Zoom in the box shown in the top panel.

It was noted that the uniform baselines and white noise provide a reasonably accurate model of the instrument noise. One can try to improve the modelling by introducing more base functions in addition to the simple uniform baselines. Typical choices of base functions are sines and cosines (Fourier modes) or some polynomials.

The suitable frequencies of the Fourier modes are integer multiples of $1/T_{\rm period}$, where $T_{\rm period}$ is the duration of the elementary scanning period. These Fourier modes form an orthogonal set in $T_{\rm period}$. Legendre polynomials with the length $T_{\rm period}$ are orthogonal too. Eqs. (4.16) and (4.17) remain valid for multiple base functions with the modification that vector $\mathbf{a}$ now contains the amplitudes of the base functions and each column of the matrix $\mathbf{F}$ contains the values of the corresponding base function along the TOD. The matrix $\mathbf{F}^T\mathbf{F}$ is diagonal and we assume the normalization $\mathbf{F}^T\mathbf{F} = \mathrm{diag}(n_b)$ for the base functions. The number of samples in an elementary scanning period is n_b.

Delabrouille found that the addition of a small number of low order polynomials or splitting a ring in to a small number of distinct uniform baselines improved the performance of destriping ([70]). Maino et al. showed, however, that splitting a ring in two uniform baselines did not bring significant improvement ([73]). A possible reason for the different results was discussed in paper **I** ([1]). It could be due to the different noise spectra considered in these studies: $1/f$ for Maino et al. ([73]) and $1/f^2$ for Delabrouille ([70]).

We examined the maximum likelihood destriping using uniform baselines, first Fourier modes (sines and cosines with frequency $1/T_{\rm period}$) and 1st and 2nd order Legendre polynomials (paper **I** [1]). We assumed a co-added TOD leading to an elementary scanning period equal to a ring. We produced the results for three different knee frequencies (0.025 Hz, 0.1 Hz, and 0.4 Hz) of the simulated instrument noise. We found that for the two lowest knee frequencies multiple base functions did not reduce the map noise as compared to the noise of the uniform baselines only. At the highest knee frequency fitting uniform baselines and Legendre polynomials produced lower map noise than fitting uniform baselines only, but the difference was very small. Fitting first Fourier modes did not bring lower noise at the highest knee frequency.

For Fourier modes and 2nd order Legendre polynomials the CG algorithm required a remarkably large number of iterations to converge (paper **I** [1]). These problems seemed to be related to the symmetries of the used scanning strategy, which can create moderately small, but non-vanishing, eigenvalues of the matrix $\mathbf{F}^{\rm T}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ (paper **I** [1]). These small eigenvalues reduce the accuracy at which one can determine the amplitudes of the Fourier modes or 2nd order Legendre polynomials. There is clearly no advantage in fitting multiple base functions if their amplitudes contain large inaccuracy. The TODs used in paper **I** ([1]) represented the mission time of 7 months. Longer mission time or a different type of scanning strategy might increase the small eigenvalues and thus improve the accuracy of the base function amplitudes.

An alternative approach to solve the map from

$$\mathbf{y} = \mathbf{P}\mathbf{m} + \mathbf{F}\mathbf{a} + \mathbf{n}' \tag{4.19}$$

was proposed in [67]. Defining an extended map ($\mathbf{q}$)

$$\mathbf{q} = \begin{pmatrix} \mathbf{m} \\ \mathbf{a} \end{pmatrix} \tag{4.20}$$

and an extended pointing matrix ($\mathbf{Q}$)

$$\mathbf{Q} = \begin{pmatrix} \mathbf{P} & \mathbf{F} \end{pmatrix}. \tag{4.21}$$

we obtain for the TOD

$$\mathbf{y} = \mathbf{Q}\mathbf{q} + \mathbf{n}'. \tag{4.22}$$

Because we assume white noise, the minimum variance $\mathbf{q}$ is

$$\mathbf{q} = (\mathbf{Q}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Q})^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{y}. \tag{4.23}$$

$\mathbf{C}_n$ is the (diagonal) covariance matrix of the white noise. It can be shown that the map $\mathbf{m}$ and the base function amplitudes $\mathbf{a}$ obtained from Eq. (4.23) are identical to the map and the amplitudes obtained from Eqs. (4.16) and (4.17).

We may note that the maximum likelihood approach to destriping can be extended to treat polarised observations. Instead of just temperature, three Stokes parameters of the CMB anisotropy (I, Q and U, see the footnote in Sect. 3.1.4) needs to be solved for each map pixel. A row in the pointing matrix will now contain three non-zero elements that are the coefficients multiplying the Stokes parameters in the footnote. The matrix $\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P}$ will be block diagonal with 3×3 symmetric and non-negative definite matrices on its diagonals. The base function amplitudes and the output map (I, Q, U map) can still be solved from Eqs. (4.16) and (4.17). A maximum likelihood destriper (Polar) suitable for polarised observations has been developed ([78]). Some earlier implementations of the polarisation destripers exist too ([72], [80]).

4.4.3 Covariance of the Basefunction Amplitudes

Assuming that the white noise and the correlated noise components of the instrument noise of Eq. (4.14) are independent, the total noise covariance ($\mathbf{N}$) is

$$\mathbf{N} = \mathbf{C}_n + \mathbf{F}\mathbf{C}_a\mathbf{F}^{\mathrm{T}}. \tag{4.24}$$

Matrix $\mathbf{C}_a \equiv \langle \mathbf{a}\mathbf{a}^{\mathrm{T}} \rangle$ is the covariance matrix of the base function amplitudes. The amplitude covariance is ignored in the destriping. It is reasonable to expect that the noise performance of the output maps would be improved if all the available information (including the information on the amplitude covariance) is utilized when fitting the base functions. In principle, if one knows the power spectrum of the correlated part of the noise or has a good estimate of it, one is able to determine the covariance matrix $\mathbf{C}_a$.

Recently a map-making algorithm and a code MADAM (Map-making through Destriping for Anisotropy Measurements) have been introduced that exploit the covariance of the base function amplitudes ([77]). Assuming Gaussian distributions the base function amplitudes and the map can be solved by minimizing a log-likelihood function that is now (cf. Eq. (4.15)

$$\chi^2 = (\mathbf{y} - \mathbf{Pm} - \mathbf{Fa})^{\mathrm{T}} \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{Pm} - \mathbf{Fa}) + \mathbf{a}^{\mathrm{T}} \mathbf{C}_a^{-1} \mathbf{a}. \tag{4.25}$$

The equation for the base function amplitudes becomes

$$\left(\mathbf{F}^{\mathrm{T}} \mathbf{C}_n^{-1} \mathbf{Z} \mathbf{F} + \mathbf{C}_a^{-1}\right) \mathbf{a} = \mathbf{F}^{\mathrm{T}} \mathbf{C}_n^{-1} \mathbf{Z} \mathbf{y}. \tag{4.26}$$

After inserting the amplitudes, the map is solved from Eq. (4.16). A practical method to derive the amplitude covariance matrix from the known (or estimated) noise power spectrum is presented in [77]. In this method the autocorrelation function of the noise is expanded in terms of real exponential functions. The exponentials allow an efficient calculation of the amplitude covariance matrix. The exponential model holds for e.g. $1/f^{\alpha}$ type of noise ($0 < \alpha \leq 2$).

It was shown in [77] that if $\mathbf{n} = \mathbf{Fa} + \mathbf{n}'$ is an accurate description of the instrument noise the output map of the MADAM algorithm equals the minimum variance map.

The MADAM algorithm is a step aside from the traditional path of the destriping, where no prior knowledge was assumed on the noise. To gain the full advantage of the MADAM algorithm we need to estimate the noise PSD from the observed data. The iterative and non-iterative methods that were described in Sect. 4.3 are applicable here.

Temperature maps made by MADAM were compared (in [77]) to the maps of pure destriping (no amplitude covariance). Perfectly known noise PSD was assumed in MADAM. Uniform baselines, Fourier modes and Legendre polynomials were considered as base functions. Because $\mathbf{F}^{\mathrm{T}} \mathbf{C}_n^{-1} \mathbf{Z} \mathbf{F}$ is non-negative definite and $\mathbf{C}_a^{-1}$ is positive definite, the matrix $\mathbf{F}^{\mathrm{T}} \mathbf{C}_n^{-1} \mathbf{Z} \mathbf{F} + \mathbf{C}_a^{-1}$ is positive definite and thus well-behaving and no CG convergence problems occur for any set of base functions. In MADAM the increase of the number of base functions will lead to improved modelling of the correlated noise. When subtracted from the observed TOD the level of the correlated noise in the cleaned TOD is reduced leading to maps with lower noise. In the pure destriping the accuracy of the fitted base function amplitudes was poorer (due to small eigenvalues of $\mathbf{F}^{\mathrm{T}} \mathbf{C}_n^{-1} \mathbf{Z} \mathbf{F}$) and the reduction of the correlated noise was less effective. Similar behavior was experienced with uniform baselines of different lengths. At one minute and longer baselines the maps of MADAM and pure destriping had nearly the same noise level. When the length of the baseline is shortened the modelling of the correlated noise should improve and the output maps should have lower noise. This happened in the MADAM maps ([77]) but shorter baselines did not bring similar noise reduction in the maps of pure destriping (again due to poorer accuracy of the baseline magnitudes).

A qualitative explanation to these results can be given by comparing the matrices $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ and $\mathbf{C}_a^{-1}$. Let us consider one minute uniform baselines first. We use the noise properties of an LFI 30 GHz detector as an example (see Fig. 4.2) and assume that the diagonals of $\mathbf{C}_n$ are all equal (σ^2). Matrices $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ and $\mathbf{C}_a^{-1}$ are diagonally dominant. The order of magnitude of the diagonal elements of $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ is n_b/σ^2. The diagonals of $\mathbf{C}_a^{-1}$ are $\sim 1/\sigma_{a_i}^2$, where $\sigma_{a_i}^2$ is the variance of the baseline amplitude a_i. For an LFI 30 GHz detector we have $n_b = 1950$ and $\sigma_{a_i}^2 \approx 2\sigma^2$ for one minute uniform baselines[1]. Thus the diagonals of $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$ are about three orders of magnitude larger than the diagonals of $\mathbf{C}_a^{-1}$ making the matrix $\mathbf{C}_a^{-1}$ nearly insignificant relative to the matrix $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$.

When the length of the uniform baselines is made shorter, the factor n_b becomes smaller, but, due to the $1/f^\alpha$ spectral behavior of the correlated noise, the change of the variance of the baseline amplitudes is marginal. This leads to the increase of the relative importance of the amplitude covariance (compared to $\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}$). Thus the baseline amplitudes of MADAM and pure destriping will be different for short uniform baselines. Due to the $1/f^\alpha$ spectral behavior of the correlated noise, the variances of the amplitudes of the Fourier modes will decrease with increasing Fourier mode frequency leading to the increase of the relative importance of $\mathbf{C}_a^{-1}$ and the amplitudes will become different for MADAM and for pure destriping.

A version of MADAM that makes (I,Q,U) maps from the polarisation observations is currently under development ([78]).

4.5 Comparison of GLS and Destriping

A comparison of temperature maps made by GLS map-making algorithm and destriping was carried out in this thesis (paper **III** [3]). In this study GLS and destriped maps were made from the same TODs and then compared. A considerable part of this work was carried out in two meetings of PLANCK Working Group 3 (WG3) that were held in June 2002 and in January 2003. Simulated one year TODs (signal+noise) representing a single LFI 100 GHz detector were generated. The TODs were made using the Level S simulation package ([43]). At the time of the meetings the LFI plan included 24 detectors (12 horn antennas) at 100 GHz but they were later discarded due to budgetary reasons.

The target of this section of the thesis is to give a summary of the key findings of the comparison study described in paper **III** ([3]).

It should be noted that a similar comparison study but covering, e.g., multiple PLANCK detectors, polarisation observations and more realistic foregrounds, telescope beams and satellite spin axis nutations is currently on-going in WG3 ([79]).

[1]For uniform baselines $\sigma_{a_i}^2 = 2\int_0^{f_s/2} S(f)|G(f)|^2 df$, where $S(f)$ is the $1/f^\alpha$ part of the noise PSD (Eq. (3.20)) and $G(f) = \frac{\sin(\pi f T)}{\pi f T}$, where T is the length (in time) of the uniform baselines.

4.5.1 Relation between GLS and Destriping Algorithms

This section contains a summary of the Appendix A of paper **III** ([3]). We assumed that the expected power of the white noise is identical in every TOD sample. This leads to a uniform diagonal covariance matrix for the white noise: $\mathbf{C}_n = \mathrm{diag}(\sigma^2)$. In the destriping the base function amplitudes and the output map are solved from the equations (see Eqs. (4.16) and (4.17))

$$\mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{F}\mathbf{a} = \mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{y} \tag{4.27}$$

and

$$\mathbf{m} = (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}(\mathbf{y} - \mathbf{F}\mathbf{a}). \tag{4.28}$$

The matrix $\mathbf{Z}$ is now $\mathbf{Z} = \mathbf{I} - \mathbf{P}(\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}$.

The observed TOD ($\mathbf{y}$) contains two components: signal $\mathbf{s}$ and instrument noise $\mathbf{n}$ (Eq. (4.1)). The term $\mathbf{Z}\mathbf{y}$ (in Eq. (4.27)) can now be split in two parts

$$\mathbf{Z}\mathbf{y} = \mathbf{Z}\mathbf{s} + \mathbf{Z}\mathbf{n}. \tag{4.29}$$

After writing out, the first term on the right hand side is

$$\mathbf{Z}\mathbf{s} = \mathbf{s} - \mathbf{P}(\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{s}. \tag{4.30}$$

Apart from the sign, this is the pixelization noise introduced in Sect. 4.2 ([55]). We denote the pixelization noise by a vector $\mathbf{p}$ and write $\mathbf{p} = -\mathbf{Z}\mathbf{s}$. The split of $\mathbf{Z}\mathbf{y}$ in two components means that the base function amplitudes can be split analogously: $\mathbf{a} = \mathbf{a}_p + \mathbf{a}_n$, where the source of $\mathbf{a}_p$ is the pixelization noise and the source of $\mathbf{a}_n$ is the instrument noise. The amplitudes are solutions to equations

$$\mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{F}\mathbf{a}_p = -\mathbf{F}^{\mathrm{T}}\mathbf{p} \tag{4.31}$$

and

$$\mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{F}\mathbf{a}_n = \mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{n}. \tag{4.32}$$

Inserting $\mathbf{y} = \mathbf{s} + \mathbf{n}$ and $\mathbf{a} = \mathbf{a}_p + \mathbf{a}_n$ in Eq. (4.28) we obtain for the output map

$$\mathbf{m} = (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{s} - (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{F}\mathbf{a}_p + (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}(\mathbf{n} - \mathbf{F}\mathbf{a}_n). \tag{4.33}$$

The first term on the right side is the binned noiseless map. Ideally, we would like the output map to be equal to the binned noiseless map. The rest of the terms bring error. The sum of the second and the third terms is called the *reconstruction error map*. The second term is the signal component of the reconstruction error map and the third term is its noise component, respectively.

It was shown in the Appendix A of paper **III** ([3]) that the output map of the GLS map-making (that is a solution to $\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{P}\mathbf{m} = \mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{y}$) can be put in a similar form

$$\mathbf{m} = (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{s} - (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{H}\widetilde{\mathbf{\Delta}}_p + (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}(\mathbf{n} - \mathbf{H}\widetilde{\mathbf{\Delta}}_n). \tag{4.34}$$

Here the complex valued matrix $\mathbf{H}$ (dimension equal to the number of samples in the TOD) is an inverse Discrete Fourier Transform (DFT) operator that converts frequency domain vectors to time domain (to TOD domain): $\mathbf{y} = \mathbf{H}\widetilde{\mathbf{y}}$, where $\widetilde{\mathbf{y}}$ is the frequency domain counterpart of $\mathbf{y}$. The inverse operator to $\mathbf{H}$ is its Hermitian conjugate $\mathbf{H}^{\dagger}$. We have assumed that the matrix $\mathbf{H}$ is normalized to $\mathbf{H}\mathbf{H}^{\dagger} = \mathbf{H}^{\dagger}\mathbf{H} = \mathbf{I}$. The lengths of the frequency domain vectors $\widetilde{\mathbf{\Delta}}_p$ and $\widetilde{\mathbf{\Delta}}_n$ equal the length of the TOD. The source of $\widetilde{\mathbf{\Delta}}_p$ is the pixelization noise and the source of $\widetilde{\mathbf{\Delta}}_n$ is the instrument noise. They are solutions to equations

$$[\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{H} + (\widetilde{\mathbf{N}} - \mathbf{I})^{-1}]\widetilde{\mathbf{\Delta}}_p = -\mathbf{H}^{\dagger}\mathbf{p} \tag{4.35}$$

and

$$[\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{H} + (\widetilde{\mathbf{N}} - \mathbf{I})^{-1}]\widetilde{\mathbf{\Delta}}_n = \mathbf{H}^{\dagger}\mathbf{Z}\mathbf{n}. \tag{4.36}$$

The matrix $\widetilde{\mathbf{N}}$ is the Fourier transform of the noise covariance matrix $\mathbf{N}$ normalized with the power of the white noise: $\widetilde{\mathbf{N}} = \mathbf{H}^{\dagger}\mathbf{N}\mathbf{H}/\sigma^2$. Because $\mathbf{N}$ is circulant, $\widetilde{\mathbf{N}}$ is diagonal. The frequency bins of the noise power spectrum are the diagonal elements of $\mathbf{H}^{\dagger}\mathbf{N}\mathbf{H}$. Eqs. (4.34), (4.35) and (4.36) were the main theoretical results of paper **III** ([3]).

For the MADAM algorithm, that utilizes the covariance of the baseline amplitudes (matrix $\mathbf{C}_a$), the vectors $\mathbf{a}_p$ and $\mathbf{a}_n$ are solved from (cf. Eqs. (4.31) and (4.32))

$$\left(\mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{F} + \sigma^2\mathbf{C}_a^{-1}\right)\mathbf{a}_p = -\mathbf{F}^{\mathrm{T}}\mathbf{p} \tag{4.37}$$

and

$$\left(\mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{F} + \sigma^2\mathbf{C}_a^{-1}\right)\mathbf{a}_n = \mathbf{F}^{\mathrm{T}}\mathbf{Z}\mathbf{n}. \tag{4.38}$$

If these equations are compared to Eqs. (4.35) and (4.36) a close resemblance can be detected. If we would set the length of the base functions equal to the length of the TOD and use the complete set of Fourier mode base functions, the amplitude vectors $\mathbf{a}_p$ and $\mathbf{a}_n$ solved from Eqs. (4.37) and (4.38) would be identical to the vectors $\widetilde{\mathbf{\Delta}}_p$ and $\widetilde{\mathbf{\Delta}}_n$ that are solutions to Eqs. (4.35) and (4.36). The GLS and MADAM maps would be identical in this case (minimum variance maps). One can say that GLS map-making is related to MADAM in the following sense: GLS map-making effectively fits a complete orthogonal set of Fourier modes to the correlated part of the noise, subtracts the Fourier modes, after weighting them with the fitted amplitudes, from the original observed TOD and bins the output map from the cleaned TOD.

4.5.2 Comparison of Temperature Maps

We compared temperature maps that were made using three distinct map-making algorithms: two GLS codes and one destriping code (paper **III** [3]). The GLS algorithms were the Roma Optimal Mapmaking Algorithm (ROMA, [60], [61]) and MapCUMBA (originally introduced in [55], current version based on preconditioned conjugate gradient principle). The destriping code fitted a uniform baseline to a ring, that was co-added from 60 scan circles. The destriping code used in this study is described in paper **I** ([1]).

The simulation parameters of the TOD generation are given in Table 4.1. Signal-only and signal+noise output maps were made using all three map-making methods. The resolution of the output maps was $N_{\rm side} = 512$. The sky coverage at this resolution and for the applied scanning strategy was 100%. An example of an output map of destriping is shown in Fig. 4.5.

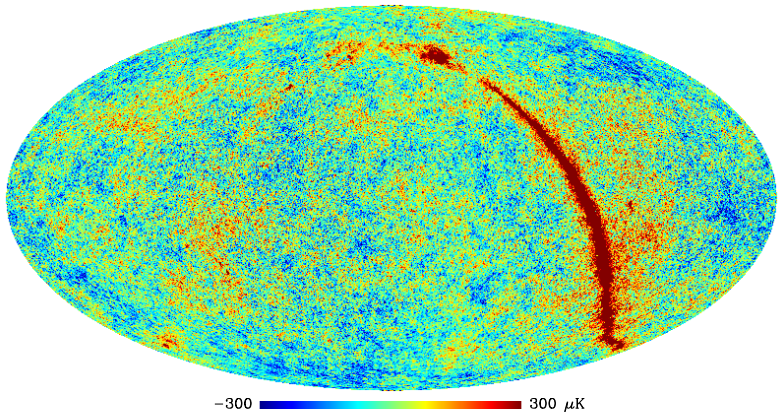

Figure 4.5: A temperature map made from the simulated observed TOD that contained CMB, foreground and instrument noise (case 2). The map-making method was destriping. The map represents an observation of a single LFI 100 GHz detector. An elliptic telescope beam was assumed. For this plot the map resolution was degraded to $N_{\rm side} = 256$. The color scale is in antenna temperature units.

It was shown in Eqs. (4.33) and (4.34) that the output maps of GLS and destriping contain three parts: binned noiseless map ($1^{\rm st}$ term), signal component of the reconstruction error map ($2^{\rm nd}$ term) and noise component of the reconstruction error map ($3^{\rm rd}$ term). The $1^{\rm st}$ term is the wanted map and the $2^{\rm nd}$ and $3^{\rm rd}$ terms bring error. The goal was to compare the magnitudes of the $2^{\rm nd}$ and $3^{\rm rd}$ terms produced by GLS and destriping algorithms. The pixel std of the three components of the output maps are shown in Table 4.2. The results are for the elliptic beam. The results for the symmetric beam are available in paper **III** ([3]).

Table 4.1: Simulation parameters in the comparison of the temperature maps (paper **III** [3]). Two TODs (signal-only and signal+noise) were generated for all 4 simulation cases leading to 8 TODs in total. The TODs are sums of CMB (C), foreground (F), and/or noise (N) as indicated in the table. The differences between the cases were the knee frequencies of the instrument noise, beams and that the foreground emissions were in cases 1 and 2 only. The TODs were generated using Level S software ([43]).

Parameters common to all cases		
Detector		LFI 100 GHz
Number of detectors		1
Mission time		12 months
Scanning (a)		Cycloidal
Noise		
	σ (b)	3957.26 μK
	$f_{\rm min}$	10^{-4} Hz
	$f_{\rm s}$	108.3 Hz

Parameters for	Case 1	Case 2	Case 3	Case 4
TOD (S only)	C+F	C+F	C	C
TOD (S+N)	C+F+N	C+F+N	C+N	C+N
Beam (c,d)	Symmetric	Elliptic	Symmetric	Elliptic
Noise - $f_{\rm k}$	0.03 Hz	0.03 Hz	0.1 Hz	0.1 Hz

(a) 10 deg amplitude and 6 months period.
(b) White noise std of the detector TOD (in antenna temp. units).
(c) Symmetric Gaussian beam:
Full width half maximum (FWHM) = 10.6551 arcmin.
(d) Elliptic Gaussian beam:
$\mathrm{FWHM}_{\rm major}$ = 11.8652 arcmin, $\mathrm{FWHM}_{\rm minor}$ = 9.5684 arcmin.

The data in the last column of Table 4.2 is for the residual map of the instrument noise. It was produced by subtracting the output map of the signal-only (noiseless) observations from the output map of the signal+noise observations. As expected, the GLS map-making produces lower map noise than destriping. The difference is, however, small: 0.15% in case 2 and 0.35% in case 4. The higher noise level in case 4 is caused by the higher knee frequency. The noise in the map gives a bias to the angular power spectrum and impacts the error bars of the power spectrum estimate. Assuming that the magnitude of the noise bias is accurately determined (using e.g. noise-only Monte Carlo simulations) and removed from the power spectrum the impact on the error bars remains. It was shown in paper **III** ([3]) that due to the higher map noise the error bars of destriping would be (in this case) $\sim$5% larger than the error bars of GLS.

Table 4.2: Results of the comparison of the temperature maps (paper **III** [3]). The numbers are the pixel std of the maps and they are given in the antenna temperature units (at 100 GHz). The data is for the elliptic beam TODs. The corresponding data for the symmetric beam (cases 1 and 3) are given in paper **III** ([3]). The numbers for ROMA and MapCUMBA were nearly identical. The ROMA data is given under the title "GLS". The numbers were calculated from $N_{\mathrm{side}} = 512$ full sky maps. The columns refer to the 1^{st}, 2^{nd} and 3^{rd} terms of Eqs. (4.33) and (4.34). The pixel std of the white noise maps were calculated as in Fig. 4.3. The white noise values are slightly different between the cases due to small differences in the scannings.

	Binned noiseless	Rec. error: signal	Rec. error: noise
Case 2			
GLS	174.057 μK	0.6140 μK	138.246 μK
Destriping	174.057 μK	0.2518 μK	138.454 μK
White noise			137.219 μK
Case 4			
GLS	78.209 μK	0.4064 μK	138.910 μK
Destriping	78.209 μK	0.1691 μK	139.397 μK
White noise			137.400 μK

It can be noted that the map noise of a single LFI 100 GHz detector ($\sim$139 μK) is slightly lower than the noise of the first year W band frequency map ($\sim$142 μK) of the WMAP experiment ([48]). The W band map combines data from eight WMAP detectors (four differencing assemblies) operating at a center frequency of 94 GHz.

The data in the middle column of Table 4.2 is the map-making error due to the signal itself (signal component of the reconstruction error map). These error maps were obtained by subtracting the binned noiseless map from the output maps of the signal-only (noiseless) observations. The error is larger in the GLS maps than in the destriped maps. The presence of the foreground signal increases the signal component of the reconstruction error (case 2 in Table 4.2). Although the mean level (std) of the reconstruction error is small compared to the error due to the instrument noise, foreground signal produces large errors in some individual pixels. This can be seen in Table 4.3, which gives the minimum and maximum pixel temperatures of the signal component of the reconstruction error. Especially the extreme pixel errors of GLS are significant (case 2).

Error maps for GLS and destriping are shown in Fig. 4.6. It shows that the large errors of the GLS maps tend to locate in the vicinity of the strongest foreground signal. A galactic cut in the GLS map might reduce these errors. For destriping this error appears as erroneous baselines corresponding to an offset of an entire ring. Therefore a similar spatial correlation between the errors and the foreground signal

Table 4.3: Minimum and maximum pixel temperatures of the signal component of the reconstruction error map (2^{nd} term of Eqs. (4.33) and (4.34)). The numbers are in the antenna temperature units. The data is for the elliptic beam. The data for the symmetric beam (cases 1 and 3) is available in paper **III** ([3]). The ROMA data is given under the title "GLS". The numbers were calculated from $N_{\text{side}} = 512$ full sky maps.

Rec. error: signal	Min	Max
Case 2		
GLS	-62.2 μK	37.8 μK
Destriping	-2.3 μK	1.8 μK
Case 4		
GLS	-2.8 μK	3.1 μK
Destriping	-0.9 μK	0.8 μK

is not visible in the destriped map, although we expect that the baselines mainly originate from where the foreground signal is strong.

The source of the signal component of the reconstruction error (ε_p) is the pixelization noise (see the right hand sides of Eqs. (4.31) and (4.35)). We can expect that in the destriping the uniform baselines of the pixelization noise contribute to this error. Because we assumed an exact repetition of the pointings of the 60 circles of a ring, the pixelization noise of those 60 circles contains the same signal repeating itself 60 times. The length of the repeating signal is $T_{\text{period}} = 60$ s. A periodic signal (with period T_{period}) has a Fourier series representation whose Fourier mode frequencies are multiples of $1/T_{\text{period}}$. The lowest (zero) mode contributes to the uniform baselines. In the GLS map-making, however, Fourier modes up to the knee frequency (0.1 Hz) of the instrument noise contribute to this error (paper **III** [3]). This is caused by the noise filter that is determined from the known noise PSD. The difference in the contributing pixelization noise power (larger in GLS) explains why the signal component of the reconstruction error is larger for the GLS map-making than for the destriping.

The galactic foreground signal has a larger spatial variation than the CMB signal. This causes a higher pixelization noise, which explains the increased error when the foreground signal is included in the simulations.

The map-making error due to the CMB signal will cause a bias in the CMB angular power spectrum estimate. In this study the bias was marginal for destriping but more notable for GLS. One can use signal-only Monte Carlo simulations to determine the magnitude of this bias, which can then be used to correct the obtained power spectra. The bias due to the foregrounds was not examined.

Finally, it should be noted that the magnitude of the signal component of the reconstruction error will increase in destriping if the number of base functions is

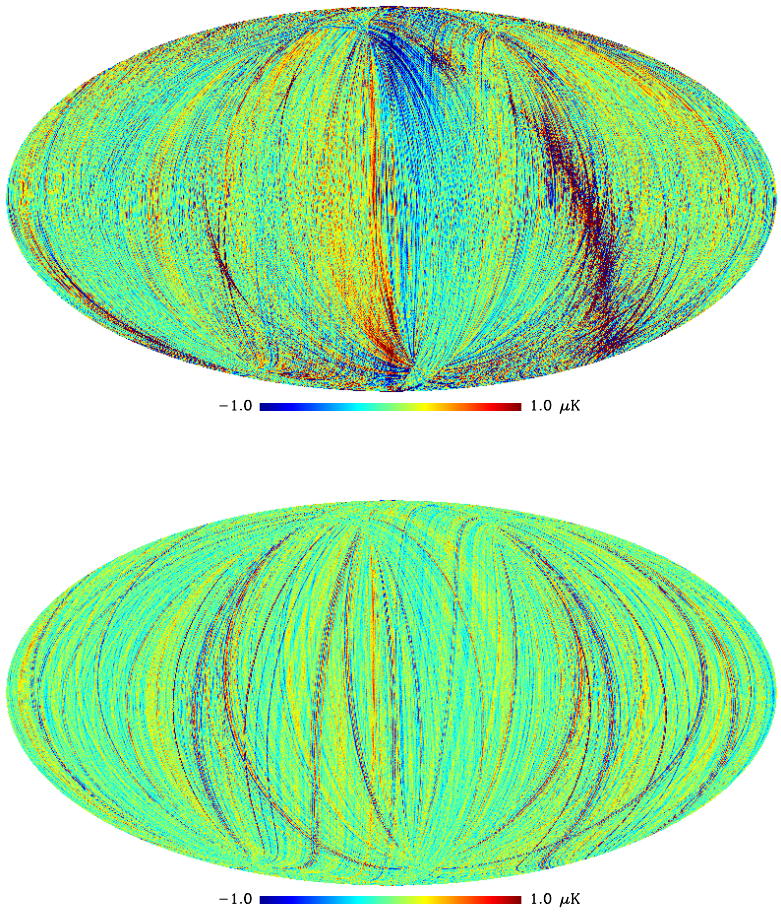

Figure 4.6: Map-making error due to the signal (signal component of the reconstruction error). These maps were obtained by subtracting the binned noiseless map from the output maps of the signal-only (noiseless) observations. The signal contained CMB and foreground emissions and an elliptic beam was assumed (case 2). The maps represent observations of a single LFI 100 GHz detector. The map resolutions are $N_{\mathrm{side}} = 512$. The color scale is in antenna temperature units. **Top:** GLS (ROMA). **Bottom:** Destriping.

increased ([77]) or if the uniform baselines are made shorter. However, the error remains typically small compared to the error due to the instrument noise.

4.6 Deconvolution Map-making and Destriping

Destriping algorithms described in this thesis produce maps that are not estimates of the true sky directly but they are estimates of the true sky convolved with the

instrument response. No destriping algorithm has been introduced until now that could produce the estimate of the true sky directly. Modifications to the existing destriping algorithms are proposed here that would produce these estimates.

In principle the base function amplitudes should be solved from Eq. (4.17) or (4.26) using the observation matrix $\mathbf{A}$ (introduced in Sect. 4.1) instead of the simple pointing matrix $\mathbf{P}$. Because the observation matrix is not sparse and because the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{A}$ (required to calculate the matrix $\mathbf{Z}$) is large and non-diagonal, this way to solve the amplitudes is a computationally heavy task.

The destriping studies described and referenced in this thesis show clearly that the base function amplitudes are dominated by the instrument noise, which is not influenced by the beam used in the experiment. The baselines solved using the pointing matrix provide a good approximation of the correlated part of the instrument noise. We can assume that in most of the cases the baseline magnitudes would be essentially the same whether we use the observation matrix or the pointing matrix when solving them. We suggest that when the asymmetric beams are introduced in the destriping we go on solving the base function amplitudes using the simple pointing matrix. As before, the cleaned TOD ($\mathbf{d}$) is obtained by subtracting the base functions from the observed TOD ($\mathbf{y}$)

$$\mathbf{d} = \mathbf{y} - \mathbf{F}\mathbf{a}. \tag{4.39}$$

Instead of simply binning the output map, we need to solve it from the equation

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{m} = \mathbf{A}^{\mathrm{T}}\mathbf{d}. \tag{4.40}$$

A direct inversion of the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ at the required map resolutions is too demanding a job for the current computers. A method to solve the map without storing the observation matrix has recently been proposed ([81]). It is called *deconvolution mapmaking*. A short description of that approach is given here.

Because the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is symmetric and non-negative definite, CG iteration is a suitable method to solve the map from Eq. (4.40). We need to evaluate the right hand side of Eq. (4.40) first. It is a map domain vector. During the CG phase we need to apply the operator $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ to a map once in every iteration.

An observation matrix acting on a map ($\mathbf{A}\mathbf{m}$) produces a TOD that encodes the scan path and whose samples are convolutions of the telescope response and the map. The operation $\mathbf{A}\mathbf{m}$ is called *convolution* ([81]). The details of the convolution algorithm proposed in [81] are given in [82]. The same algorithm is applied in the totalconvolve module of the Level S software ([43]).

Let us assume that we have a temperature field $T(\mathbf{n})$ on the sky, where $\mathbf{n} = (\theta, \varphi)$ are the spherical coordinates of a point in the celestial sphere. The spherical harmonic expansion coefficients of $T(\mathbf{n})$ are $a_{\ell m}$. The instrument beam response is $B(\mathbf{n}, \mathbf{n}_0, \psi)$, where $\mathbf{n}_0$ is the pointing of the center of the beam and ψ is its orientation angle (see Sect. 3.1.3). The beam response has a spherical harmonic expansion given in Eq. (3.4). The temperature of the sky observed through the instrument is (cf. Eq.

(3.19))[2]

$$\widetilde{T}(\mathbf{n}_0,\psi)=\int_{4\pi}d\Omega_{\mathbf{n}}B^*(\mathbf{n},\mathbf{n}_0,\psi)T(\mathbf{n})=\sum_{\ell m}b^*_{\ell m}(\mathbf{n}_0,\psi)a_{\ell m}, \tag{4.41}$$

where asterisk indicates the complex conjugate. The expansion coefficients of the beam depend on its pointing and orientation according to Eq. (3.5). Inserting in Eq. (4.41) we obtain

$$\widetilde{T}(\varphi_0,\theta_0,\psi)=\sum_{\ell=0}^{\ell_{\max}}\sum_{m,m'=-\ell}^{\ell}b^*_{\ell m'}a_{\ell m}D^{\ell *}_{mm'}(\varphi_0,\theta_0,\psi), \tag{4.42}$$

where $b_{\ell m}$ are the expansion coefficients of the beam in its reference pointing and orientation. It is assumed that both the sky and the beam are band limited to some $\ell_{\max}$. Using the definition of the Wigner D-functions ([83])

$$D^{\ell}_{mm'}(\varphi_0,\theta_0,\psi)=e^{-im\varphi_0}d^{\ell}_{mm'}(\theta_0)e^{-im'\psi} \tag{4.43}$$

and the relation (Eq. (10) in p. 113 of [83])

$$D^{\ell}_{mm'}(\varphi_0,\theta_0,\psi)=\sum_{m''=-\ell}^{\ell}d^{\ell}_{mm''}(\pi/2)d^{\ell}_{m''m'}(\pi/2)e^{-i(m\varphi_0+m''\theta_0+m'\psi)} \tag{4.44}$$

the observed temperature can be expressed as

$$\widetilde{T}(\varphi_0,\theta_0,\psi)=\sum_{m,m'',m'=-\ell_{\max}}^{\ell_{\max}}T_{mm''m'}e^{i(m\varphi_0+m''\theta_0+m'\psi)}. \tag{4.45}$$

The reduced Wigner matrices are denoted by $d^{\ell}_{mm'}$ and $T_{mm''m'}$ is the spherical harmonic representation of the observed temperature field $\widetilde{T}(\varphi_0,\theta_0,\psi)$

$$T_{mm''m'}=\sum_{\ell=0}^{\ell_{\max}}b^*_{\ell m'}a_{\ell m}d^{\ell}_{mm''}(\pi/2)d^{\ell}_{m''m'}(\pi/2). \tag{4.46}$$

In the case of a mild beam asymmetry we do not have to consider the full $(-\ell_{\max},\ell_{\max})$ range of m' values but it is sufficient to limit to $(-m_{\max},m_{\max})$, where $m_{\max}\ll\ell_{\max}$. The inverse Fourier transform of $T_{mm''m'}$ creates a discrete temperature field $\widetilde{T}(\varphi_0,\theta_0,\psi)$, which is tabulated in an equally spaced three dimensional array (3D array). The array size is $2\ell_{\max}+1$ in the φ_0 and θ_0 dimensions and $2m_{\max}+1$ in the ψ dimension. Each point of the scan path has a distinct $(\varphi_0,\theta_0,\psi)$. The TOD is generated by picking values of the 3D array that fall closest to $(\varphi_0,\theta_0,\psi)$. Polynomial or Fourier series interpolation can be applied to improve the accuracy of the TOD samples ([43]). This completes the convolution part of the $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ operation.

[2]It can be shown that if $F(\mathbf{n})$ and $G(\mathbf{n})$ are two fields in the sky and if $f_{\ell m}$ and $g_{\ell m}$ are their spherical harmonic expansion coefficients then $\int_{4\pi}d\Omega_{\mathbf{n}}F^*(\mathbf{n})G(\mathbf{n})=\sum_{\ell m}f^*_{\ell m}g_{\ell m}$.

The matrix $\mathbf{A}^{\mathrm{T}}$ acting on a TOD is called *a transpose convolution* ([81]). The result of this operation is a map domain vector. In terms of the 3D array of the observed temperatures the map produced by the transpose convolution can be given by

$$\widehat{T}(\mathbf{n}_p) = \sum_{\varphi_0,\theta_0,\psi} B(\mathbf{n}_p, \varphi_0, \theta_0, \psi)\widetilde{T}(\varphi_0, \theta_0, \psi). \tag{4.47}$$

Here $\mathbf{n}_p$ is the pointing to a pixel p and the sums are taken over the dimensions of the 3D array. Because B and $\widetilde{T}$ fields can be represented by three dimensional Fourier series in $(\varphi_0, \theta_0, \psi)$, the sums can be carried out easily and the spherical harmonic expansion coefficients of $\widehat{T}(\mathbf{n}_p)$ are

$$\hat{a}_{\ell m} = \sum_{m'',m'} b_{\ell m'} d^{\ell}_{mm''}(\pi/2) d^{\ell}_{m''m'}(\pi/2) T_{mm''m'}. \tag{4.48}$$

This completes the transpose convolution.

The tools for processing the right hand side of Eq. (4.40) are available now. For each sample of the cleaned TOD ($\mathbf{d}$) the scanning strategy specifies $(\varphi_0, \theta_0, \psi)$. As a first step we bin $\mathbf{d}$ in an initially empty 3D array which is then Fourier transformed to produce its $T_{mm''m'}$. Applying Eq. (4.48) the spherical harmonic coefficients $(\gamma_{\ell m})$ of the map $\mathbf{A}^{\mathrm{T}}\mathbf{d}$ are obtained.

Assume that $s_{\ell m}$ represents an initial guess map ($\mathbf{s}$) of the CG iteration. The following steps are carried out.

1. Using Eq. (4.46) calculate $T_{mm''m'}$.
2. Inverse Fourier transform $T_{mm''m'}$ to produce its 3D array $(\widetilde{T}(\varphi_0, \theta_0, \psi))$.
3. Scan the 3D array to a TOD.
4. Bin the TOD back to an initially empty 3D array.
5. Fourier transform the 3D array to produce its $T_{mm''m'}$.
6. Using Eq. (4.48) calculate the spherical harmonic coefficients $(s'_{\ell m})$ of the map $\mathbf{s}' = \mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{s}$.
7. Use $s'_{\ell m}$ and $\gamma_{\ell m}$ to steer the iteration to produce a better guess of the map $(s_{\ell m})$.
8. When the iterations stop we will have the $s_{\ell m}$ of the output map.

The above list shows the formal steps of the $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ operation and it might not be a good recipe for the development of an efficient software code. During the actual software implementation the steps shall be reconsidered and some of them might be combined and arranged differently.

Because the beam response washes out the small scale structures of the sky, its high-ℓ components are small which may lead to an ill-conditioned $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ matrix ([81]). It can be regularized using $b_{\ell m}/G_\ell$ instead of $b_{\ell m}$ in Eqs. (4.46) and (4.48). G_ℓ is a symmetric smoothing function (e.g. symmetric Gaussian beam that approximates the actual beam). In this case the output map will be the true sky convolved with the symmetric beam G_ℓ.

The number of CG iterations can be reduced using a preconditioning matrix. A diagonal matrix that had the inverse of the number of hits in its diagonal was applied in [81].

Chapter 5

Angular Power Spectrum Estimation

5.1 Estimation Problem

We assume that the CMB observations have been reduced to a temperature map $\mathbf{x}$. The temperature of a pixel is a sum of signal and instrument noise: $x_i = s_i + n_i$. Subscript i indexes the pixels and $N_{\rm pix}$ is the number of pixels in the sky. Because the signal is statistically independent from the noise, we obtain for the map covariance

$$\mathbf{C} \equiv \langle \mathbf{x}\mathbf{x}^{\rm T} \rangle = \langle \mathbf{s}\mathbf{s}^{\rm T} \rangle + \langle \mathbf{n}\mathbf{n}^{\rm T} \rangle = \mathbf{S} + \mathbf{N}. \tag{5.1}$$

Here $\langle \cdot \rangle$ is the ensemble mean. The size of the covariance matrix is $N_{\rm pix} \times N_{\rm pix}$ and the maps are real-valued column vectors. The matrices $\mathbf{C}$, $\mathbf{S}$ and $\mathbf{N}$ are symmetric.

We assume that the signal is the CMB temperature anisotropy signal that has been observed through the telescope response. The signal of the map can be expanded in terms of spherical harmonics

$$s_i = \sum_{\ell m} a_{\ell m} B_\ell Y_{\ell m}(\mathbf{n}_i). \tag{5.2}$$

Here $\mathbf{n}_i$ is a unit vector pointing to the center of the pixel i. This model of the map signal does not cover the effects that arise from the distribution of the observations across the pixel area. We will discuss those effects in Sect. 5.5.3. Here we assume a radially symmetric telescope response which is modelled by the factor B_ℓ. The coefficients $a_{\ell m}$ represent the CMB temperature fluctuations of the sky. We assume that they are Gaussian distributed zero mean random variables and they are statistically isotropic in the sky (see Sect. 2.5)

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell^{\rm th}. \tag{5.3}$$

Here asterisk indicates the complex conjugate. The underlying "theoretical" angular power spectrum of the CMB sky is $C_\ell^{\rm th}$. To simplify the notation we define $C_\ell \equiv$

$B_\ell^2 C_\ell^{\rm th}$ and use this definition in Chapter 5 and in Appendices A and B. Because we assume a perfect knowledge of the beam, the information contents of $C_\ell^{\rm th}$ and C_ℓ are the same.

The beam response B_ℓ^2 is a smoothing function, whose value decreases with increasing ℓ. We assume that B_ℓ^2 makes C_ℓ band-limited: C_ℓ will contain insignificant power beyond some $\ell_{\rm max}$.

Using Eqs. (5.2) and (5.3) the covariance matrix of the signal is

$$S_{ij} = \langle s_i s_j \rangle = \langle s_i s_j^* \rangle = \sum_\ell C_\ell P_{ij}^\ell. \tag{5.4}$$

The matrix P_{ij}^ℓ is defined as ([85])

$$P_{ij}^\ell = \sum_{m=-\ell}^{\ell} Y_{\ell m}(\mathbf{n}_i) Y_{\ell m}^*(\mathbf{n}_j). \tag{5.5}$$

It is determined by the pixelization of the sky and it is symmetric: $P_{ji}^\ell = P_{ij}^\ell$.

In a CMB experiment the ℓ-modes 0 and 1 of the CMB anisotropy signal are buried under the large monopole and non-cosmological dipole (arising from the motion of the observer relative to the CMB radiation field, see Sect. 3.3.1) and cannot be distinguished from them. We will not consider the monopole and dipole signals here and we assume that they do not appear in our maps. To model this we set $C_\ell = 0$ for $\ell = 0$ and $\ell = 1$. Unless otherwise noted the ℓ-sums of this chapter go from 0 to ∞.

We assume that the instrument noise in the map is Gaussian distributed, zero mean and its covariance is $N_{ij} \equiv \langle n_i n_j \rangle = \langle n_i n_j^* \rangle$.

In the case of partial sky coverage we will discard the unobserved pixels of the map and assume that the map vector $\mathbf{x}$ contains the observed pixels only. In this case the length of the map vector will be reduced from $N_{\rm pix}$ and the sizes of the covariance matrices will be correspondingly smaller.

We can now state the problem of the estimation of the CMB angular power spectrum: Given the map $\mathbf{x}$ we need to find an unbiased spectrum estimate $\widehat{C}_\ell^{\rm E}$ that minimizes the variance of the estimation error $\Delta\widehat{C}_\ell^{\rm E} = \widehat{C}_\ell^{\rm E} - C_\ell$. For an unbiased estimator $\langle \widehat{C}_\ell^{\rm E} \rangle = C_\ell$.

5.2 Maximum Likelihood Estimator

Under the assumptions of a Gaussian distribution and zero mean the probability density function of the map $\mathbf{x}$ is

$$\mathrm{p}(\mathbf{x}) = (2\pi)^{-\mathrm{N_{pix}}/2} \exp\left(-\frac{1}{2} (\mathbf{x}^{\rm T} \mathbf{C}^{-1} \mathbf{x} + \mathrm{Tr}[\ln \mathbf{C}]) \right). \tag{5.6}$$

The log-likelihood function $\mathcal{L}$ is the logarithm of the inverted probability density function. Dropping out an irrelevant additive constant we obtain

$$\mathcal{L} = \frac{1}{2}\mathrm{Tr}[\ln \mathbf{C} + \mathbf{C}^{-1}\mathbf{D}]. \tag{5.7}$$

Here $\mathrm{Tr}[\cdot]$ is the trace of a matrix and $\mathbf{D} = \mathbf{x}\mathbf{x}^{\mathrm{T}}$. The *maximum likelihood estimator* $\widehat{C}_\ell^{\mathrm{ML}}$ of the angular power spectrum C_ℓ is solved from the set of equations $\mathcal{L}_{,\ell} = 0$, where $\mathcal{L}_{,\ell} \equiv \frac{\partial \mathcal{L}}{\partial C_\ell}$. Note that $\mathbf{C}$ is a function of C_ℓ.

Under the assumption of Gaussian distributed map $\mathbf{x}$, the maximum likelihood estimator can be unbiased ($\langle \widehat{C}_\ell^{\mathrm{ML}} \rangle = C_\ell$) and, from all practical estimators, it typically has the smallest variance of the estimation error $\Delta\widehat{C}_\ell^{\mathrm{ML}} = \widehat{C}_\ell^{\mathrm{ML}} - C_\ell$ ([24], [89]).

The variance of the estimation error is related to the *Fisher information matrix*. That matrix is defined as an expectation value of the second derivative of the log-likelihood function ([24], [89])

$$F_{\ell\ell'} \equiv \langle \mathcal{L}_{,\ell\ell'} \rangle. \tag{5.8}$$

A formula for the Fisher information matrix is given in [89]

$$F_{\ell\ell'} = \frac{1}{2}\mathrm{Tr}[\mathbf{C}^{-1}\mathbf{C}_{,\ell}\mathbf{C}^{-1}\mathbf{C}_{,\ell'}] = \frac{1}{2}\mathrm{Tr}[\mathbf{C}^{-1}\mathbf{P}^{\ell}\mathbf{C}^{-1}\mathbf{P}^{\ell'}]. \tag{5.9}$$

The latter form is obtained after inserting the derivative of Eq. (5.4). The matrix is symmetric: $F_{\ell\ell'} = F_{\ell'\ell}$.

The significance of the Fisher information matrix is that it gives the minimum variance of the estimation error $\Delta\widehat{C}_\ell = \widehat{C}_\ell - C_\ell$ that we can obtain, regardless of which unbiased method we are using in the estimation. The lower limit of the variance is ([24], [89])

$$\langle (\Delta\widehat{C}_\ell)^2 \rangle \geq (\mathbf{F}^{-1})_{\ell\ell}. \tag{5.10}$$

This is called the Cramér-Rao bound ([89]). Generally, if an estimator exists that is unbiased and whose error variance equals the Cramér-Rao bound, that estimator will be the maximum-likelihood estimator. However, a maximum-likelihood estimator is not necessarily unbiased or will saturate to the Cramér-Rao bound ([90]).

A number of authors have considered the maximum likelihood power spectrum estimators in CMB data analysis applications (e.g. [58], [84] - [88]). The set of equations $\mathcal{L}_{,\ell} = 0$ for the power spectrum estimate is non-linear and e.g. the Newton-Raphson iterative method ([53]) to solve $\widehat{C}_\ell^{\mathrm{ML}}$ has been applied ([58], [86], [87]). In this method the result of iteration n ($\widehat{C}_\ell^{\mathrm{ML(n)}}$) is obtained from the result of the previous iteration

$$\widehat{C}_\ell^{\mathrm{ML(n)}} = \widehat{C}_\ell^{\mathrm{ML(n-1)}} - \sum_{\ell'} (\mathbf{F}^{-1})_{\ell\ell'} \mathcal{L}_{,\ell'}. \tag{5.11}$$

The derivative of $\mathcal{L}$ and $\mathbf{F}^{-1}$ are evaluated at $\widehat{C}_\ell^{\mathrm{ML(n-1)}}$.

The challenges to solve the maximum likelihood power spectrum estimator are twofold. Firstly, the initial guess $\widehat{C}_\ell^{\mathrm{ML}(0)}$ should fall close enough to the correct value for the iteration to converge towards the global minimum ([58]). Secondly, the evaluations of the inverse Fisher matrix and the derivative $\mathcal{L}_{,\ell}$ require inversions of several matrices with size $N_{\mathrm{pix}} \times N_{\mathrm{pix}}$. This becomes a computationally demanding job in PLANCK-like experiments, where $N_{\mathrm{pix}} = 10^6 \ldots 10^8$.

5.3 Quadratic Estimator

To find alternative estimators to the maximum likelihood power spectrum estimator and to understand their relation to it, we shall next consider the *quadratic estimator*, that is a quadratic function of the map pixels. We will follow the approach introduced by Tegmark ([85]). We are considering here CMB temperature anisotropy observations only, but Tegmark's formalism can be extended to the power spectrum estimation of the polarized CMB observations ([91]).

The quadratic estimator is defined as ([85])

$$\widetilde{C}_\ell = \mathbf{x}^{\mathrm{T}} \mathbf{E}^\ell \mathbf{x}. \tag{5.12}$$

Here $\mathbf{E}^\ell$ is an $N_{\mathrm{pix}} \times N_{\mathrm{pix}}$ symmetric matrix ($E^\ell_{ji} = E^\ell_{ij}$). It is undetermined at this time. The expectation value of the estimator is

$$\langle \widetilde{C}_\ell \rangle = \langle \mathbf{x}^{\mathrm{T}} \mathbf{E}^\ell \mathbf{x} \rangle = \mathrm{Tr}[\langle \mathbf{x}^{\mathrm{T}} \mathbf{E}^\ell \mathbf{x} \rangle] = \mathrm{Tr}[\mathbf{E}^\ell \langle \mathbf{x}\mathbf{x}^{\mathrm{T}} \rangle]. \tag{5.13}$$

Replacing $\langle \mathbf{x}\mathbf{x}^{\mathrm{T}} \rangle$ with $\sum_{\ell'} C_{\ell'} \mathbf{P}^{\ell'} + \mathbf{N}$ (see Sect. 5.1) we obtain for the expectation value

$$\langle \widetilde{C}_\ell \rangle = \sum_{\ell'} \mathrm{Tr}[\mathbf{E}^\ell \mathbf{P}^{\ell'}] \mathrm{C}_{\ell'} + \mathrm{Tr}[\mathbf{E}^\ell \mathbf{N}]. \tag{5.14}$$

The quantity $\langle \widetilde{N}_\ell \rangle \equiv \mathrm{Tr}[\mathbf{E}^\ell \mathbf{N}]$ is the *noise bias*. We define a matrix $M_{\ell\ell'} \equiv \mathrm{Tr}[\mathbf{E}^\ell \mathbf{P}^{\ell'}]$, which is the *mode coupling matrix* (kernel matrix). It describes the coupling between the ℓ-modes of the estimator due to e.g. the geometry of the sky coverage (cf. [93]). It should be noted that it describes the mode coupling accurately for the ensemble means, not for a single estimator. Inserting the symbols of the mode coupling matrix and noise bias, we have

$$\langle \widetilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell\ell'} C_{\ell'} + \langle \widetilde{N}_\ell \rangle. \tag{5.15}$$

Arranging the power spectrum coefficients (C_ℓ, $\widetilde{C}_\ell$, $\widetilde{N}_\ell$) into column vectors, the equation can be expressed as

$$\langle \widetilde{\mathbf{c}} \rangle = \mathbf{M}\mathbf{c} + \langle \widetilde{\mathbf{n}} \rangle. \tag{5.16}$$

This equation suggests an estimator $\widehat{\mathbf{c}}$ for the spectrum $\mathbf{c}$ of the CMB sky

$$\widehat{\mathbf{c}} = \mathbf{M}^{-1} (\widetilde{\mathbf{c}} - \langle \widetilde{\mathbf{n}} \rangle). \tag{5.17}$$

We will discuss later (in Sect. 5.5.4) how an estimate for the noise bias can be obtained. Assuming that we have found a correct noise bias estimate, $\widehat{\mathbf{c}}$ is an unbiased estimator of $\mathbf{c}$.

Eq. (5.17) makes sense if the mode coupling matrix is numerically non-singular[1]. PLANCK-like scanning strategies provide observations over the full or nearly full sky (see Sect. 3.2). To reduce contamination of the CMB signal from galactic emissions, we may have to cut out the galactic region from the maps prior to the CMB power spectrum estimation. In the largest galactic cuts, that we consider in this study, the galactic region at ± 20 deg from the galactic plane ($|b| \leq 20$ deg) is cut out from the map (removing $\sim 35\%$ of the sky). As a rule of thumb, the inverse of the mode coupling matrix exists if the two-point correlation function can be determined at all angular scales from the data within the uncut part of the map ([94]). A sky with ± 20 deg galactic cut meets this requirement. We demonstrate in Sect. 5.5.2 that the mode coupling matrix of the ± 20 deg galactic cut is numerically non-singular and its inverse exists.

The pixels of the map $\mathbf{x}$, that were not observed by the scanning or that were intentionally cut out from the map, are called unobserved pixels in this study. If the number of the unobserved pixels is increased, the mode coupling matrix will become numerically singular at some point. In this case the matrix may become well-behaving again, if several ℓ-modes are summed to an ℓ-bin ([93]). Another possibility to circumvent the singularity problem is to use the quadratic estimator $\widetilde{C}_\ell$ in the estimation of the cosmological parameters (instead of the actual power spectrum estimates $\widehat{C}_\ell$) ([94]). The estimation of the cosmological parameters is beyond the scope of this thesis and we will not discuss this issue any further.

Because the map $\mathbf{x}$ was assumed zero mean and Gaussian distributed, the covariance matrix $\widetilde{\mathbf{V}}$ of $\widetilde{\mathbf{c}}$ can be expressed as ([85])

$$\widetilde{V}_{\ell\ell'} \equiv \langle \widetilde{C}_\ell \widetilde{C}_{\ell'} \rangle - \langle \widetilde{C}_\ell \rangle \langle \widetilde{C}_{\ell'} \rangle = 2\mathrm{Tr}[\mathbf{C}\mathbf{E}^\ell \mathbf{C}\mathbf{E}^{\ell'}]. \tag{5.18}$$

Using Eq. (5.17) the covariance matrix of the estimator $\widehat{\mathbf{c}}$ becomes ([94])

$$\widehat{\mathbf{V}} = \mathbf{M}^{-1}\widetilde{\mathbf{V}}(\mathbf{M}^{-1})^{\mathrm{T}}. \tag{5.19}$$

The matrix $\mathbf{E}^\ell$ is still undetermined. Tegmark shows how an optimal matrix $\mathbf{E}^\ell$ is obtained that minimizes $\widetilde{V}_{\ell\ell}$ subject to a condition $M_{\ell\ell} = 1$ ([85]). That matrix is

$$\mathbf{E}^\ell = \frac{1}{2F_{\ell\ell}} \mathbf{C}^{-1}\mathbf{P}^\ell \mathbf{C}^{-1}. \tag{5.20}$$

The factor $F_{\ell\ell}$ is the diagonal element of the Fisher information matrix. This $\mathbf{E}^\ell$ leads to $\mathbf{M} = (\mathbf{F}^{\mathrm{D}})^{-1}\mathbf{F}$ and $\widetilde{\mathbf{V}} = (\mathbf{F}^{\mathrm{D}})^{-1}\mathbf{F}(\mathbf{F}^{\mathrm{D}})^{-1}$, where $\mathbf{F}$ is the Fisher information

[1] A square matrix is numerically non-singular if its condition number (absolute value of the ratio of its largest and smallest eigenvalues) is smaller than the inverse of the floating point precision of the computer used (condition number $\lesssim 10^6$ for single precision and $\lesssim 10^{12}$ for double precision, see [53]).

matrix (Eq. (5.9)) and $\mathbf{F}^{\mathrm{D}}$ is its diagonal version ($F_{\ell\ell}$'s in the diagonals and zeros in the off-diagonals). Inserting $\mathbf{M}$ and $\widetilde{\mathbf{V}}$ in to Eq. (5.19) and using the fact that $\mathbf{F}$ and $\mathbf{F}^{\mathrm{D}}$ are symmetric, we obtain for the covariance matrix of the estimator $\widehat{\mathbf{c}}$

$$\widehat{\mathbf{V}} = \mathbf{F}^{-1}. \tag{5.21}$$

So, with this choice of $\mathbf{E}^{\ell}$ (Eq. (5.20)), $\widehat{\mathbf{c}}$ is really an optimal estimator in the sense that its error bars (std of the estimation error) equal the Cramér-Rao bound and are thus the lowest attainable. Because the error bars of the estimator $\widehat{\mathbf{c}}$ equal the error bars of the maximum likelihood estimator, the estimator $\widehat{\mathbf{c}}$ is statistically equivalent to the maximum likelihood estimator. It is worth noting that the $\frac{1}{2F_{\ell\ell}}$ factor in $\mathbf{E}^{\ell}$ is a normalization that we could have chosen differently. The covariance matrix of $\widehat{\mathbf{c}}$ is not affected if $\mathbf{E}^{\ell}$ is multiplied by a constant.

We need to note that the covariance matrix of Eq. (5.21) is obtained only if we use the true noise covariance matrix $\mathbf{N}$ and the true CMB angular power spectrum C_ℓ when we determine $\mathbf{C}$. In a real experiment we have only estimates (or guesses) of these quantities available. Therefore the estimation error of the quadrature estimator will typically be larger than the Cramér-Rao bound.

The CMB temperature map is typically signal dominated at large scales (low ℓ) and noise dominated at small scales (high ℓ). In these extreme situations the map covariance $\mathbf{C}$ can be approximated by either $\mathbf{S}$ (signal dominance) or by $\mathbf{N}$ (noise dominance) only. The formulas for $\mathbf{E}^{\ell}$ (Eq. (5.20)) in these two cases are derived in Appendix A. The result can be given in a single equation (see Eqs. (A.16) and (A.19))

$$E^{\ell}_{ij} = \frac{\Omega_p^2}{2\ell+1} w_i w_j P^{\ell}_{ij}. \tag{5.22}$$

We assume HEALPix pixelization ([52]) where every pixel of the map $\mathbf{x}$ has the same area $\Omega_p = 4\pi/N_{\mathrm{pix}}$. The quantities w_i are called pixel weights. They are discussed in more detail in Sect. 5.4. The equation is valid for every pixel of the sky provided that we select $w_i = 0$ for the unobserved pixels.

For the noise dominated map, $\mathbf{E}^{\ell}$ of Eq. (5.22) gives a power spectrum estimator that is statistically equivalent to the maximum likelihood estimator if we assume that the map noise covariance is diagonal $N_{ij} = \sigma_i^2 \delta_{ij}$ and if we set $w_i = 1/\sigma_i^2$ (inverse noise variance weighting). Under these assumptions this estimator is statistically equivalent to the maximum likelihood estimator at all sky coverages (full or partial).

With this model the best we can do for the signal dominated map is to use uniform pixel weights (e.g. $w_i = 1$). In that case the estimator is statistically equivalent to the maximum likelihood estimator at full sky, but only an approximation of it in the partial sky coverage. For narrow galactic cuts (like $\pm$ 20 deg) its performance is, however, close to the maximum likelihood performance ([94]).

The particular $\mathbf{E}^{\ell}$ given in Eq. (5.22) defines a set of estimators called *pseudo-C_ℓ estimators* with pixel weights w_i as user selectable parameters ([94]). We shall discuss pseudo-C_ℓ estimators more in the next section.

Finally, it is of interest to derive the mode coupling matrix ($\mathbf{M}$) corresponding to $\mathbf{E}^\ell$ of Eq. (5.22). Inserting $\mathbf{E}^\ell$ to the definition of $\mathbf{M}$ (Eq. (5.14)) we obtain

$$M_{\ell\ell'} = \frac{\Omega_p^2}{2\ell+1} \sum_{i,j} w_i w_j P_{ij}^\ell P_{ji}^{\ell'} \tag{5.23}$$

and from that we finally have

$$M_{\ell\ell'} = \frac{1}{2\ell+1} \sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'} |\Omega_p \sum_i w_i Y_{\ell m}(\mathbf{n}_i) Y^*_{\ell' m'}(\mathbf{n}_i)|^2. \tag{5.24}$$

The latter form is obtained after replacing the matrix $\mathbf{P}^\ell$ with its definition from Eq. (5.5). The i and j sums are over the entire set of pixels that cover the sky (w_i = 0 for the unobserved pixels). This matrix is nearly identical to the mode coupling matrix introduced in the MASTER (Monte Carlo Apodized Spherical Transform EstimatoR) power spectrum estimation method ([93]). For the full sky and with unit pixel weights $M_{\ell\ell'}$ is close to a unit matrix (see Sect. 5.5.2).

5.4 Pseudo-C_ℓ Estimator

The spherical harmonic transform of the observed map $\mathbf{x}$ is defined as ([94])

$$\widetilde{a}_{\ell m} = \Omega_p \sum_{k=0}^{N_{\rm pix}-1} x_k w_k Y_{\ell m}(\mathbf{n}_k). \tag{5.25}$$

We can construct the angular power spectrum of the $\widetilde{a}_{\ell m}$ coefficients

$$\widetilde{C}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |\widetilde{a}_{\ell m}|^2 = \sum_{i,j=0}^{N_{\rm pix}-1} x_i \left(\frac{\Omega_p^2}{2\ell+1} \sum_{m=-\ell}^{\ell} w_i w_j Y_{\ell m}(\mathbf{n}_i) Y^*_{\ell m}(\mathbf{n}_j) \right) x_j. \tag{5.26}$$

Using the definition of the matrix $\mathbf{P}^\ell$ (Eq. (5.5)) it is easy to see that the quantity inside the parenthesis is identical to the matrix $\mathbf{E}^\ell$ of Eq. (5.22). Therefore $\widetilde{C}_\ell$ is a pseudo-C_ℓ estimator (see Sect. 5.3). We remind that the pseudo-C_ℓ estimator belongs to a larger set of quadratic estimators (cf. Eq. (5.12)).

The quantity $\widetilde{C}_\ell$ of Eq . (5.26) is often called simply the *pseudo spectrum*. We use the term pseudo spectrum for $\widetilde{C}_\ell$ in this thesis. The estimate $\widehat{C}_\ell$ of C_ℓ is obtained from the pseudo spectrum by subtracting the noise bias estimate from it and deconvolving the difference with the mode coupling matrix (see Eq. (5.17)).

The benefit of the $\widehat{C}_\ell$ estimator (compared to the maximum likelihood estimator) is its simple implementation. The discrete spherical harmonic transform on a map with $N_{\rm pix}$ pixels (to produce $\widetilde{C}_\ell$) demands far less computing resources and time

than the inversions of $N_{\rm pix} \times N_{\rm pix}$ matrices required by the maximum likelihood estimator.

Several authors have considered pseudo-C_ℓ estimators in the CMB experiments. A discussion of the statistical properties of the pseudo-C_ℓ estimators and their use in estimating the full sky CMB spectrum from the partial sky observations is given in [95]. The MASTER method developed by Hivon et al. is described in [93]. MASTER has been applied in e.g. the BOOMERANG data analysis ([66], [96]). Pseudo-C_ℓ estimator was applied in the WMAP data analysis too ([97]). Pseudo-C_ℓ estimators to recover the true shape of the power spectrum from the apodized sky using Gabor transforms were developed in [98] and [99]. The use of pseudo-C_ℓ estimators in estimating the CMB power spectrum from the GLS maps is described in [100].

Estimators that use the two-point correlation function in the sky and that can then be converted to give an estimate of the angular power spectrum have been proposed ([101], [102]). An extension of the two-point correlation function estimators to the analysis of CMB polarisation observations has been proposed too ([103]). The two-point correlation function and the angular power spectrum represent the same thing in different domains (two-point correlation function in the map domain and angular power spectrum in the spherical harmonic domain). Thus the pseudo-C_ℓ estimators and the estimators based on two-point correlation function are (mathematically) nearly equivalent.

It was shown in the previous section (Sect. 5.3) that with a proper selection of the pixel weights w_i and under some particular assumptions, the $\widehat{C}_\ell$ estimator is statistically equivalent to the maximum likelihood estimator. The fact that different weights are optimal at different signal and noise conditions (e.g. $w_i = 1$ for signal dominated maps and $w_i = 1/\sigma_i^2$ for noise dominated maps) suggests that we should produce several $\widehat{C}_\ell$ estimators, with different weights, from the same observed map and use different estimators at different ranges of ℓ.

For low ℓ, where CMB typically dominates, $w_i = 1$ can be used, although this estimator performs as a maximum likelihood estimator in the full sky only. In the cut sky it is an approximation of the maximum likelihood estimator. For high ℓ, where noise dominates, $w_i = 1/\sigma_i^2$ can be used, although it is optimal only for white noise. On the other hand, because the map noise is typically close to white noise at small scales, these are probably close to the actual optimal weights.

The intermediate range of ℓ is the most problematic case. For that range it is difficult to derive an analytic formula for the optimal weights ([94]). As an example, WMAP team used uniform weights at $\ell < 200$, inverse noise variance weights ($w_i = N_i$, where N_i is the number of hits in pixel i) at $\ell > 450$ and "heuristic" weights

$$w_i = \frac{1}{\frac{1}{\langle N \rangle} + \frac{1}{N_i}} \tag{5.27}$$

at $200 < \ell < 450$ ([97]). Here $\langle N \rangle$ is the mean number of hits over the observed pixels. When $N_i \gg \langle N \rangle$ ("signal dominated pixel") $w_i \approx \langle N \rangle$ (constant) and when

$N_i \ll \langle N \rangle$ ("noise dominated pixel") $w_i \approx N_i$.

Different power spectrum estimates at different ranges of ℓ lead to abrupt changes at the boundaries of the estimates. A smooth combined estimate $\widehat{C}_\ell^{\rm c}$ that is optimum at each range of ℓ can be obtained from the component estimates $\widehat{C}_\ell^{\alpha}$, when they are combined in a maximum likelihood sense. Here α indexes a component estimate (e.g. $\alpha = 1, 2, 3$ in WMAP). The combined estimate is solved by minimizing the likelihood ([94])

$$\chi^2 = \sum_{\ell,\ell'} \sum_{\alpha,\beta} \left(\widehat{C}_\ell^{\alpha} - \widehat{C}_\ell^{\rm c} \right) \mathcal{F}_{\ell\ell'}^{\alpha\beta} \left(\widehat{C}_{\ell'}^{\beta} - \widehat{C}_{\ell'}^{\rm c} \right), \tag{5.28}$$

where $\mathcal{F}_{\ell\ell'}^{\alpha\beta}$ is the inverse of the covariance matrix $\langle \Delta\widehat{C}_\ell^{\alpha} \Delta\widehat{C}_{\ell'}^{\beta} \rangle$. At a given ℓ the estimate $\widehat{C}_\ell^{\alpha}$, that has the smallest uncertainty, has the largest weight in the combined estimate $\widehat{C}_\ell^{\rm c}$. The auto and cross covariance matrices $\langle \Delta\widehat{C}_\ell^{\alpha} \Delta\widehat{C}_{\ell'}^{\beta} \rangle$ of the component estimates are the difficult part of this approach. Their determination may not be a trivial task in a real CMB experiment.

So far we have made no distinction whether the observed map $\mathbf{x}$ is a single detector map, a frequency map from the observations of multiple detectors or a CMB map after the component separation from many frequency maps. The power spectrum estimation methods described so far can be applied without modification to any of these maps.

An alternative $\widehat{C}_\ell$ estimator suitable for the observations of multiple detectors has been proposed and applied e.g. in WMAP data processing ([97], [104]). In this method the maps are made for each detector separately, their pseudo-$\widetilde{a}_{\ell m}$ coefficients are determined, angular cross-power spectrum for every detector pair is calculated and the combined pseudo spectrum is obtained as an inverse covariance weighted sum of the cross-spectra. Assuming that the instrument noise is uncorrelated between the detectors, the noise biases of the cross-spectra are zero and no noise bias estimation is required. The final spectrum estimate $\widehat{C}_\ell$ is obtained by deconvolving the combined pseudo spectrum with the mode coupling matrix. The error bars of this method are larger than the error bars of the $\widehat{C}_\ell$ estimate derived from the (multidetector) frequency map. For large number of detectors the difference is, however, small.

5.5 Power Spectrum Estimation and Destriping

The estimation of the power spectra from the maps made by destriping was studied in this thesis (paper **II** [2]). This section gives a summary of that study and a discussion on some of its topics.

5.5.1 Simulation Description

We considered two different underlying "theoretical" CMB skies whose angular power spectra were denoted by $C_\ell^{\rm th}$. They were ΛCDM (cosmological constant +

Cold Dark Matter) and OCDM (Open CDM) models. We show here results of ΛCDM only. The results of OCDM are available in paper **II** ([2]). The simulation and map-making parameters used in the study are given in Table 5.1.

We considered the originally planned LFI 100 GHz detectors. The nominal scanning strategy (see Sect. 3.2) was applied. The hit map of the applied scanning strategy is shown in Fig. 5.1 (upper panel). Due to nominal scanning strategy (satellite spin axis always in the ecliptic plane) small areas near the ecliptic poles remained without observations. Thus we call this case "nearly full sky". In addition, we considered a case called "galactic cut", where the galactic region ($|b| < 20$ deg) was cut out from the maps. The hit map of this case is shown in Fig. 5.1 as well.

The TODs for the map-making were generated using computer simulations. A TOD was a sum of CMB and instrument noise. The CMB component was scanned from a high-resolution *input map*. The TOD samples were picked from the input map pixels where the pointings of the scan path fell. The resolution of the input map was $N_{\rm side} = 1024$ and it was generated from a realization ($a_{\ell m}$) of the theoretical spectrum $C_\ell^{\rm th}$. The temperature (d_i) of the pixel i of the input map was given as

$$d_i = \sum_{\ell m} a_{\ell m} B_\ell D_\ell(1024) Y_{\ell m}(\mathbf{n}_i). \tag{5.29}$$

Here B_ℓ is the response of the symmetric beam and $D_\ell(1024)$ is the HEALPix pixel window function for $N_{\rm side} = 1024$ ([105], see also Appendix B of this thesis). Due to the pixel window function, d_i is the beam smoothed CMB temperature field that is averaged over the pixel area. The unit vector pointing to the center of the pixel is $\mathbf{n}_i$. The noise component of the TOD was a sum of white and $1/f$ noise. The white noise std was the single detector std divided by $\sqrt{24}$. The TODs we generated represent the case of 24 LFI 100 GHz detectors if we assume that their noises are uncorrelated and they have identical scan paths. For Monte Carlo (MC) studies we generated a number of independent CMB-only, noise-only and CMB+noise TODs.

The input maps were made using the Synfast code of the HEALPix package. When running Synfast the user needs to choose between two operating modes. In the first mode the user gives Synfast the angular power spectrum, random number seed, beam parameters and map $N_{\rm side}$. Synfast uses its internal random number generator to create the $a_{\ell m}$ and outputs the map as in Eq. (5.29). Synfast sets the pixel window function according to $N_{\rm side}$ chosen by the user. In the second mode the user inputs $a_{\ell m}$ directly and Synfast calculates the pixel temperatures as $d_i = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\mathbf{n}_i)$. In this mode Synfast ignores the pixel window function and the beam. The first operating mode was used in this study when we generated the input maps ($C_\ell^{\rm th}$ was the input spectrum to Synfast).

We used a simple pointing matrix in destriping. It contained only ones and zeros with a single non-zero element in every row. The non-zero elements referred to the input map pixels where the detector had been pointing. The resolution of the output maps was $N_{\rm side} = 512$. The pseudo spectra $\widetilde{C}_\ell$ for $\ell \in [2, 1500]$ were calculated from these maps. We considered only unit pixel weights in this study ($w_i = 1$ for the

Table 5.1: Simulation parameters of the power spectrum estimation (paper **II** [2]). The maps were made using destriping. The CMB component of the TOD was scanned from a high resolution input map ($N_{\rm side} = 1024$). The CMB components of the detectors were identical, because we assumed identical detector scan paths. The TODs of the full set of 24 LFI 100 GHz detectors were modelled by one TOD whose white noise std was the single detector std divided by $\sqrt{24}$. The noise TODs were made using the Level S code ([43]). Destriping operated on co-added TODs, where 60 circles of a repointing period had been averaged to a ring.

Detector		LFI 100 GHz
Number of detectors		24
Mission time		7 months
Scanning (a)		Nominal
Beam (b)		Symmetric
$C_\ell^{\rm th}$		ΛCDM
		OCDM (Open CDM)
TOD - CMB		Scanned from a CMB input map
TOD - Noise	σ (c)	979.8 μK
	Knee freq.	0.1 Hz
	Minimum freq.	4×10^{-6} Hz
	Sampling freq.	108.3 Hz
	$1/f$ slope (α)	1
Output maps	$N_{\rm side}$	512
	Destriping	One minute uniform baselines
	Sky coverage (d)	98.5% / 64.6%
	Coordinates	Ecliptic

(a) Perfect repetition of the pointings of the 60 circles of a ring.
Identical detector scan paths were assumed.
(b) Symmetric Gaussian beam, FWHM = 10.0 arcmin.
(c) White noise std (in CMB temperature units):
Single detector std (4800 μK) divided by $\sqrt{24}$.
(d) Nearly full sky / $\pm$ 20 deg galactic cut.

observed pixels and $w_i = 0$ for the unobserved pixels). The impact of weights on the accuracy of the power spectrum estimates is discussed in Sect. 5.5.5. We further define $C_\ell^{\rm B}$, which is the angular power spectrum of the binned noiseless map. That map was obtained by binning the samples of the noiseless CMB-only TOD to the pixels of the $N_{\rm side} = 512$ map.

The ensemble mean of the CMB pseudo spectrum depends on $C_\ell^{\rm th}$. Hivon et al.

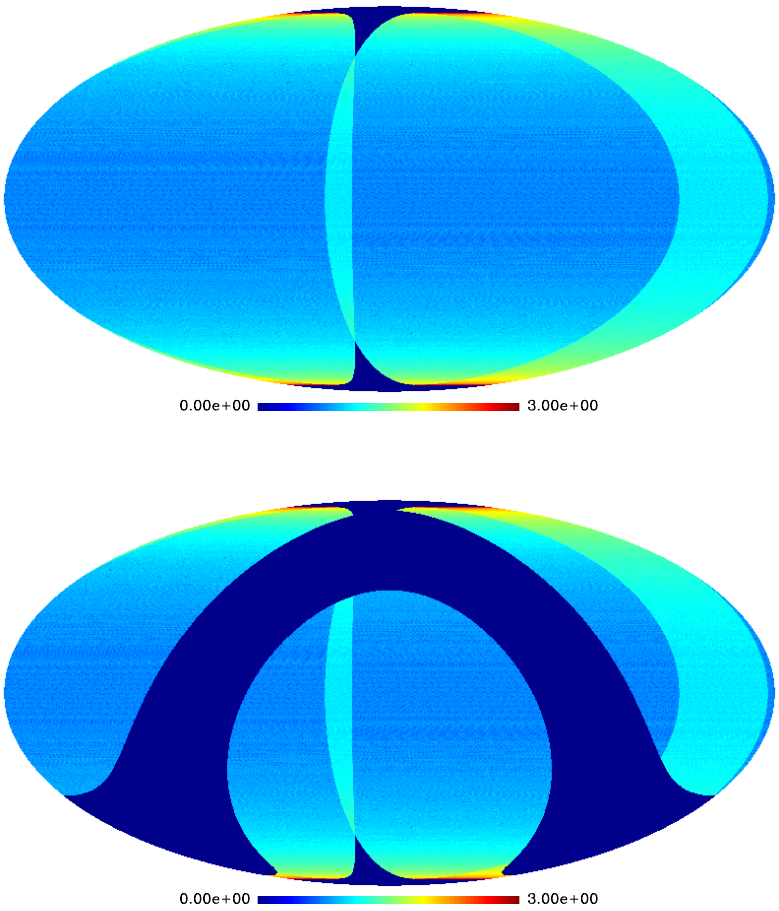

Figure 5.1: Number of hits per pixel for the nearly full sky (**top**) and for the galactic cut (**bottom**). Note that the scale is $\log_{10}(\mathrm{n_{hit}})$, where $\mathrm{n_{hit}}$ is the number of hits in a pixel. The maps are from paper **II** ([2]).

express this relation in MASTER as ([93])

$$\langle \widetilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell\ell'} F_{\ell'} B^2_{\ell'} D^2_{\ell'} C^{\mathrm{th}}_{\ell'} + \langle \widetilde{N}_\ell \rangle. \tag{5.30}$$

The matrix $M_{\ell\ell'}$ is the mode coupling matrix determined by the applied sky cut and pixel weighting ([93], see also Eq. (5.24)). The smoothing of the symmetric beam is modelled by B^2_ℓ. Output map pixelization introduces additional smoothing, which is represented by the pixel window factor D^2_ℓ. The filter function F_ℓ represents a possible distorting effect of map-making and the noise bias $\langle \widetilde{N}_\ell \rangle$ the remaining noise. We will discuss these factors in the subsequent sections of this thesis. In this study we have assumed a symmetric beam but a brief theoretical discussion on the

deconvolution of the asymmetric beams is given in Appendix B.

5.5.2 Mode Coupling Matrix

The mode coupling matrix ($M_{\ell\ell'}$) used in the MASTER equation (Eq. (5.30)) is given in Eq. (5.24). That formula is not very practical for the calculation of the matrix elements. An approximate formula that is better suited for this calculation was derived in [93].

$$M_{\ell\ell'} = \frac{2\ell'+1}{4\pi} \sum_{\ell''=0}^{2\ell_{\max}} (2\ell''+1) W_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2, \tag{5.31}$$

where $\begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix}$ is the Wigner $3-j$ symbol. The angular power spectrum of the sky coverage map is W_ℓ. The sky coverage map contains the weights w_i in its pixels. This equation provides $M_{\ell\ell'}$ up to $\ell, \ell' = \ell_{\max}$. Due to the properties of the Wigner $3-j$ symbols, W_ℓ needs to be evaluated only up to $2\ell_{\max}$. Eq. (5.31) for $M_{\ell\ell'}$ can be derived from Eq. (5.24) if the sum over the pixels (i-sum) is approximated by an integral over the celestial sphere. Eq. (5.31) shows that the mode coupling matrix $M_{\ell\ell'}$ is solely determined by the sky coverage map and that $M_{\ell\ell'} \geq 0$ for every element of the matrix.

In this study we used unit weights for the observed pixels. The power spectra of the coverage maps of the nearly full sky and galactic cut are shown in Fig. 5.2. For comparison, we show the power spectrum of the full sky too ($w_i = 1$ in every pixel).

It could be expected that the angular power spectrum of the full sky coverage map would contain a monopole and zero for the rest of the multipoles. This is not, however, the case, because the spherical harmonics are not a complete set of orthogonal functions in the discrete sphere. The HEALPix pixelization has a symmetry with respect to the ecliptic plane which exhibits itself in the alternating multipoles. Every second multipole has a large value and every second has a small value (see the full sky spectrum of Fig. 5.2).

Due to scanning the nearly full sky coverage map has a strong symmetry with respect to the ecliptic plane. The galactic cut creates a symmetry with respect to the galactic plane. These symmetries explain the alternating multipoles of these two cases.

All spectra show an increase of power at high ℓ. This is a consequence of the incompleteness of the spherical harmonics in the discrete sphere.

Some rows of the mode coupling matrices are shown in Fig. 5.3. As in the angular power spectra of the coverage maps, the elements of a row alternate between large and small values. We can see that the mode coupling matrices are diagonally dominant in all considered sky cuts. Their diagonal elements are nearly constant with mean values 0.97 (nearly full sky) and 0.46 (galactic cut). The diagonal elements of the full sky matrix are close to 1. The ℓ-to-ℓ variation of the diagonal elements is

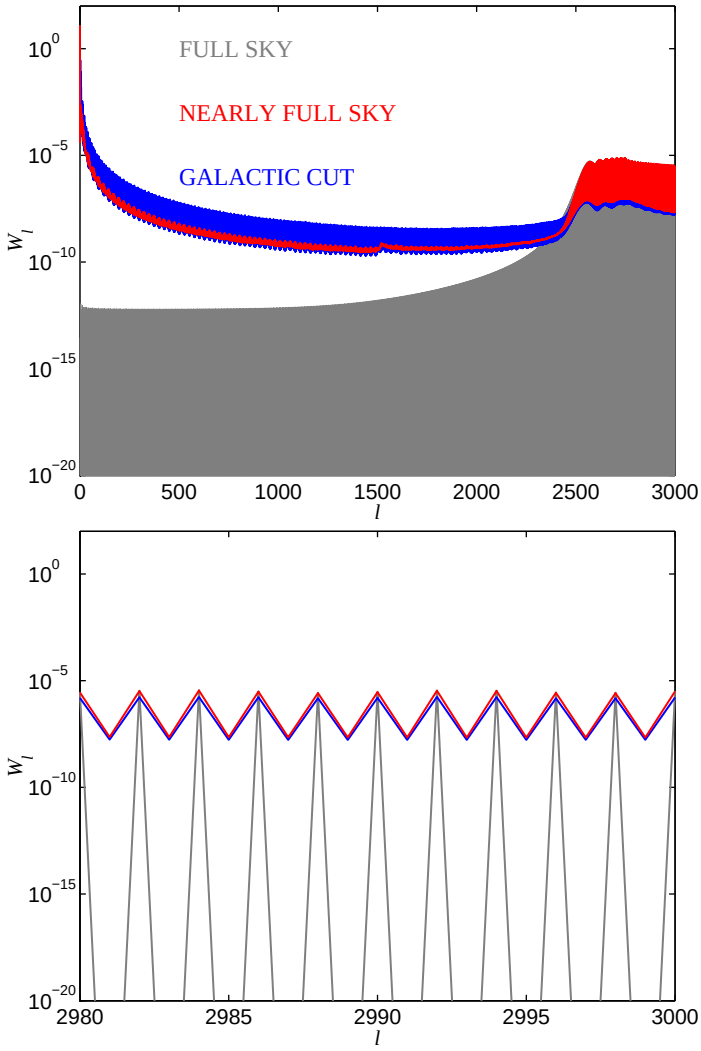

Figure 5.2: Angular power spectra W_ℓ of the sky coverage maps ($w_i = 1$ for the observed pixels and $w_i = 0$ for the unobserved pixels). The corresponding hit maps are shown in Fig. 5.1. For comparison the power spectrum of the full sky ($w_i = 1$ in every pixel) is shown too. The resolutions of the coverage maps were $N_{\mathrm{side}} = 512$. The monopole (not clearly visible) is $4\pi f_{\mathrm{sky}}^2$, where f_{sky} is the fraction of the sky covered with observations. **Top:** Full spectra. **Bottom:** Zoom to a narrow range of ℓ.

largest in the galactic cut being $\sim$2% there. There is no simple relation between f_{sky} and the magnitude of the diagonal elements, but their magnitudes decrease with decreasing f_{sky}. The condition numbers of the mode coupling matrices are $\sim$1.09 (nearly full sky) and $\sim$1.7 (galactic cut) showing that they are well-behaving and can be inverted for the estimation of C_ℓ^{th} (cf. Eq. (5.17)).

The non-zero off-diagonal elements of $M_{\ell\ell'}$ cause coupling between the ℓ-modes

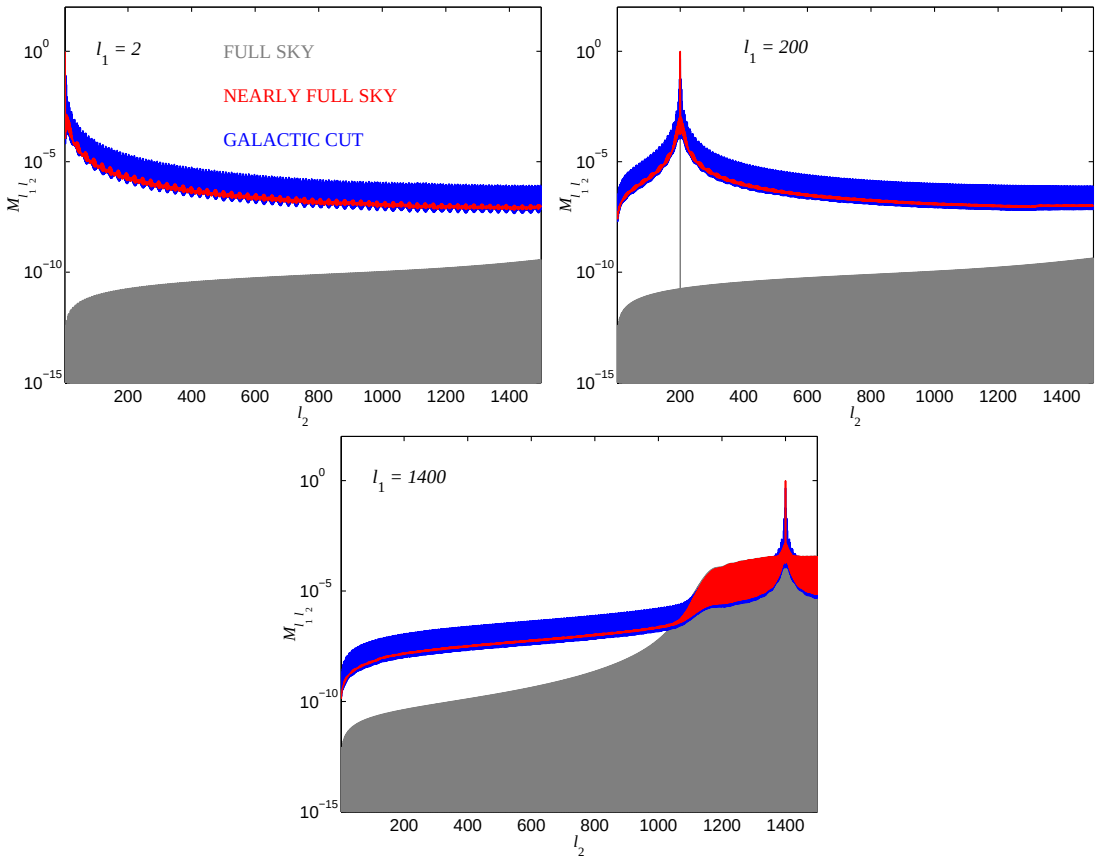

Figure 5.3: Three rows of the mode coupling matrices corresponding to the full sky, nearly full sky and galactic cut coverage maps. The element values of a row alternate between large and small values. At the used x-axis scale this alternation shows up as large colored areas instead of stand out curves.

of the power spectrum of the cut sky. The mean power spectrum of the cut sky depends on the mean power spectrum of the full sky as

$$\langle C_\ell^{\rm cutsky} \rangle = M_{\ell\ell} \langle C_\ell^{\rm fullsky} \rangle + \sum_{\ell' \neq \ell} M_{\ell\ell'} \langle C_{\ell'}^{\rm fullsky} \rangle. \tag{5.32}$$

At high ℓ, where CMB has low power, the second term on the right hand side may become significant with respect to the first term. This leads to a high-ℓ excess power. At low ℓ, where CMB has a lot of power, the first term typically dominates over the second term. Because tiny non-zero off-diagonals exist for the pixelized full sky too (due to incompleteness of the spherical harmonics), some small high-ℓ excess power will occur in this case as well. The mode coupling can be removed (from the ensemble means) if the spectrum of the cut sky is deconvolved with the mode coupling matrix.

5.5.3 Pixel Window and Pointing Distribution Effects

The relation (Eq. (5.30)) between the ensemble mean of the CMB pseudo spectra $\langle \widetilde{C}_\ell \rangle$ and C_ℓ^{th} suggests a similar relation between the ensemble mean of the spectra of the binned noiseless maps and C_ℓ^{th}

$$\langle C_\ell^{\mathrm{B}} \rangle = \sum_{\ell'} M_{\ell\ell'} B_{\ell'}^2 D_{\ell'}^2 C_{\ell'}^{\mathrm{th}}. \tag{5.33}$$

The relation

$$\langle \widetilde{C}_\ell \rangle = F_\ell \langle C_\ell^{\mathrm{B}} \rangle + \langle \widetilde{N}_\ell \rangle. \tag{5.34}$$

remains between the mean pseudo spectrum and the mean spectrum of the binned noiseless map (paper **III** [3]). The filter function F_ℓ accounts here for the map-making errors only. Their origin is the pixelization noise and F_ℓ reflects the signal component of the reconstruction error that we found in the map domain (see Sect. 4.5.1). We expect F_ℓ to be close to one.

Let us examine the relation between $\langle C_\ell^{\mathrm{B}} \rangle$ and C_ℓ^{th} in more detail. (The following discussion is presented in paper **III** ([3]) in a slightly different form.) In our study (paper **II** [2]) the pointings of the observations coincided with the pixels of the input map. The binned noiseless map with its larger pixel size was produced from these observations. The input map had $N_{\mathrm{side}} = 1024$, whereas the binned noiseless map had $N_{\mathrm{side}} = 512$ so that each pixel of the binned noiseless map includes four subpixels that are the pixels of the input map. This discussion goes beyond the simple signal model of the observed map that was introduced in Sect. 5.1. The signal in the binned noiseless map is influenced by the distribution of the hits in the pixels and therefore cannot be accurately described by the model of Eq. (5.2) given in Sect. 5.1.

We consider the samples d_i (i indexes the sample) of the CMB-only TOD that fall in a pixel k of the binned noiseless map. The number of hits in that pixel is N_k. The temperature d_i was given in Eq. (5.29). We repeat it here

$$d_i = \sum_{\ell m} a_{\ell m} B_\ell D_\ell(1024) Y_{\ell m}(\mathbf{n}_i). \tag{5.35}$$

The unit vector $\mathbf{n}_i$ points to the center of the input map pixel the detector is pointing at. The following discussion applies to a real CMB experiment too, if we discard the pixel window function $D_\ell(1024)$ (use 1 instead of $D_\ell(1024)$) and consider the vector $\mathbf{n}_i$ as the actual pointing of the center of the detector beam.

The temperature of the pixel k of the binned noiseless map is

$$T_k^{\mathrm{B}} = \frac{1}{N_k} \sum_{i \in k} d_i = \sum_{\ell m} a_{\ell m} B_\ell D_\ell(1024) \frac{1}{N_k} \sum_{i \in k} Y_{\ell m}(\mathbf{n}_i), \tag{5.36}$$

where $i \in k$ refers to those TOD samples that hit the pixel k.

The expansion coefficients of the binned noiseless map are obtained as (we assume unit weights for the observed pixels)

$$a^{\rm B}_{\ell m} = \Omega_p \sum_k T^{\rm B}_k Y^*_{\ell m}(\mathbf{q}_k). \tag{5.37}$$

The unit vector pointing to the center of the pixel k (of the binned noiseless map) is $\mathbf{q}_k$ and the k-sum is taken over the observed pixels.

After inserting $T^{\rm B}_k$ from Eq. (5.36) we obtain for the $a_{\ell m}$ of the binned noiseless map

$$a^{\rm B}_{\ell m} = \sum_{\ell' m'} a_{\ell' m'} B_{\ell'} D_{\ell'}(1024) \Omega_p \sum_k \frac{1}{N_k} \sum_{i \in k} Y_{\ell' m'}(\mathbf{n}_i) Y^*_{\ell m}(\mathbf{q}_k). \tag{5.38}$$

This equation defines an $a_{\ell m}$ coupling matrix (paper **III** [3])

$$K^{\rm B}_{\ell m \ell' m'} \equiv \Omega_p \sum_k \frac{1}{N_k} \sum_{i \in k} Y_{\ell' m'}(\mathbf{n}_i) Y^*_{\ell m}(\mathbf{q}_k) \tag{5.39}$$

between the $a_{\ell m}$ of the binned noiseless map and the CMB sky. Properties of these coupling matrices are discussed in [92].

Using the statistical isotropy of the CMB sky (Eq. (5.3)) we obtain for the ensemble mean of the angular power spectrum of the binned noiseless map

$$\langle C^{\rm B}_\ell \rangle = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle |a^{\rm B}_{\ell m}|^2 \rangle = \sum_{\ell'} M^{\rm B}_{\ell \ell'} B^2_{\ell'} D^2_{\ell'}(1024) C^{\rm th}_{\ell'}, \tag{5.40}$$

where $M^{\rm B}_{\ell \ell'}$ is the mode coupling matrix of the binned noiseless map (paper **III** [3])

$$M^{\rm B}_{\ell \ell'} = \frac{1}{2\ell+1} \sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'} |K^{\rm B}_{\ell m \ell' m'}|^2. \tag{5.41}$$

This mode coupling matrix describes the effects of sky coverage (distribution of the observed and unobserved pixels in the sky) and sampling of the observed pixels (distribution of the pointings in the pixels). Deconvolving the mean spectrum of the binned noiseless map with this mode coupling matrix, squared pixel window function $D^2_\ell(1024)$ and beam would produce an estimate of $C^{\rm th}_\ell$.

Practical calculation of the elements $M^{\rm B}_{\ell \ell'}$ requires that every pointing of the mission should be considered. Therefore we can expect that the calculation of $M^{\rm B}_{\ell \ell'}$ requires considerably more computing resources and time than the calculation of $M_{\ell \ell'}$ (that was given in Eq. (5.31)).

No attempts to calculate $M^{\rm B}_{\ell \ell'}$ were made in this thesis. Instead we approximated $M^{\rm B}_{\ell \ell'} D^2_{\ell'}(1024)$ by $M_{\ell \ell'} D^2_{\ell'}$, which makes Eq. (5.40) into Eq. (5.33). We need to determine the pixel window factor D^2_ℓ of this approximation. Because $M_{\ell \ell'}$ is insensitive to the sampling of the pixels those effects need to be covered by D^2_ℓ.

If we assume that each of the 4 input map pixels, that make a pixel of the binned noiseless map, had been hit the same number of times, the resulting binned noiseless map would just be a downgraded (to $N_{\rm side} = 512$) version of the input map. In this case the pixel window factor of the binned noiseless map would be $D_\ell^2 = D_\ell^2(512)$, where $D_\ell(512)$ is the HEALPix pixel window function for $N_{\rm side} = 512$ ([105]).

Let us consider D_ℓ^2 in a more general way. Under the approximation $M^{\rm B}_{\ell\ell'} D^2_{\ell'}(1024) \approx M_{\ell\ell'} D^2_{\ell'}$ the two-point correlation function between a pair of observed pixels (pixels k and k') of the binned noiseless map is (see e.g. [107])

$$\chi_{kk'} = \sum_\ell \frac{2\ell+1}{4\pi} B_\ell^2 D_\ell^2 C_\ell^{\rm th} P_\ell(\mathbf{n}_k \cdot \mathbf{n}'_k). \tag{5.42}$$

Here $P_\ell(x)$ is the Legendre polynomial and $\mathbf{n}_k \cdot \mathbf{n}'_k$ is the dot product of the two unit vectors pointing to the centers of the pixels. The variance ($\sigma^2_{\rm map}$) of the observed pixels of the binned noiseless map is obtained by setting $k = k'$ (zero lag). Because $P_\ell(1) = 1$, the variance is

$$\sigma^2_{\rm map} = \chi_{kk} = \sum_\ell \frac{2\ell+1}{4\pi} B_\ell^2 D_\ell^2 C_\ell^{\rm th}. \tag{5.43}$$

On the other hand, because the pixel temperature $T^{\rm B}_k$ of the binned noiseless map is a zero mean Gaussian distributed random variable, the variance of the observed pixels can also be expressed as

$$\sigma^2_{\rm map} = \frac{1}{N_{\rm pix} f_{\rm sky}} \sum_k \langle |T^{\rm B}_k|^2 \rangle, \tag{5.44}$$

where the k-sum is over the observed pixels. Inserting $T^{\rm B}_k$ (from Eq. (5.36)) to Eq. (5.44), we obtain for the pixel variance

$$\sigma^2_{\rm map} = \sum_\ell \frac{2\ell+1}{4\pi} B_\ell^2 \left[\frac{1}{N_{\rm pix} f_{\rm sky}} \sum_k \frac{1}{N_k^2} \sum_{i,j\in k} P_\ell(\mathbf{n}_i \cdot \mathbf{n}_j) D_\ell^2(1024) \right] C_\ell^{\rm th}. \tag{5.45}$$

This equation is similar to Eq. (5.43) and a formula for D_ℓ^2 can be extracted

$$D_\ell^2 = \frac{1}{N_{\rm pix} f_{\rm sky}} \sum_k \frac{1}{N_k^2} \sum_{i,j\in k} P_\ell(\mathbf{n}_i \cdot \mathbf{n}_j) D_\ell^2(1024). \tag{5.46}$$

In our study the hits ($\mathbf{n}_i$ and $\mathbf{n}_j$) were distributed in the centers of the four subpixels of the pixel of the binned noiseless (or output) map. If the 4 subpixels are sampled uniformly (in every pixel of the binned noiseless map), the factor

$$D_\ell'^2 = \frac{1}{N_{\rm pix} f_{\rm sky}} \sum_k \frac{1}{N_k^2} \sum_{i,j\in k} P_\ell(\mathbf{n}_i \cdot \mathbf{n}_j). \tag{5.47}$$

multiplying $D_\ell^2(1024)$ in Eq. (5.46) is relatively simple to evaluate. We did it in a simplified case, where we assumed that every pixel of the binned noiseless map is a square with an area of $\Omega_p = 4\pi/N_{\rm pix}$. We found that the value of $D_\ell^{\prime 2}$ was close to $D_\ell^2(512)/D_\ell^2(1024)$. We did not evaluate Eq. (5.47) for the actual hits but instead used the approximation $D_\ell^2 \approx D_\ell^2(512)$ in our study (paper **II** [2]).

Eq. (5.46) can be used in a real experiment too, if we need to estimate the pixel window factor of the output map. In a real experiment we discard $D_\ell^2(1024)$ and use the actual pointings of the beam center in $\mathbf{n}_i$ and $\mathbf{n}_j$.

To find out how well our approximation $M^{\rm B}_{\ell\ell'} D^2_{\ell'}(1024) \approx M_{\ell\ell'} D^2_{\ell'}(512)$ works, we produced $N_{\rm MC} = 450$ CMB-only TODs from $C_\ell^{\rm th}$ (with different $a_{\ell m}$ realizations), binned the TODs to maps, calculated the spectrum $C_\ell^{\rm B}$ from every map and produced the mean spectrum $\langle C_\ell^{\rm B} \rangle_{\rm MC}$. We further deconvolved the mean spectrum with the mode coupling matrix $M_{\ell\ell'}$, symmetric beam and $D_\ell^2 = D_\ell^2(512)$ leading to a spectrum $C_\ell^{\rm BB}$

$$C_\ell^{\rm BB} = \frac{\sum_{\ell'} (M^{-1})_{\ell\ell'} \langle C_{\ell'}^{\rm B} \rangle_{\rm MC}}{B_\ell^2 D_\ell^2(512)}. \tag{5.48}$$

The spectrum $C_\ell^{\rm BB}$ is an estimate of $C_\ell^{\rm th}$. The spectra $C_\ell^{\rm BB}$ and $C_\ell^{\rm th}$ and their ratio for the nearly full sky are shown in Fig. 5.4. It shows that the approximation $M^{\rm B}_{\ell\ell'} D^2_{\ell'}(1024) \approx M_{\ell\ell'} D^2_{\ell'}(512)$ works well up to $\ell \sim 800$. The results of paper **II** ([2]) were produced using this approximation.

The remaining effect in $C_\ell^{\rm BB}$, which blows up at high ℓ, is due to the non-uniform sampling of the pixel area of the binned noiseless map. The approximation $M^{\rm B}_{\ell\ell'} D^2_{\ell'}(1024) \approx M_{\ell\ell'} D^2_{\ell'}(512)$ would be an accurate description of the relation between $\langle C_\ell^{\rm B} \rangle$ and $C_\ell^{\rm th}$ if the number of hits in a pixel (of the binned noiseless map) would be large (approaching ∞), pointings would be distributed uniformly over the pixel area and pixels would be circular in shape. This is, however, not the case here (hits are only in four subpixels of the output map pixel) leading to ℓ-mode coupling that cannot be described by $M_{\ell\ell'}$.

The high-ℓ excess power due to pixel sampling was discussed in paper **II** ([2]), where it was modelled with a bias (*signal bias*) in the pseudo spectrum $\widetilde{C}_\ell$ of the output map. We showed in paper **II** ([2]) that the signal bias was a sum of two parts, where one of the parts was nearly independent from ℓ. We derived analytic formulas to approximate these parts (paper **II** [2]). It was shown that the ℓ-independent part dominates the signal bias at high ℓ.

We calculated the value of the analytic approximation of the ℓ-independent part of the signal bias, subtracted it from $\langle C_\ell^{\rm B} \rangle_{\rm MC}$ and reproduced the spectrum $C_\ell^{\rm BB}$. Although the signal bias is a bias in $\widetilde{C}_\ell$, we used it as an approximation of the bias in $C_\ell^{\rm B}$, because the filter function F_ℓ (see Eq. (5.34)) due to map-making errors is nearly 1 for destriping (paper **II** [2], paper **III** [3]). The ratio of the reproduced $C_\ell^{\rm BB}$ and $C_\ell^{\rm th}$ is shown in Fig. 5.5 ("signal bias removed"). The other ratio ("no signal bias") is the same curve as the one in the bottom panel of Fig. 5.4. We can see that the analytic signal bias models the mode coupling reasonably well up to $\ell \sim 1300$.

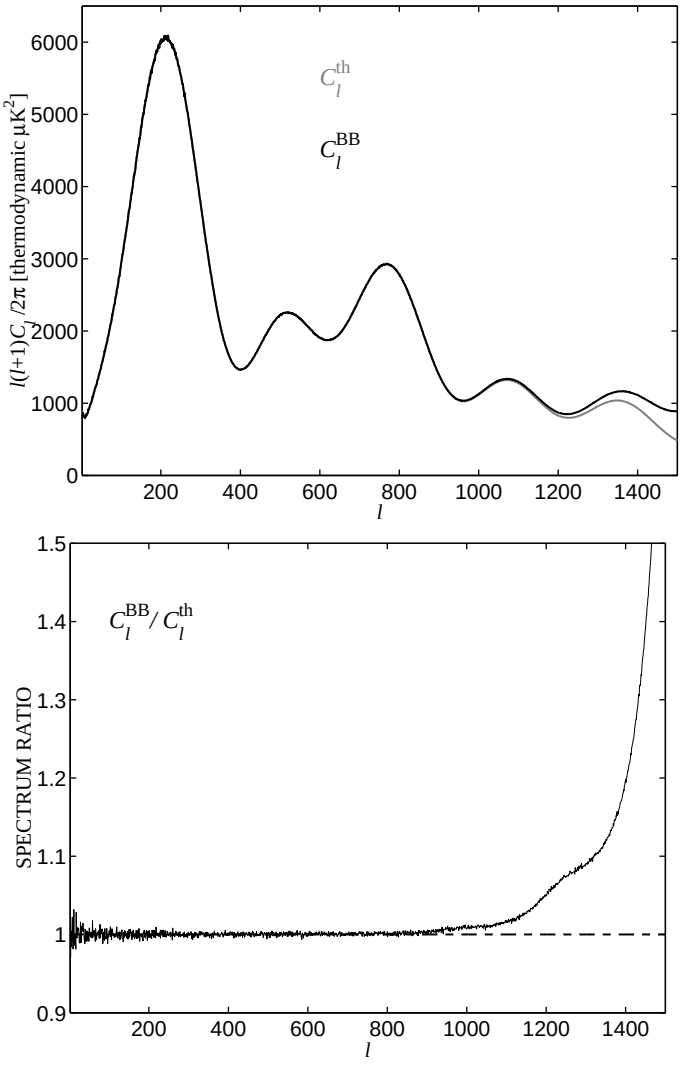

Figure 5.4: **Top:** The angular spectrum $C_\ell^{\rm BB}$ (black curve) is obtained by deconvolving the MC mean of the spectra of the binned noiseless maps ($\langle C_\ell^{\rm B}\rangle_{\rm MC}$) with the mode coupling matrix $M_{\ell\ell'}$, symmetric beam and $N_{\rm side} = 512$ pixel window function (see Eq. (5.48)). The number of MC realizations to produce the mean was $N_{\rm MC} = 450$. The $C_\ell^{\rm B}$ spectra were derived from the nearly full sky maps. The theoretical (ΛCDM) CMB spectrum $C_\ell^{\rm th}$ is shown too (gray curve). **Bottom:** The ratio of the spectra.

The facts, that the filter function is close to 1 in destriping and the signal bias provides a reasonably good model of the high-ℓ performance of the pseudo spectrum, suggested that we could modify the relation between the ensemble mean of the

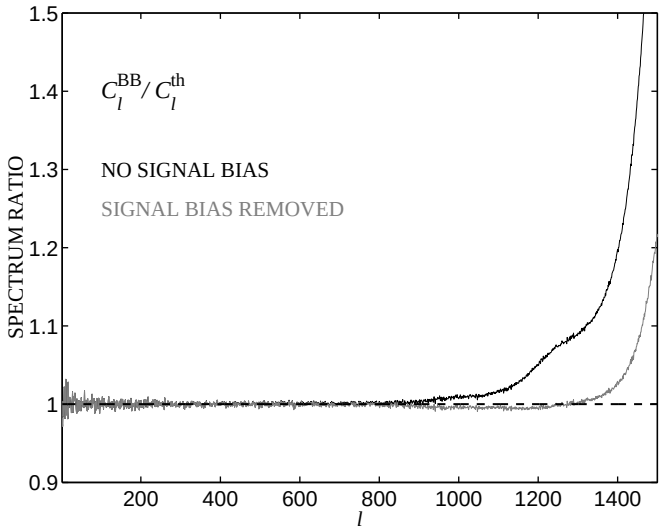

Figure 5.5: The ratio of the spectra $C_\ell^{\rm BB}$ and $C_\ell^{\rm th}$. For the black curve (no signal bias) no signal bias was removed from the mean spectra of the binned noiseless maps ($\langle C_\ell^{\rm B}\rangle$). It is the same as the curve in the bottom panel of Fig. 5.4. For the gray curve (signal bias removed) the analytic approximation of the signal bias was subtracted from $\langle C_\ell^{\rm B}\rangle_{\rm MC}$ before $C_\ell^{\rm BB}$ was calculated.

pseudo spectra and $C_\ell^{\rm th}$ to the following form (paper **II** [2])

$$\langle \widetilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell\ell'} B_{\ell'}^2 D_{\ell'}^2 C_{\ell'}^{\rm th} + S_\ell + \langle \widetilde{N}_\ell \rangle. \tag{5.49}$$

Here S_ℓ is the signal bias. In paper **II** ([2]) we determined the signal biases for the nearly full sky and galactic cut cases using signal-only MC simulations, where we produced $N_{\rm MC} = 450$ CMB-only TODs (with different $a_{\ell m}$ realizations), destriped maps from them and produced $\langle \widetilde{C}_\ell \rangle_{\rm MC}$ from those maps. The signal bias was solved from the above equation ($\langle \widetilde{N}_\ell \rangle = 0$). In this approach any (tiny) map-making errors were immersed in the signal bias. The MC signal biases and Eq. (5.49) were used when the results of paper **II** ([2]) were produced.

5.5.4 Noise Bias

The instrument noise in the output map causes a bias (noise bias) in the pseudo spectrum (see Eqs. (5.30) and (5.49)). To obtain an unbiased estimator of $C_\ell^{\rm th}$ the noise bias needs to be estimated and the estimate removed from the pseudo spectrum.

The definition of the noise bias was given in Eq. (5.14) ([85]). We repeat it here

$$\langle \widetilde{N}_\ell \rangle = {\rm Tr}[\mathbf{E}^\ell \mathbf{N}]. \tag{5.50}$$

For pseudo-C_ℓ estimators the matrix $\mathbf{E}^\ell$ is given by Eq. (5.22) and $\mathbf{N}$ is the noise covariance of the output map. Inserting Eq. (5.22) in place of $\mathbf{E}^\ell$ we obtain for the noise bias

$$\langle \widetilde{N}_\ell \rangle = \frac{\Omega_p^2}{2\ell+1} \sum_{i,j} w_i w_j N_{ij} P_{ij}^\ell = \frac{\Omega_p^2}{2\ell+1} \sum_{i,j} w_i w_j N_{ij} \sum_{m=-\ell}^{\ell} Y_{\ell m}(\mathbf{n}_i) Y_{\ell m}^*(\mathbf{n}_j). \quad (5.51)$$

Here the latter form is obtained after inserting the definition of P_{ij}^ℓ (Eq. (5.5)). The i- and j-sums are over the observed pixels of the output map.

For white noise ($N_{ij} = \sigma_i^2 \delta_{ij}$, where σ_i^2 is the noise variance in pixel i) the noise bias simplifies to

$$\langle \widetilde{N}_\ell \rangle_{\rm white} = \frac{\Omega_p^2}{2\ell+1} \sum_i w_i^2 \sigma_i^2 \sum_{m=-\ell}^{\ell} Y_{\ell m}(\mathbf{n}_i) Y_{\ell m}^*(\mathbf{n}_i) = \Omega_p \frac{1}{N_{\rm pix}} \sum_i w_i^2 \sigma_i^2, \quad (5.52)$$

where the latter form is obtained after replacing $\sum_{m=-\ell}^{\ell} Y_{\ell m}(\mathbf{n}_i) Y_{\ell m}^*(\mathbf{n}_i)$ with $\frac{2\ell+1}{4\pi}$ and Ω_p with $4\pi/N_{\rm pix}$. The total number of output map pixels in the sky is $N_{\rm pix}$. For unit pixel weigths ($w_i = 1$) the noise bias can be expressed as

$$\langle \widetilde{N}_\ell \rangle_{\rm white} = f_{\rm sky} \Omega_p \sigma_m^2, \quad (5.53)$$

where σ_m is the pixel std (from the observed pixels) of the white noise map

$$\sigma_m^2 = \frac{1}{N_{\rm pix} f_{\rm sky}} \sum_i \sigma_i^2 = \frac{\sigma^2}{N_{\rm pix} f_{\rm sky}} \sum_i \frac{1}{N_i}. \quad (5.54)$$

The latter form assumes that the white noise variance of the TOD samples is uniform (σ^2). Here N_i is the number of hits in an output map pixel i.

To calculate an estimate of the noise bias we would need an estimate of the noise covariance matrix ($\mathbf{N}$) of the output map first. This matrix is difficult to produce in a PLANCK-like experiment, where the number of pixels is large ($N_{\rm pix} = 10^6 \ldots 10^8$). Because white noise usually dominates the temperature anisotropy maps at small scales, Eqs. (5.52) - (5.54) provide a reasonably accurate high-ℓ estimate of the noise bias. At low ℓ the noise covariance is increased (from its white noise value) due to temporal correlations of the $1/f$ noise. The effect of $1/f$ noise in the map covariance matrix is difficult to calculate accurately. An analysis of these issues has been carried out for some scanning strategies ([108]).

For destriping an analytic method has been proposed that can provide an estimate of the noise bias ([109]). However, in paper **II** ([2]) we used MC simulations to obtain an estimate for it. We assumed a perfect knowledge of the parameter values of the noise PSD (see Table 5.1). We generated $N_{\rm MC} = 100$ independent noise-only TODs from the noise PSD, destriped maps from them, calculated the pseudo spectra $\widetilde{N}_\ell$ from the maps and produced the mean $\langle \widetilde{N}_\ell \rangle_{\rm MC}$. It was used in Eq. (5.49) instead of $\langle \widetilde{N}_\ell \rangle$.

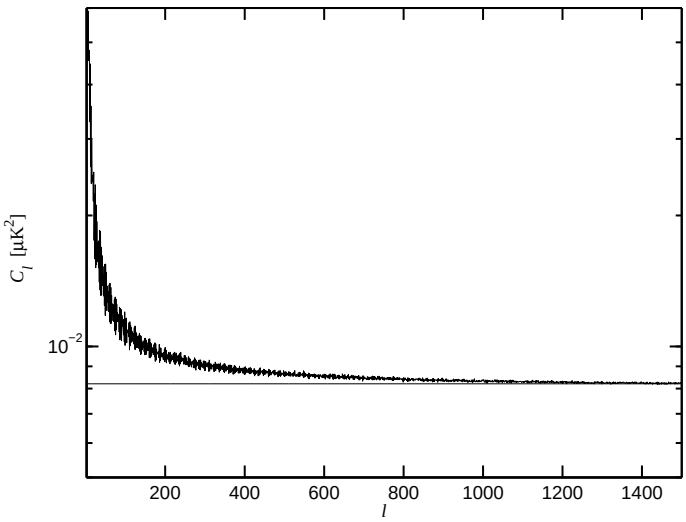

Figure 5.6: Noise bias resulting from noise-only MC simulations (top curve). It is a mean power spectrum of $N_{\rm MC} = 100$ independent noise realizations. We assumed that the noise model parameter values (see Table 5.1) were known. Solid horizontal line is the theoretical white noise spectrum (from Eqs. (5.53) and (5.54)). Noise spectra refer to the modelled set of 24 LFI 100 GHz detectors and nearly full sky coverage. The y-axis is in the CMB temperature scale. The figure is from paper **II** ([2]).

The MC noise bias estimate $\langle \widetilde{N}_\ell \rangle_{\rm MC}$, that we produced for the nearly full sky, is shown in Fig. 5.6. As expected, the theoretical white noise spectrum is a reasonably good approximation of the noise bias at high ℓ, but some excess power appears at low ℓ due to the residues of the $1/f$ noise. This MC noise bias estimate was used when the nearly full sky results of paper **II** ([2]) were produced.

In a real CMB experiment the instrument noise PSD needs to be estimated from the observed data. Iterative and non-iterative methods to do that were described briefly in Sect. 4.3.

The MC noise bias estimate is itself a random variable, whose expectation value is identical to the true noise bias $\langle \widetilde{N}_\ell \rangle$, if we know the noise PSD perfectly and there are no implementation errors in our MC simulations. The variance of the MC noise bias estimate can be approximated as ([110], [111])

$$\langle (\Delta \widetilde{N}_\ell)^2 \rangle_{\rm MC} \approx \frac{2}{(2\ell+1) f_{\rm sky} N_{\rm MC}} \langle \widetilde{N}_\ell \rangle^2_{\rm MC}. \tag{5.55}$$

This uncertainty will increase the magnitudes of the error bars (std of the CMB power spectrum estimate) by a factor $\approx (1 + \frac{1}{2N_{\rm MC}})$ in the noise dominated region (usually at high ℓ) of the spectrum and by a smaller factor at other regions. The number of noise-only MC realizations used in this study ($N_{\rm MC} = 100$) leads only to $\lesssim 0.5\%$ increase in the error bars, which we can consider quite insignificant.

5.5.5 Covariance Matrix of the Power Spectrum Estimate

We carried out signal+noise MC simulations to estimate the covariance matrix of the power spectrum estimate (paper **II** [2]). We produced $N_{\rm MC} = 450$ signal+noise TODs with different CMB and noise realizations, destriped maps from them and produced the pseudo spectra $\widetilde{C}_\ell$ from the maps. Using the noise and signal bias estimates, pixel window factor, beam response and mode coupling matrix $M_{\ell\ell'}$ described earlier in this section, $N_{\rm MC}$ estimates $\widehat{C}_\ell$ (of $C_\ell^{\rm th}$) were produced from the inverted Eq. (5.49). Thereafter the estimate of the covariance matrix was calculated as

$$(\widehat{V}_{\ell\ell'})_{\rm MC} = \langle (\widehat{C}_\ell - \langle \widehat{C}_\ell \rangle_{\rm MC})(\widehat{C}_{\ell'} - \langle \widehat{C}_{\ell'} \rangle_{\rm MC}) \rangle_{\rm MC}. \tag{5.56}$$

In the nearly full sky and galactic cut cases the magnitudes of the covariance matrix elements dropped rapidly when going away from the diagonal. The relatively small number of MC realizations did not allow us to calculate accurate estimates for all elements of the covariance matrix. Therefore we studied only the diagonal elements and some off-diagonal elements close to the diagonal.

Our results showed that the error bars of the power spectrum estimates were considerably larger for the galactic cut than for the nearly full sky (paper **II** [2]). The effective number of independent m-modes (for a given ℓ) is reduced by the sky cut ([92]) which shows up as an increase in the variance of $\widehat{C}_\ell$. The variances obtained from the MC simulations were close to the variances obtained from their analytic model (paper **II** [2]).

In paper **II** ([2]) we applied unit pixel weights ($w_i = 1$) only. We described in Sect. 5.3 how a more accurate power spectrum estimator could be designed, if a number of power spectrum estimates (with appropriate sets of pixel weights) would be produced and they would be combined to an estimate in a maximum-likelihood sense. Using the analytical model of the covariance matrix $\widehat{\mathbf{V}}$ (see Sect. 5.3) we examined the magnitudes of its diagonal elements for different sets of pixel weights w_i to see what effect the weights have in the accuracy of the estimate. This study is not described in paper **II** ([2]), but it is presented below.

We assumed white instrument noise (no $1/f$ noise). Otherwise the parameter values were as in Table 5.1. We assumed that our noise and signal bias estimates were accurate and solved the power spectrum estimate $\widehat{C}_\ell$ from Eq. (5.49). For simplicity we deconvolved the pseudo spectrum with $M_{\ell\ell'}$ only leading to an estimate whose expectation value is $\langle \widehat{C}_\ell \rangle = D_\ell^2(512) B_\ell^2 C_\ell^{\rm th}$. We considered the covariance matrices of the galactic cut only.

The covariance matrix of the pseudo spectrum $\widetilde{C}_\ell$ is given in Eq. (5.18) ([85]). We repeat that equation here:

$$\widetilde{V}_{\ell\ell'} = 2{\rm Tr}[\mathbf{C}\mathbf{E}^\ell\mathbf{C}\mathbf{E}^{\ell'}]. \tag{5.57}$$

Because the output map covariance $\mathbf{C}$ is a sum of signal and noise covariance matrices (Eq. (5.1)), $\widetilde{V}_{\ell\ell'}$ is a sum of three terms: signal $\times$ signal (SS), signal $\times$ noise

(SN) and noise × noise (NN) terms

$$\widetilde{V}_{\ell\ell'} = \widetilde{V}^{\rm SS}_{\ell\ell'} + \widetilde{V}^{\rm SN}_{\ell\ell'} + \widetilde{V}^{\rm NN}_{\ell\ell'}. \tag{5.58}$$

The covariance matrix $\widehat{\mathbf{V}}$ of the CMB power spectrum estimate $\widehat{C}_\ell$ is calculated as in Eq. (5.19).

The evaluation of $\widetilde{\mathbf{V}}$ from Eq. (5.57) is a straightforward but tedious task. The necessary techniques to carry out this calculation can be found in e.g. [94] and [97] (also given in Appendix A of this thesis). We will not show the details of this calculation here but merely give the resulting formulas that we obtained.

Before we could calculate $\widetilde{\mathbf{V}}$ and $\mathbf{M}$ (mode coupling matrix) we needed to define three maps. Their resolution was the same as the resolution of the output map ($N_{\rm side}$ = 512 in this case). The values of their observed pixels were: w_i (map 1), w_i^2 (map 2) and w_i^2/N_i (map 3). The unobserved pixels had value 0 in all maps. The number of hits (N_i) was obtained from the hit count map of the galactic cut case (see the bottom panel of Fig. 5.1). Additionally, we needed the angular power spectra of these maps (W_ℓ^{11}, W_ℓ^{22}, W_ℓ^{33}) and the cross spectrum between the maps 2 and 3: $W_\ell^{23} = \frac{1}{2\ell+1}\sum_m a^{(2)}_{\ell m} a^{(3)*}_{\ell m}$. We further defined a set of matrices (cf. Eq. (5.31))

$$\Gamma^{ij}_{\ell\ell'} = \frac{2\ell'+1}{4\pi}\sum_{\ell''}(2\ell''+1)W^{ij}_{\ell''}\begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{5.59}$$

Here superscripts ij refer to the angular spectra of the maps 1, 2 and 3 and to their cross spectra. The mode coupling matrix is $M_{\ell\ell'} = \Gamma^{11}_{\ell\ell'}$.

The components of the covariance matrix $\widetilde{\mathbf{V}}$ were derived and they can be written as

$$\widetilde{V}^{\rm SS}_{\ell\ell'} = \frac{2\langle\widehat{C}_\ell\rangle\langle\widehat{C}_{\ell'}\rangle}{2\ell'+1}\Gamma^{22}_{\ell\ell'}, \tag{5.60}$$

$$\widetilde{V}^{\rm SN}_{\ell\ell'} = \frac{4\Omega_p\sigma^2\sqrt{\langle\widehat{C}_\ell\rangle\langle\widehat{C}_{\ell'}\rangle}}{2\ell'+1}\Gamma^{23}_{\ell\ell'} \tag{5.61}$$

and

$$\widetilde{V}^{\rm NN}_{\ell\ell'} = \frac{2\Omega_p^2\sigma^4}{2\ell'+1}\Gamma^{33}_{\ell\ell'}. \tag{5.62}$$

The components were summed up to $\widetilde{\mathbf{V}}$ and the covariance matrix $\widehat{\mathbf{V}}$ was then calculated using Eq. (5.19).

We produced the variances of $\widehat{C}_\ell$ (diagonal elements $\widehat{V}_{\ell\ell}$) for different sets of pixel weights w_i. We normalized them with the reference values for the variances (paper **II** [2], see also [110], [111])

$$\langle(\Delta\widehat{C}_\ell)^2\rangle_{\rm ref} = \frac{2}{(2\ell+1)f_{\rm sky}}\left(\langle\widehat{C}_\ell\rangle + N_\ell\right)^2. \tag{5.63}$$

Here $N_\ell = \langle \widetilde{N}_\ell \rangle_{\rm white}/f_{\rm sky}$. $\langle \widetilde{N}_\ell \rangle_{\rm white}$ was evaluated from Eq. (5.53) and its full sky equivalent spectrum was produced by multiplying it with $1/f_{\rm sky}$.

The normalized variances

$$R_{\ell\ell} = \frac{\widehat{V}_{\ell\ell}}{\langle (\Delta \widehat{C}_\ell)^2 \rangle_{\rm ref}} \tag{5.64}$$

were initially evaluated for three sets of pixel weigths: $w_i = 1$, $w_i = N_i$ and $w_i =$ heuristic WMAP weights ([97], see also Eq. 5.27). The calculation of $\widehat{\mathbf{V}}$ for the inverse noise variance weights ($w_i = N_i$) was, however, numerically unstable. Similar observation was made by Efstathiou ([94]). There the numerical problems were caused by the cuspiness of the hit count map. The nominal scanning strategy (that we used) leads to large hit counts in a relatively small areas close to the ecliptic poles which makes the hit count map cuspy in this case too. This may explain the numerical problems that we observed. The inverse noise variance weighting was feasible in the WMAP data analysis, because in WMAP the hits are more uniformly distributed than in Planck leading to less cuspy hit count maps.

Efstathiou proposed regularized weights to prevent the numerical problems associated with the non-uniform hit distribution ([94]). These weights are given in terms of a regularizing parameter $\epsilon_{\rm f}$

$$w_i = \frac{1}{\frac{1}{N_i} + \frac{\epsilon_{\rm f}}{\langle N \rangle}}. \tag{5.65}$$

At $\epsilon_{\rm f} = 1$ Eq. (5.65) gives heuristic WMAP weights and at $\epsilon_{\rm f} = 0$ it gives the inverse noise variance weights.

We replaced the inverse noise variance weights with the regularized weights (using $\epsilon_{\rm f} = 0.5$) and reproduced $R_{\ell\ell}$. The normalized variances for the three sets of weights (unit, heuristic WMAP and regularized weights) are shown in Fig. 5.7. Smaller $R_{\ell\ell}$ leads to smaller error bars for $\widehat{C}_\ell$. Uniform weights give the smallest error bars in the signal dominated part of the spectrum ($\ell \lesssim 600$). In the noise dominated region the heuristic or regularized weights give clearly a better performance than the uniform weights. The difference is $\sim 40\%$ in variances ($\sim 20\%$ in error bars). By reducing $\epsilon_{\rm f}$ the high-ℓ performance of the regularized weights could be improved with the increased risk for numerical instabilities that occur latest when $\epsilon_{\rm f}$ reaches zero.

Fig. 5.7 can be compared to the upper panel of Fig. 16 (of paper **II** [2]), where $R_{\ell\ell}$ obtained from the MC simulations is displayed. The "$w_i = 1$" curve of Fig. 5.7 is nearly the same as the horizontal line of Fig. 16 (galactic cut). There is a reasonably good fit between the $R_{\ell\ell}$ of the MC simulations and analytical model.

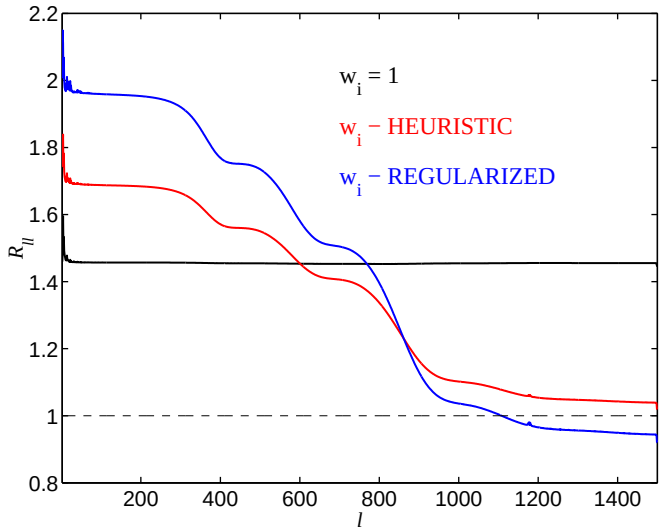

Figure 5.7: Normalized variance $R_{\ell\ell} = \widehat{V}_{\ell\ell}/\langle(\Delta\widehat{C}_\ell)^2\rangle_{\rm ref}$ for three different sets of pixel weights: $w_i = 1$, $w_i =$ heuristic WMAP weights (Eq. (5.27), [97]) and $w_i =$ regularized weights (Eq. (5.65) with $\epsilon_{\rm f} = 0.5$, [94]). The curves were derived from an analytical model of the covariance matrix $\widehat{\mathbf{V}}$ whose details were derived in this thesis but that used results from [94] and [97]. The variances are for the galactic cut.

Chapter 6

Conclusions

The objective of this thesis has been to study map-making and angular power spectrum estimation from the CMB temperature anisotropy observations of PLANCK-like experiments. The map-making approach considered in this thesis is destriping. The main benefit of destriping is its simplicity as compared to the GLS map-making algorithms. Destriping requires no prior information of the instrument noise, but still it can produce output maps whose noise is only slightly higher than the noise of the GLS output maps. Under the assumption of Gaussian distributed noise GLS output maps fall close to the minimum variance map.

We have presented a maximum-likelihood formulation of the destriping and implemented destriping codes that were applied to the simulated PLANCK-like data sets (paper **I** [1]). We compared the output maps of our destriping method to the output maps of two already existing destriping methods. The difference in these methods is in the weights they assign to the map pixels. We found that our method produces lower noise maps than the method presented in [69], [71], [73], although the difference was rather small. The map noise of our method and the method described in [70] was nearly the same.

We examined the possibility to improve the accuracy of destriping by fitting more base functions than simple uniform baselines (paper **I** [1]). For the assumed instrument noise (white + $1/f$ noise), we found that multiple base functions did not systematically reduce the map noise as compared to the uniform baselines. For typical detector knee frequencies ($\lesssim 0.1$ Hz) we found no improvement. For high knee frequencies (0.4 Hz in our case) fitting uniform baselines and Legendre polynomials produced lower map noise than fitting uniform baselines alone, but the difference was very small. The optimal set of base functions seemed to depend on the instrument noise characteristics. It was suggested in paper **I** ([1]) that the symmetries of the scanning may reduce the accuracy at which the amplitudes of the added base functions can be determined. A longer mission time or steeper slope of the $1/f^{\alpha}$ noise (we had 7 months mission time and slope $\alpha = 1$ in paper **I** [1]) may be more favorable cases for additional base functions beyond the uniform baselines. In our later studies we found that multiple base functions can systematically lower the

map noise if the prior covariance of the base function amplitudes is used when solving the estimates of the amplitudes ([77]). Amplitude covariance requires, however, prior information of the instrument noise.

In this thesis we compared the output maps of our destriping code to the output maps of two GLS map-making codes (paper **III** [3]). The maps were made from TODs representing one year of observations of a PLANCK LFI detector. As expected, the noise of the GLS output maps was lower than the noise of the destriped maps. The difference was, however, rather small. The pixelization noise caused by the spread of the observations in the output map pixels leads to a signal dependent map-making error in the output maps. That error was larger in the GLS maps than in the destriped maps, because only the uniform baselines of the pixelization noise contribute to this error in destriping but in the GLS map-making also the higher frequency components of the pixelization noise contribute.

In this thesis we applied a MASTER-like pseudo-C_ℓ power spectrum estimation method to the output maps of destriping (paper **II** [2]). We used simulated data of a PLANCK LFI detector and demonstrated that this approach works well leading to accurate estimates of the true power spectrum of the CMB temperature anisotropy. We found that the signal dependent destriping error (due to pixelization noise) does not cause a significant error in the power spectrum estimates. Instead we discovered a more significant power spectrum error, that does not depend on the map-making algorithm. It is due to the non-uniform distribution of the hits in the output map pixels. This causes both spectral smoothing and some excess power at high ℓ. In paper **II** ([2]) the spectral smoothing was approximated with a HEALPix pixel window function and corrected from the power spectrum estimates. To describe the high-ℓ excess power the concept of signal bias of the power spectrum estimate was introduced in paper **II** ([2]). Typically the signal bias was some orders of magnitude smaller than the noise bias of the pseudo spectrum. The excess power was less pronounced for multiple detector observations than for a single detector case, since the observations of multiple detectors covered the pixels more uniformly (paper **II** [2]). Large part of the high-ℓ excess power could be removed by subtracting an estimate of the signal bias that was produced using MC simulations. We also derived a reasonably accurate analytical estimate for the signal bias (paper **II** [2]).

Appendix A

Quadratic Power Spectrum Estimator for Signal and Noise Dominated Maps

In Sect. 5.3 the quadratic power spectrum estimator

$$\widetilde{C}_\ell = \mathbf{x}^{\mathrm{T}}\mathbf{E}^\ell\mathbf{x}. \tag{A.1}$$

was introduced. The vector $\mathbf{x}$ is the map containing a sum of CMB and noise. Its covariance matrix is $\mathbf{C} = \mathbf{S} + \mathbf{N}$ (see Sect. 5.1).

With a proper choice of matrix $\mathbf{E}^\ell$ the quadratic estimator becomes statistically equivalent to the maximum likelihood solution. The matrix $\mathbf{E}^\ell$ that gives that performance is

$$\mathbf{E}^\ell = \frac{1}{2F_{\ell\ell}}\mathbf{C}^{-1}\mathbf{P}^\ell\mathbf{C}^{-1}. \tag{A.2}$$

The matrix $\mathbf{F}$ is the Fisher information matrix (Eq. (5.9)) and the matrix $\mathbf{P}^\ell$ is defined in Eq. (5.5).

At large scales (low ℓ) the map $\mathbf{x}$ is typically dominated by the CMB signal, whereas the noise typically dominates at small scales (high ℓ). We will derive the formulas for the matrix $\mathbf{E}^\ell$, that apply in the signal and noise dominated maps.

A.1 Signal Dominated Map

We can approximate the map covariance as $\mathbf{C} \approx \mathbf{S}$. The matrix $\mathbf{S}$ is given by Eq. (5.4). Unless otherwise noted the ℓ-sums go from 0 to ∞.

We need to evaluate $\mathbf{C}^{-1}$ first. Let us try an ansatz ([94], [97])

$$\widehat{\mathbf{C}^{-1}} = \Omega_p^2 \sum_\ell \frac{1}{C_\ell}\mathbf{P}^\ell. \tag{A.3}$$

We specified that $C_0 = C_1 = 0$ (see Sect. 5.1). In order to avoid problems we assume for the time being that C_0 and C_1 are non-zero and let them approach zero at the end.

We calculate next

$$(\widehat{\mathbf{C}^{-1}}\mathbf{C})_{ij} = \sum_k (\widehat{\mathbf{C}^{-1}})_{ik}\mathbf{C}_{kj} = \Omega_p^2 \sum_{\ell,\ell'} \frac{C_\ell}{C_{\ell'}} \sum_k P_{ik}^\ell P_{kj}^{\ell'}. \tag{A.4}$$

The k-sum is over the pixels of the observed sky. The k-sum

$$\sum_k P_{ik}^\ell P_{kj}^{\ell'} = \sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'} \sum_k Y_{\ell m}(\mathbf{n}_i) Y_{\ell m}^*(\mathbf{n}_k) Y_{\ell' m'}(\mathbf{n}_k) Y_{\ell' m'}^*(\mathbf{n}_j) \tag{A.5}$$

contains a coupling sum $\sum_k Y_{\ell m}^*(\mathbf{n}_k) Y_{\ell' m'}(\mathbf{n}_k)$ that, for the full sky and for moderate sky cuts, peaks sharply at $|\ell - \ell'| \ll \ell$ ([94], [97]). Therefore we may write

$$\sum_k Y_{\ell m}^*(\mathbf{n}_k) Y_{\ell' m'}(\mathbf{n}_k) \approx \frac{1}{\Omega_p} \delta_{\ell\ell'} \delta_{mm'}, \tag{A.6}$$

which is a good approximation for the full sky and for moderate sky cuts. The coupling sum is closely related to the mode coupling matrix (Eq. (5.24)). It is shown in Sect. 5.5.2 that the mode coupling matrix is close to a unit matrix for the full sky and diagonally dominant for a cut sky, where the galactic region at ± 20 deg from the galactic plane has been cut out from the map. Therefore, the approximation Eq. (A.6) is justified in these cases. Using the approximation in Eq. (A.5) leads to

$$\sum_k P_{ik}^\ell P_{kj}^{\ell'} = \frac{1}{\Omega_p} P_{ij}^\ell \delta_{\ell\ell'}. \tag{A.7}$$

Inserting this back to Eq. (A.4) gives

$$(\widehat{\mathbf{C}^{-1}}\mathbf{C})_{ij} = \Omega_p \sum_\ell P_{ij}^\ell = \Omega_p \sum_{\ell m} Y_{\ell m}(\mathbf{n}_i) Y_{\ell m}^*(\mathbf{n}_j) = \delta_{ij}, \tag{A.8}$$

where the Kronecker delta results from the completeness relation of the spherical harmonic functions.

Next we obtain for the matrix $\mathbf{E}^\ell$

$$2F_{\ell\ell} E_{ij}^\ell = \sum_{k,q} (\mathbf{C}^{-1})_{ik} P_{kq}^\ell (\mathbf{C}^{-1})_{qj} = \sum_{\ell',\ell''} \frac{1}{C_{\ell'} C_{\ell''}} \sum_{k,q} P_{ik}^{\ell'} P_{kq}^\ell P_{qj}^{\ell''}. \tag{A.9}$$

Applying Eq. (A.7) first for the k-sum and then for the q-sum, we obtain

$$2F_{\ell\ell} E_{ij}^\ell = \frac{1}{\Omega_p^2 C_\ell^2} P_{ij}^\ell. \tag{A.10}$$

Next we calculate the elements of the Fisher information matrix (see Eq. (5.9))

$$F_{\ell\ell'} = \frac{1}{2}\mathrm{Tr}[\mathbf{C}^{-1}\mathbf{P}^{\ell}\mathbf{C}^{-1}\mathbf{P}^{\ell'}] = \sum_{\ell_1,\ell_2}\frac{1}{2C_{\ell_1}C_{\ell_2}}\sum_{i,j,k,q}P^{\ell_1}_{ij}P^{\ell}_{jk}P^{\ell_2}_{kq}P^{\ell'}_{qi}. \tag{A.11}$$

The j-sum ($\sum_j Y^*_{\ell_1 m_1}(\mathbf{n}_j)Y_{\ell m}(\mathbf{n}_j)$) and q-sum ($\sum_q Y^*_{\ell_2 m_2}(\mathbf{n}_q)Y_{\ell' m'}(\mathbf{n}_q)$) are coupling sums that peak at $|\ell - \ell_1| \ll \ell$ and $|\ell' - \ell_2| \ll \ell'$. Thus we may pull C_{ℓ_1} and C_{ℓ_2} out from the ℓ_1, ℓ_2-sum and replace them with C_ℓ and $C_{\ell'}$ ([94], [97]). Inserting this approximation we obtain

$$F_{\ell\ell'} = \frac{1}{2C_\ell C_{\ell'}}\sum_{\ell_1,\ell_2}\sum_{i,j,k,q}P^{\ell_1}_{ij}P^{\ell}_{jk}P^{\ell_2}_{kq}P^{\ell'}_{qi}. \tag{A.12}$$

Applying the completeness relation to ℓ_1- and ℓ_2-sums ($\sum_{\ell_1} P^{\ell_1}_{ij} = \delta_{ij}/\Omega_p$ and $\sum_{\ell_2} P^{\ell_2}_{kq} = \delta_{kq}/\Omega_p$) we obtain for the Fisher matrix

$$F_{\ell\ell'} = \frac{1}{2C_\ell C_{\ell'}\Omega_p^2}\sum_{i,k}P^{\ell}_{ik}P^{\ell'}_{ki} = \frac{2\ell+1}{2C_\ell C_{\ell'}\Omega_p^4}W_{\ell\ell'}, \tag{A.13}$$

where we have defined a matrix

$$W_{\ell\ell'} = \frac{\Omega_p^2}{2\ell+1}\sum_{i,k}P^{\ell}_{ik}P^{\ell'}_{ki} = \frac{1}{2\ell+1}\sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'}\left|\Omega_p\sum_i Y_{\ell m}(\mathbf{n}_i)Y_{\ell' m'}(\mathbf{n}_i)\right|^2. \tag{A.14}$$

The latter form is obtained by inserting the definition of the matrix $\mathbf{P}^\ell$ (from Eq. (5.5)). The i and j sums are taken over the uncut pixels of the sky. It is worth noting that the matrix $W_{\ell\ell'}$ is identical to the mode coupling matrix (with unit pixel weights) used in the MASTER power spectrum estimation method ([93], see also Eq. (5.24) of this thesis).

We can now insert Eq. (A.13) in Eq. (A.10) and solve $\mathbf{E}^\ell$. The result is

$$E^{\ell}_{ij} = \frac{\Omega_p^2}{(2\ell+1)W_{\ell\ell}}P^{\ell}_{ij}. \tag{A.15}$$

Due to the freedom of normalization we use 1 instead of $W_{\ell\ell}$ and finally obtain for the matrix $\mathbf{E}^\ell$

$$E^{\ell}_{ij} = \frac{\Omega_p^2}{2\ell+1}P^{\ell}_{ij}. \tag{A.16}$$

The spectrum C_ℓ does not appear in our final E^ℓ_{ij} and therefore the limits $C_0 \to 0$ and $C_1 \to 0$ are trivial.

A.2 Noise Dominated Map

Now the map covariance can be approximated as $\mathbf{C} \approx \mathbf{N}$. We assume white map noise with diagonal noise covariance $N_{ij} = \sigma_i^2 \delta_{ij}$, where σ_i^2 is the noise variance in pixel i. We calculate

$$2F_{\ell\ell}E_{ij}^{\ell} = \sum_{k,q} (\mathbf{C}^{-1})_{ik} P_{kq}^{\ell} (\mathbf{C}^{-1})_{qj}, \tag{A.17}$$

where k and q sums are over the observed pixels of the sky. Noting that $(\mathbf{C}^{-1})_{ij} \approx 1/\sigma_i^2 \delta_{ij}$ the equation can be written as

$$2F_{\ell\ell}E_{ij}^{\ell} = \frac{1}{\sigma_i^2 \sigma_j^2} P_{ij}^{\ell}. \tag{A.18}$$

Due to the freedom of normalization we can select scaling $\frac{\Omega_p^2}{2\ell+1}$ (instead of $\frac{1}{2F_{\ell\ell}}$) and obtain the matrix $\mathbf{E}^{\ell}$ in a similar form as in the signal dominated case

$$\mathbf{E}^{\ell} = \frac{\Omega_p^2}{2\ell+1} \frac{1}{\sigma_i^2 \sigma_j^2} \mathbf{P}^{\ell}. \tag{A.19}$$

We assume a HEALPix pixelization ([52]), where every pixel of the map $\mathbf{x}$ has the same area $\Omega_p = 4\pi/N_{\mathrm{pix}}$.

Appendix B

Beam Deconvolution

In the discussion of the power spectrum estimation (see Ch. 5) we systematically assumed a perfectly radially symmetric response of the telescope main beam. The deconvolution of the power spectrum with a symmetric beam is simple. It becomes considerably more difficult if the symmetric beam assumption is relaxed and asymmetric beam responses are considered. Studies (e.g. [40], [106], [107]) exist where methods for the asymmetric beam deconvolution are discussed. In the case of an asymmetric beam, the beam smoothings of the output map pixels will vary (from pixel to pixel) due to different orientations of the beam. Some of the existing methods attempt to determine an effective symmetric beam response that approximates the mean of the smoothings of the pixels (see e.g. [107]).

In Sect. 5.5.3 we examined the relation of the ensemble mean angular power spectrum of the binned noiseless map ($\langle C_\ell^{\mathrm{B}} \rangle$) and the theoretical CMB temperature anisotropy spectrum (C_ℓ^{th}). A symmetric beam was assumed in that discussion. The intention of this Appendix is to repeat that discussion using an asymmetric beam. Our goal is to find a general mode coupling matrix that would give a full description of the relation between these two spectra. We call it the mode coupling matrix of the experiment and use a symbol $M_{\ell\ell'}^{\mathrm{E}}$ for it. Using this matrix the relation between these two spectra could be written as (ℓ-sums go from 0 to ∞ in this Appendix)

$$\langle C_\ell^{\mathrm{B}} \rangle = \sum_{\ell'} M_{\ell\ell'}^{\mathrm{E}} C_{\ell'}^{\mathrm{th}}. \tag{B.1}$$

The mode coupling matrix of the experiment describes correctly the effects of sky coverage, weighting of the pixels, sampling of the observed pixels (distribution of observations in them) and beam.

The mode coupling matrix $M_{\ell\ell'}^{\mathrm{E}}$ would be useful in a real (noisy) CMB experiment. Let us assume that we have an accurate estimate of the noise bias and the map-making errors (that were accounted for by the filter function F_ℓ, see Sect. 5.5.3) are so small that they can be ignored (according to papers **II** [2] and **III** [3] this is a good assumption at least for destriping). Under these assumptions an estimate of C_ℓ^{th}, with an insignificant bias, could be obtained by deconvolving the difference

of the pseudo spectrum and noise bias estimate with $M^{\mathrm{E}}_{\ell\ell'}$. In this scheme no pixel window factors, signal bias estimates or symmetric beam approximations would be required.

As in Sect. 5.5.3 we consider the samples d_i of the CMB-only TOD that fall in a pixel k of the binned noiseless map. In addition to the pointing of the beam center ($\mathbf{n}_i$), we need now to consider the beam orientation as well. The beam orientation in a TOD sample is given by an orientation angle ψ_i (see Sect. 3.1.3). In this Appendix we will consider arbitrary detector pointings (not in the pixels of an input map).

The temperature d_i can be given as (see Eq. (4.41))

$$d_i = \sum_{\ell' m'} a_{\ell' m'} b^*_{\ell' m'}(\mathbf{n}_i, \psi_i). \tag{B.2}$$

Here the $a_{\ell m}$ are a realization of C^{th}_ℓ and $b_{\ell m}(\mathbf{n}_i, \psi_i)$ are the beam coefficients when the beam center is pointing to $\mathbf{n}_i$ and its orientation angle is ψ_i. The beam coefficients are given in Eq. (3.5). We repeat that formula here

$$b_{\ell' m'}(\mathbf{n}_i, \psi_i) = \sum_{m''=-\ell'}^{\ell'} b_{\ell' m''} D^{\ell'}_{m' m''}(\mathbf{n}_i, \psi_i). \tag{B.3}$$

Here $b_{\ell m}$ are the beam coefficients in its reference pointing (usually towards the north ecliptic pole) and orientation. $D^{\ell}_{m m'}(\mathbf{n}_i, \psi_i)$ are the Wigner D-functions. We assume that the beam response is band-limited having insignificant (~ 0) power beyond some ℓ_{max}.

The temperature of the pixel k of the binned noiseless map is (we assume unit weights for the observed pixels)

$$T^{\mathrm{B}}_k = \frac{1}{N_k} \sum_{i \in k} d_i = \sum_{\ell' m'} a_{\ell' m'} \frac{1}{N_k} \sum_{i \in k} b^*_{\ell' m'}(\mathbf{n}_i, \psi_i), \tag{B.4}$$

where $i \in k$ refers to those TOD samples that hit the pixel k and N_k is the number of hits in that pixel. The $a_{\ell m}$ of the binned noiseless map are obtained as

$$a^{\mathrm{B}}_{\ell m} = \Omega_p \sum_k T^{\mathrm{B}}_k Y^*_{\ell m}(\mathbf{q}_k). \tag{B.5}$$

The unit vector pointing to the center of the pixel k (of the binned noiseless map) is $\mathbf{q}_k$ and the k-sum is taken over the observed pixels. The area of a pixel is Ω_p (same in every pixel).

Inserting T^{B}_k from Eq. (B.4) we obtain for the $a^{\mathrm{B}}_{\ell m}$

$$a^{\mathrm{B}}_{\ell m} = \sum_{\ell' m'} a_{\ell' m'} \Omega_p \sum_k \frac{1}{N_k} \sum_{i \in k} b^*_{\ell' m'}(\mathbf{n}_i, \psi_i) Y^*_{\ell m}(\mathbf{q}_k). \tag{B.6}$$

This equation defines the $a_{\ell m}$ coupling matrix of the experiment

$$K^{\mathrm{E}}_{\ell m \ell' m'} \equiv \Omega_p \sum_k \frac{1}{N_k} \sum_{i \in k} b^*_{\ell' m'}(\mathbf{n}_i, \psi_i) Y^*_{\ell m}(\mathbf{q}_k). \tag{B.7}$$

It defines a relation between the $a_{\ell m}$ of the binned noiseless map and CMB sky

$$a^{\mathrm{B}}_{\ell m} = \sum_{\ell' m'} a_{\ell' m'} K^{\mathrm{E}}_{\ell m \ell' m'}. \tag{B.8}$$

Due to a band-limited beam, $K^{\mathrm{E}}_{\ell m \ell' m'}$ has significant elements up to $\ell, \ell' = \ell_{\max}$. Although the elements have four indeces, a usual practice is to combine (ℓ, m) to a single index $i(\ell, m)$ ([92]). In this scheme the size of the matrix $K^{\mathrm{E}}_{\ell m \ell' m'}$ is $(\ell_{\max} + 1)^2 \times (\ell_{\max} + 1)^2$.

Using the statistical isotropy of the CMB sky (Eq. (5.3)) we obtain for the ensemble mean of the angular power spectrum of the binned noiseless map

$$\langle C^{\mathrm{B}}_\ell \rangle = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \langle |a^{\mathrm{B}}_{\ell m}|^2 \rangle = \sum_{\ell'} \left[\frac{1}{2\ell + 1} \sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'} |K^{\mathrm{E}}_{\ell m \ell' m'}|^2 \right] C^{\mathrm{th}}_{\ell'}. \tag{B.9}$$

The mode coupling matrix of the experiment can now be extracted

$$M^{\mathrm{E}}_{\ell \ell'} = \frac{1}{2\ell + 1} \sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'} |K^{\mathrm{E}}_{\ell m \ell' m'}|^2. \tag{B.10}$$

It can also be expressed in the following form (after inserting Eqs. (B.3) and (B.7))

$$M^{\mathrm{E}}_{\ell \ell'} = \frac{1}{2\ell + 1} \sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'} \left| \Omega_p \sum_{m''=-\ell'}^{\ell'} b^*_{\ell' m''} \sum_k \left[\frac{1}{N_k} \sum_{i \in k} D^{\ell' *}_{m' m''}(\mathbf{n}_i, \psi_i) \right] Y^*_{\ell m}(\mathbf{q}_k) \right|^2. \tag{B.11}$$

All information exists to derive the elements $M^{\mathrm{E}}_{\ell \ell'}$. We expect that their calculation is not a trivial task, because it involves the pointings and beam orientations of the entire mission. We have not made an attempt to calculate $M^{\mathrm{E}}_{\ell \ell'}$ for actual pointings and beam orientations in this thesis.

We can, however, examine $M^{\mathrm{E}}_{\ell \ell'}$ in a simplified situation. Let us assume that $N_k \gg 1$ in every pixel and pixels have a regular circular shape. We further assume that the pixel area and the range of beam orientations ($\psi_i \in [0 \ldots 2\pi]$) are sampled uniformly in every pixel. Under these assumptions the i-sum of Eq. (B.11) can be approximated by an integral over the pixel area and beam orientations. That integral has a simple result

$$\frac{1}{N_k} \sum_{i \in k} D^{\ell' *}_{m' m''}(\mathbf{n}_i, \psi_i) \approx \sqrt{\frac{4\pi}{2\ell' + 1}} G_{\ell'}(k) Y_{\ell' m'}(\mathbf{q}_k) \delta_{m'' 0}, \tag{B.12}$$

where

$$G_{\ell'}(k) = \frac{1}{\Omega_p}\sqrt{\frac{4\pi}{2\ell'+1}}\int_{k\to z} d\Omega_{\mathbf{n}} Y_{\ell' 0}(\mathbf{n}). \tag{B.13}$$

The symbol "$k \to z$" in the integration limit means that the integration is over the area of the pixel k (of the binned noiseless map) but the map has been rotated before the integration in such a way that $\mathbf{q}_k$ (unit vector pointing to the center of the pixel k) points now towards the north pole (towards +z-axis). Because we assumed a circular pixel shape the function $G_{\ell'}(k)$ has the same value in every pixel.

It is of interest to find the relation between $G_\ell(k)$ and the HEALPix pixel window function $D_\ell(N_{\rm side})$ ([105]) of the pixel size of the binned noiseless map ($N_{\rm side}$ was 512 in Sect. 5.5). The HEALPix pixel window function describes the spectral smoothing of a map where a pixel temperature is the underlying temperature field averaged over the pixel area. The HEALPix pixel window function is given in Eq. (30) of [105], where a symbol w_ℓ is used for it. We use a symbol $D_\ell(N_{\rm side})$ for it in this thesis. It can be shown that in the case of an identical circular shape of the pixels, Eq. (30) of [105] would give a pixel window function that is identical to the function $G_\ell(k)$ that we introduced above. Therefore it is a reasonable approximation to use $D_{\ell'}(N_{\rm side})$ in place of $G_{\ell'}(k)$ in Eq. (B.12).

Inserting Eq. (B.12) (with $D_{\ell'}(N_{\rm side})$ in place of $G_{\ell'}(k)$) to Eq. (B.11) we obtain

$$M^{\rm E}_{\ell\ell'} \approx \frac{4\pi}{2\ell'+1} b^2_{\ell' 0} D^2_{\ell'}(N_{\rm side}) \frac{1}{2\ell+1} \sum_{m,m'=-\ell,-\ell'}^{\ell,\ell'} \left| \Omega_p \sum_k Y_{\ell' m'}(\mathbf{q}_k) Y^*_{\ell m}(\mathbf{q}_k) \right|^2. \tag{B.14}$$

In this simplified case the contributions of different effects separate out. The pixel window factor $D^2_\ell(N_{\rm side})$ describes the spectral smoothing due to sampling of the observed pixels. The mean response of the beam is $\frac{4\pi}{2\ell+1} b^2_{\ell 0}$ (the same formula was derived in [92] using a different approach). It is symmetric because all orientations were sampled uniformly. The remaining part is the mode coupling matrix introduced by MASTER ([93], see also Eq. (5.24)). It describes the effects of sky coverage and weighting of the pixels. It was discussed in detail in Sect. 5.5.2.

Finally, it is worth to consider, how Eq. (B.8) could be applied in the deconvolution map-making. A deconvolution map-making method (not based on the technique discussed here) has been proposed in [81]. It was briefly described in Sect. 4.6 of this thesis.

Let us assume that we have the $K^{\rm E}_{\ell m \ell' m'}$ values up to an $\ell_{\rm max}$ given by the band-limitation of the beam. We can invert that matrix provided that it is numerically non-singular and we have sufficient computing resources to do it (i.e., $\ell_{\rm max}$ is not too large). If we compute the pseudo $\widetilde{a}_{\ell m}$ of the (noisy) output map of the map-making and multiply it with the inverse of $K^{\rm E}_{\ell m \ell' m'}$, we obtain an estimate ($\widehat{a}_{\ell m}$) of the $a_{\ell m}$ of the CMB sky. If the output map has a full sky coverage this method should work, in principle, without problems. In the case of partial sky coverage the accuracy of the map estimate (derived from $\widehat{a}_{\ell m}$) would be deteriorated in the vicinity of the boundary of the sky cut ([92]). In both cases the noise covariance of the map estimate

would be complicated due to mode coupling arising from the multiplication with the inverse of $K^{\rm E}_{\ell m\ell' m'}$.

The $a_{\ell m}$ coupling matrix $K^{\rm E}_{\ell m\ell' m'}$ can be numerically singular due to one (or both) of two reasons: 1) The beam response is a smoothing function (its value decreases with increasing ℓ) which tends to increase the condition number of $K^{\rm E}_{\ell m\ell' m'}$. 2) For a partial sky coverage the matrix $K^{\rm E}_{\ell m\ell' m'}$ is numerically non-singular only up to some maximum ℓ, that depends on the geometry and the extent of the unobserved area of the sky ([92]). Increasing the area of the unobserved sky will decrease that maximum ℓ. As an example, [92] shows that the condition number of the $a_{\ell m}$ coupling matrix for ± 20 deg galactic cut is $\sim 5\times 10^9$, if $\ell_{\rm max}$ of the coupling matrix is $\sim$50. If $\ell_{\rm max}$ (due to band limitation of the beam) of our matrix $K^{\rm E}_{\ell m\ell' m'}$ is larger than the maximum ℓ supported by the observed area of the sky, our matrix will be numerically non-singular.

To reduce the first problem we can use a modified coupling matrix (instead of the original one) $K^{\rm E'}_{\ell m\ell' m'} = K^{\rm E}_{\ell m\ell' m'}/G_{\ell'}$, where G_ℓ is a symmetric response that approximates the asymmetric beam of the experiment. In this case the resulting map estimate will be smoothed with G_ℓ. This technique was used in [81] to regularize the matrix that was used there. To reduce the second problem we can extract a left upper block of the matrix $K^{\rm E}_{\ell m\ell' m'}$ up to such an ℓ, where the extracted block is numerically non-singular. We can then deconvolve the pseudo $\widetilde{a}_{\ell m}$ with that block. This, however, will allow us to solve our map estimate $\widehat{a}_{\ell m}$ only up to $\ell_{\rm max}$ of the extracted block.

Bibliography

[1] E. Keihänen, H. Kurki-Suonio, T. Poutanen, D. Maino, and C. Burigana, A&A **428**, 287 (2004). (Paper **I** of this thesis)

[2] T. Poutanen, D. Maino, H. Kurki-Suonio, E. Keihänen, and E. Hivon, MNRAS **353**, 43 (2004). (Paper **II** of this thesis)

[3] T. Poutanen et al., to be published in A&A, [astro-ph/0501504] (2005). (Paper **III** of this thesis)

[4] J.C. Mather, D.J. Fixsen, R.A. Shafer, C. Mosier, and D.T. Wilkinson, ApJ **512**, 511 (1999).

[5] A.A. Penzias, and R.W. Wilson, ApJ **142**, 419 (1965).

[6] G.F. Smoot et al., ApJ **396**, L1 (1992).

[7] C.L. Bennett et al., ApJS **148**, 1 (2003)

[8] http::/www.rssd.esa.int/Planck/.

[9] J.A. Tauber, in Proceedings of COSPAR meeting (2004).

[10] G. Gamow, Nature **162**, 680 (1948).

[11] G. Gamow, Phys. Rev. **74**, 505 (1948).

[12] R.A. Alpher, and R.C. Herman, Nature **162**, 774 (1948).

[13] R.A. Alpher, Phys. Rev. **74**, 1577 (1948).

[14] R.A. Alpher, and R.C. Herman, Phys. Rev. **75**, 1089 (1949).

[15] R.K. Sachs, and A.M. Wolfe, ApJ **147**, 73 (1967).

[16] P.J.E. Peebles, and J.T. Yu, ApJ **162**, 815 (1970).

[17] P.J.E. Peebles, *The Large-Scale Structure of the Universe*, (Princeton University Press, 1980).

[18] V.F. Mukhanov, H.A. Feldman, and R.H. Brandenberger, Physics Reports **215**, 203 (1992).

[19] J.M. Bardeen, Phys. Rev. D **22**, 1882 (1980).

[20] A.R. Liddle, and D.H. Lyth, *Cosmological Inflation and Large-Scale Structure*, (Cambridge University Press, 2000.)

[21] W.T. Hu, PhD Thesis, University of California at Berkeley, [astro-ph/9508126] (1995).

[22] J. Väliviita, PhD Thesis, University of Helsinki, 2005.

[23] C-P. Ma, and E. Bertschinger, ApJ **455**, 7 (1995).

[24] S. Dodelson, *Modern Cosmology*, (Academic Press, 2003).

[25] A. Kosowsky, Annals of Physics **246**, 49 (1996).

[26] U. Seljak, and M. Zaldarriaga, ApJ **469**, 437 (1996).

[27] M. Bucher, K. Moodley, and N. Turok, Phys. Rev. D **62**, 083508 (2000).

[28] http::/www.cmbfast.org/.

[29] http::/camb.info/.

[30] E.F. Bunn, and M. White, ApJ **480**, 6 (1997).

[31] A. Mennella et al., in Proceedings of International Symposium on Plasmas in the Laboratory and in the Universe: new insights and new challenges, [astro-ph/0310058] (2003).

[32] M. Seiffert et al., A&A **391**, 1185 (2002).

[33] J.M. Lamarre et al., in Proceedings of the Workshop on "The Cosmic Microwave Background and its Polarization", ed. S. Hanany & R.A. Olive, New Astronomy Reviews, [astro-ph/0308075] (2003).

[34] P.L. Richards, J. Appl. Phys. **76**, 1 (1994).

[35] S. Hanany, A.H. Jaffe, and E. Scannapieco, MNRAS **299**, 653 (1998).

[36] D.M. Brink, and G.R. Satchler, *Angular Momentum*, (Oxford University Press, Oxford, 1968).

[37] T. Risbo, Journal of Geodesy **70**, 383 (1996).

[38] C. Burigana et al., A&AS **130**, 551 (1998).

[39] P. Fosalba, O. Doré, and F.R. Bouchet, Phys. Rev. D **65**, 063003 (2002).

[40] S. Mitra, A.S. Sengupta, and T. Souradeep, Phys. Rev. D **70**, 103002 (2004).

[41] J.D. Kraus, *Radio Astronomy*, (University of New Hampshire Printing Services, Durham, 2004).

[42] J.D. Jackson, *Classical Electrodynamics*, (John Wiley & Sons, New York, 1975).

[43] M. Reinecke, K. Dolag, R. Hell, M. Bartelmann, and T. Enßlin, submitted to A&A, [astro-ph/0508522] (2005).

[44] X. Dupac, and J. Tauber, A&A **430**, 363 (2005).

[45] http::/www.physics.montana.edu/faculty/cornish/lagrange.html.

[46] http::/mathworld.wolfram.com/MollweideProjection.html.

[47] D. Scott, and G.F. Smoot, in "The Review of Particle Physics", S. Eidelman et al., Phys. Lett. **B.592**, 1 (2004).

[48] http::/map.gsfc.nasa.gov/.

[49] M.G. Hauser, and E. Dwek, ARAA **39**, 249 (2001).

[50] http::/healpix.jpl.nasa.gov/.

[51] K.M. Górski, E. Hivon, and B.D. Wandelt, in Proceedings of the MPA/ESO Cosmology Conference "Evolution of Large-Scale Structure", ed. A.J. Banday, R.S. Sheth, and L. Da Costa, PrintPartners Ipskamp, NL, 37, [astro-ph/9812350] (1999).

[52] K.M. Górski et al., ApJ **622**, 759 (2005).

[53] W.H. Press, S.A. Teukolsky, W.T. Wetterling, and B.P. Flannery, *Numerical Recipes*, (2nd Edition, Cambridge University Press, 1992).

[54] M. Tegmark, ApJ **480**, L87 (1997)

[55] O. Doré, R. Teyssier, F.R. Bouchet, D. Vibert, and S. Prunet, A&A **374**, 358 (2001).

[56] E.L. Wright, Report UCLA-ASTRO-ELW-96-03, November, [astro-ph/9612006] (1996).

[57] D. Yvon, and F. Mayet, A&A **436**, 729 (2005).

[58] J. Borrill, in Proceedings of the 5th European SGI/Gray MPP Workshop, Bologna, Italy, [astro-ph/9911389] (1999).

[59] http::/crd.lbl.gov/ cmc/MADmap/doc/.

[60] P. Natoli, G. de Gasperis, C. Gheller, and N. Vittorio, A&A **372**, 346 (2001).

[61] G. de Gasperis, A. Balbi, P. Cabella, P. Natoli, and N. Vittorio, A&A **436**, 1159 (2005).

[62] P.G. Ferreira, and A.H. Jaffe, MNRAS **312**, 89 (2000).

[63] S. Prunet, C.B. Netterfield, E. Hivon, and B.P. Crill, in "Energy densities in the Universe", ed. J. Bartlett, and J. Dumarchez, Editions Frontieres, Paris, [astro-ph/0006052] (2000).

[64] P. Natoli, D. Marinucci, P. Cabella, G. de Gasperis, and N. Vittorio, A&A **383**, 1100 (2002).

[65] A. Amblard, and J.-Ch. Hamilton, A&A **417**, 1189 (2004).

[66] C.B. Netterfield et al., ApJ **571**, 604 (2002).

[67] R. Stompor et al., Phys. Rev. D **65**, 022003 (2002).

[68] M. Bersanelli et al., COBRAS/SAMBA Report on the Phase A Study, D/SCI(96)3, ESA (1996).

[69] C. Burigana et al., Int. Rep. TeSRE/CNR, 198/1997, November [astro-ph/9906360] (1997).

[70] J. Delabrouille, A&AS **127**, 555 (1998).

[71] D. Maino et al., A&AS **140**, 383 (1999).

[72] B. Revenu et al., A&AS **142**, 499 (2000).

[73] D. Maino, C. Burigana, K.M. Górski, N. Mandolesi, and M. Bersanelli, A&A **387**, 356 (2002).

[74] M.A. Janssen et al., Report PSI-96-01, FIRE-96-01 [astro-ph/9602009] (1996).

[75] F. van Leeuwen et al., MNRAS **331**, 975 (2002).

[76] A. Mennella et al., A&A **384**, 736 (2002).

[77] E. Keihänen, H. Kurki-Suonio, and T. Poutanen, MNRAS **360**, 390 (2005).

[78] E. Keihänen et al., in preparation (2005).

[79] M.A.J. Ashdown et al., in preparation by Planck Working Group 3 (2005).

[80] C. Burigana, private communications (2003).

[81] C. Armitage, and B.D. Wandelt, Phys. Rev. D **70**, 123007 (2004).

[82] B.D. Wandelt, and K.M. Górski, Phys. Rev. D **63**, 123002 (2001).

[83] D.A. Varshalovich, A.N. Moskalev, and V.K. Khersonskii, *Quantum Theory of Angular Momentum*, (World Scientific Publishing, Singapore, 1988).

[84] K.M. Górski, ApJ **430**, L85 (1994).

[85] M. Tegmark, Phys. Rev. D **55**, 5895 (1997).

[86] J.R. Bond, A.H. Jaffe, and L. Knox, Phys. Rev. D **57**, 2117 (1998).

[87] S.P. Oh, D.N. Spergel, and G. Hinshaw, ApJ **510**, 551 (1999).

[88] O. Doré, L. Knox, and A. Peel, Phys. Rev. D **64**, 083001 (2001)

[89] M. Tegmark, A.N. Taylor, and A.F. Heavens, ApJ **480**, L22 (1997).

[90] A. Stuart, K. Ord, and S. Arnold, *Kendall's Advanced Theory of Statistics: Volume 2A - Classical Inference and the Linear Model*, (Oxford University Press, New York, 1999).

[91] M. Tegmark, and A. de Oliveira-Costa, Phys. Rev. D **64**, 063001 (2001).

[92] D.J. Mortlock, A.D. Challinor, and M.P. Hobson, MNRAS **330**, 405 (2002).

[93] E. Hivon et al., ApJ **567**, 2 (2002).

[94] G. Efstathiou, MNRAS **349**, 603 (2004).

[95] B.D. Wandelt, E.F. Hivon, and K.M. Górski, Phys. Rev. D **64**, 083003 (2001).

[96] W.C. Jones et al., submitted to ApJ, [astro-ph/0507494] (2005).

[97] G. Hinshaw et al., ApJS **148**, 135 (2003).

[98] F.K. Hansen, K.M. Górski, and E. Hivon, MNRAS **336**, 1304 (2002).

[99] F.K. Hansen, and K.M. Górski, MNRAS **343**, 559 (2003).

[100] A. Balbi, G. de Gasperis, P. Natoli, and N. Vittorio, A&A **395**, 417 (2002).

[101] I. Szapudi, S. Prunet, D. Pogosyan, A.S. Szalay, and J.R. Bond, ApJ **548**, L115 (2001).

[102] I. Szapudi, S. Prunet, and S. Colombi, ApJ **561**, L11 (2001).

[103] G. Chon, A. Challinor, S. Prunet, E. Hivon, and I. Szapudi, MNRAS **350**, 914 (2004).

[104] M. Tristram, J.F. Macías-Pérez, C. Renault, and D. Santos, MNRAS **358**, 833 (2005).

[105] Górski, K.M., Wandelt, B.D., Hivon, E., Hansen, F.K., & Banday, A.J., "The HEALPix Primer" (Version 2.00), available in http://healpix.jpl.nasa.gov (2005).

[106] T. Souradeep, and B. Ratra, ApJ **560**, 28 (2001).

[107] J.H.P. Wu et al., ApJS **132**, 1 (2001).

[108] R. Stompor, and M. White, A&A **419**, 783 (2004).

[109] G. Efstathiou, MNRAS **356**, 1549 (2005).

[110] D. Scott, M. Srednicki, and M. White, ApJ **421**, L5 (1994).

[111] M.P. Hobson, and J. Magueijo, MNRAS **283**, 1133 (1996).