Cosmic Microwave Background - Data Analysis

Komal Gupta

10 November 2018

Abstract

Data recorded during an experiment to measure CMB temperature undergoes several analysis steps before it yields information about cosmological parameters. The first step involves preprocessing the data to construct a time ordered data set, which is used to construct a map of the temperature over the sky. Foreground signals are then removed from the map, and so is the CMB dipole, so that the map contains only the temperature of CMB anisotropies. The map is used to estimate the power spectrum of CMB, which is finally used to estimate cosmological parameters. This report details several commonly used map-making techniques, followed by a brief discussion on power spectrum estimation. The present discussion focuses only on temperature maps of CMB; complete analysis involves determination of both temperature and polarization maps, which is skipped.

Contents

1 Introduction

The Cosmic Microwave Background is the leftover radiation from a period early in the history of the universe. As the universe expanded, it cooled, and once the temperature was low enough, electrons and protons could combine to form hydrogen and release photons during the era of recombination. These photons have traveled unimpeded through the universe since, and make up the CMB that we see today. Due to the expansion of the universe, the wavelength of these photons has increased since they were released, and currently lies in the microwave band. The discovery of CMB is in support of the Big Bang model of the universe, and has ruled out the Steady state model completely.

On larger scales, the CMB is uniform across the sky. However, temperature anisotropies exist at smaller scales, which are imprints of initial density fluctuations that gave rise to the large scale structures that we see today. These tiny fluctuations therefore provide a way to study the early conditions of the universe. Our efforts to measure these anisotropies began in 1989, when the Far-Infrared Absolute Spectrophotometer (FIRAS) instrument aboard the Cosmic Background Explorer (COBE) satellite confirmed the CMB to be a thermal black body spectrum at a temperature of 2.725 K[1]. Since then, a number of experiments have been conducted to measure the anisotropies on smaller scales.

NASA's Wilkinson Microwave Anisotropy Explorer (WMAP) mission provided accurate measurements of large angular scale fluctuations in the CMB during its nine years of operation from July 2001 to October 2010. In May 2009, ESA launched the Planck space observatory, which measured the temperature and polarization of CMB with a higher sensitivity and angular resolution compared to WMAP. The aim of all such experiments is to measure cosmological parameters which is done in a series of steps as shown in figure 1.

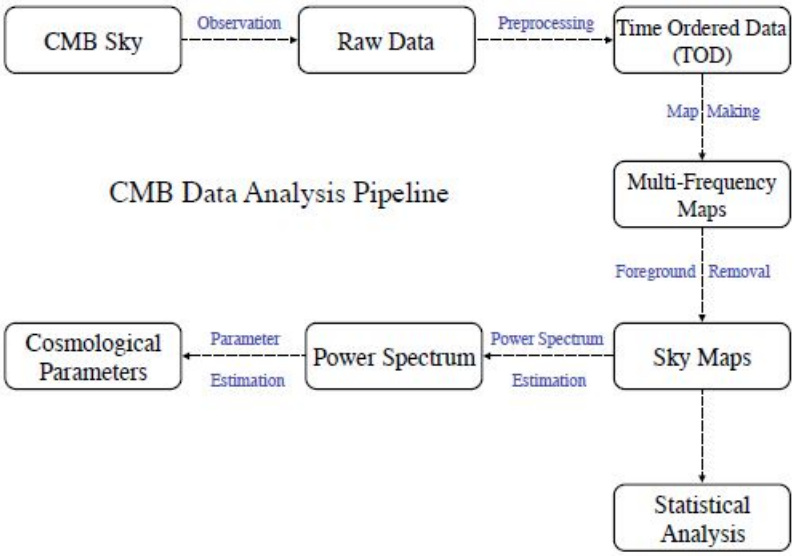

Figure 1: CMB data analysis pipeline. Source:[1]

The following chapters describe a part of the data analysis pipeline, i.e., the map-making process. Further analysis involves estimating CMB power spectrum from temperature and polarization maps, and finally deriving cosmological parameters from the power spectrum.

2 Collecting and modeling data

Experiments to measure CMB anisotropies center around measurement of temperature and polarization of the radiation. This is done by scanning the sky in a systematic manner; the exact method of scan depends on the goals of the experiment.

2.1 Scan Strategy

The scanning strategy for a particular spacecraft depends heavily upon the observation time, which is often limited by resources. The scanning strategy aims to minimize foreground contributions and atmospheric effects while allowing a decent measurement of the power spectrum. This section details the scanning strategies of two of the most important missions to measure CMB anisotropies: The WMAP mission and the Planck mission. Both of these spacecrafts orbited the second Lagrange point of the Earth-Sun system (L2), approximately 1.5 million km from Earth. The point was chosen because it is on Earth's night side and about four times further away from the Earth than the Moon ever gets, therefore allowing the spacecraft to avoid unwanted radiation from the Earth, Moon and Sun which would otherwise interfere with signal from the CMB.

WMAP Mission

The WMAP scanning strategy was chosen in order to scan a large fraction of the sky as rapidly as possible, consistent with control hardware and telemetry data rate. The WMAP was a differential experiment, in the sense that it measured the difference in temperature between two points a fixed distance apart on the sky, large enough to maintain sensitivity to signal at large angular scales. This was important for comparing WMAP results to its predecessor, COBE[3]. The scan strategy is shown in figure 3.

Planck Mission

The scan strategy for Planck was chosen to achieve the survey goals in terms of sky coverage and scanning directions. As Planck spins on its axis it observes a ring around the sky. Over the course of the year, the axis around which Planck spins moves around the sky, ensuring that the satellite is never looking towards the Sun or the Earth[6]. The scan strategy is shown in figure 2.

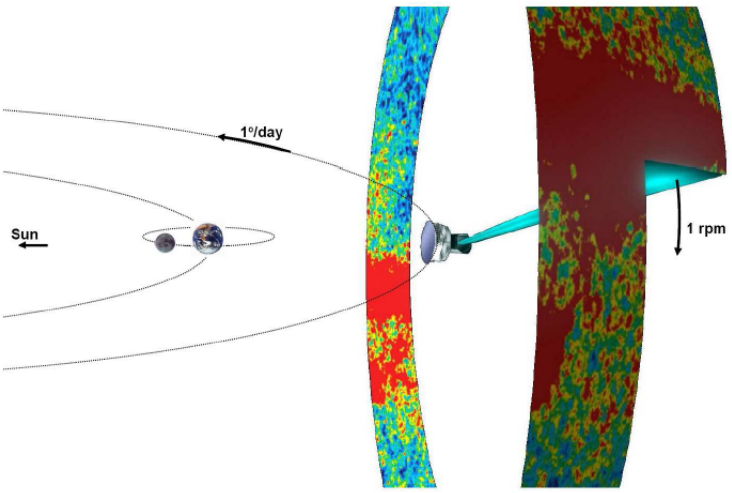

Figure 2: Planck scanning strategy. Source: [4]

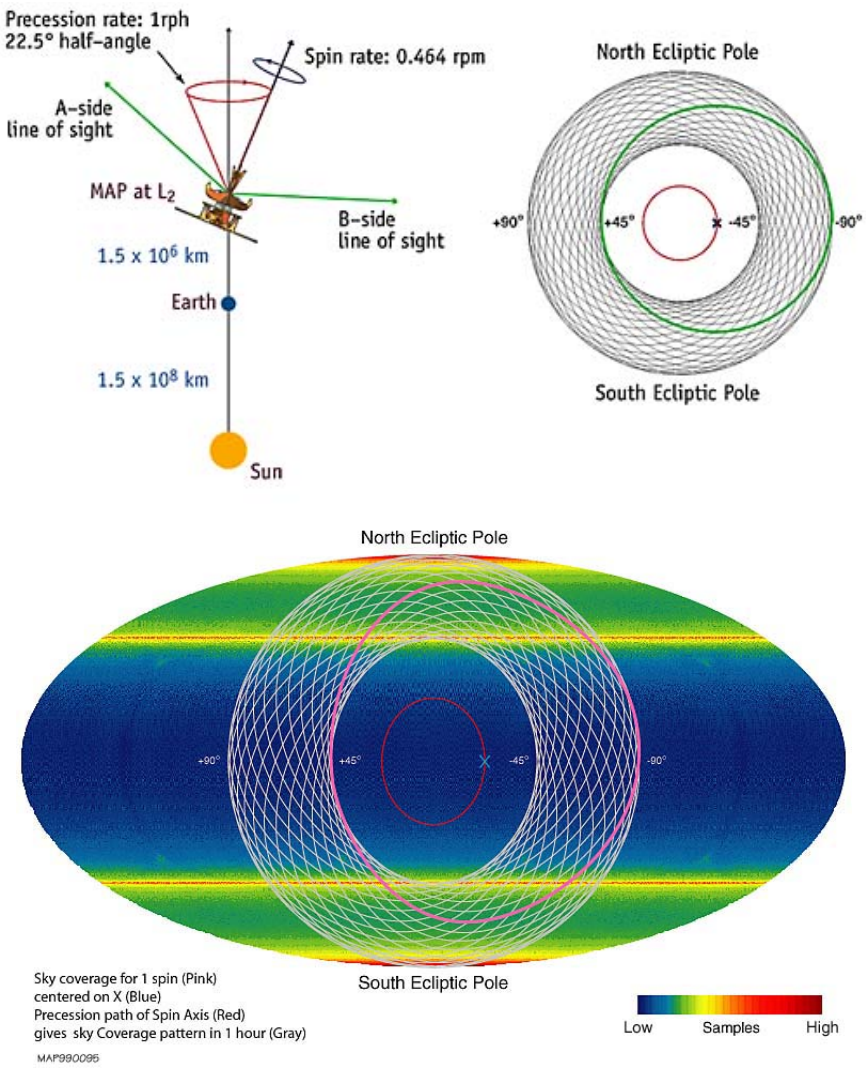

Figure 3: WMAP scanning strategy. Source: [3]

2.2 From Raw to Time Ordered Data

The raw data measured by a satellite is first preprocessed, cleaned, and checked for systematic errors. The output is a list of all positions and temperatures of all pixels observed in chronological order, known as time ordered data set (TOD). The exact constituents of the TOD depend on the scanning strategy of the experiment. In principle, the cosmological parameters can be determined directly from the TOD, which is unfeasible for large data sets[1]. It is therefore necessary to reduce the TOD to a smaller data set, the map.

2.3 Data Model

The TOD is generally modeled as

$$d = A \cdot T + n \tag{1}$$

where $\mathbf{d}$ is the time ordered data (with n_t elements corresponding to each time sample), $\mathbf{T}$ is vector representation of the map (with n_p elements representing the number of sky pixels), and $\mathbf{A}$ is a $n_t \times n_p$ matrix that relates each time sample to corresponding pixel in the sky, called the pointing matrix. For detectors that measure only the temperature and not the polarization of the CMB, the an element T_p

of the vector **T** is simply the temperature of the sky in pixel p. In case of detectors which measure polarization, each element T_p is a tuple containing three Stokes' parameters: I, Q and U. The vector **n** represents the noise added to the signal by the detector.

Noise

The presence of instrumental noise in data is unavoidable, and needs to be accounted for in the data model. The noise in TOD is assumed to be normally distributed with a covariance matrix N given by

$$N =< n \cdot n^t > \tag{2}$$

The detector noise is correlated in time. Major simplifications in the map-making problem are achieved by making the following two assumptions about noise:

Gaussianity. The probability density function of multivariate Gaussian distribution (equation (4)) can be expressed in terms of the covariance matrix. Therefore all the statistical information about noise with Gaussian distribution is contained in its covariance matrix.

Piece-wise stationarity. Stationarity implies that the probability distribution function does not change in time, i.e., the mean and variance of the distribution remain the same over time. The continuous time PDF can then be decomposed in Fourier coefficients. Therefore all the statistical information about the noise is also contained in its Fourier power spectrum.

3 Map making techniques

The process of making a map involves reduction of TOD to a smaller data set while retaining all the cosmological information, so that the parameters can be determined as accurately from the map as from the TOD. The problem is therefore to find best estimate ($\hat{\mathrm{T}}$) of T from equation (1), where d and A are known. This section explores some of the commonly used map-making techniques.

3.1 Simple map-making method

The simplest map making technique is where the detector noise is assumed to be "white". For white noise, the samples are uncorrelated, and we can simply average the data falling in each pixel. The map thus obtained is

$$\hat{\mathrm{T}} = [A^t \cdot A]^{-1} \cdot A^t \cdot d \tag{3}$$

Here, A^t projects elements of d onto the correct map pixel, while $A^t \cdot A$ is a diagonal matrix (of size $n_p \times n_p$), with the value of each diagonal element being the number of time ordered samples which have fallen into the corresponding pixel. The map obtained in this manner is optimal if the noise is white, which is not true in general. Computationally, this method is fast ($\propto n_t$), simple and robust. More complex map-making schemes that take the correlated noise in account are explained in the following subsections.

3.2 Maximum likelihood method

In this map-making scheme, the likelihood of data is maximized for a given noise model. The probability distribution of noise is given by the n_t dimensional Gaussian probability density function:

$$P(n) = \frac{1}{|(2\pi)^{n_t} N|^{1/2}} exp\left[-\frac{1}{2} n^t \cdot N^{-1} \cdot n\right] \tag{4}$$

where N is the noise covariance matrix of equation (2).
From eqn (1), we have $n = d - A \cdot T$, and $P(n) = P(T|d)$. Finally, from Bayes' theorem

$$P(T|d) \propto P(d|T) \propto \frac{1}{|(2\pi)^{n_t} N|^{1/2}} exp\left[-\frac{1}{2}(d - A \cdot T)^t \cdot N^{-1} \cdot (d - A \cdot T)\right] \tag{5}$$

Maximizing this probability w.r.t. T, i.e., $\frac{\partial \chi^2}{\partial T} = 0$, where $\chi^2 = (d - A \cdot T)^t \cdot N^{-1} \cdot (d - A \cdot T)$, we obtain the linear equation

$$A^t \cdot N^{-1} \cdot A \cdot \hat{T} = A^t \cdot N^{-1} \cdot d \tag{6}$$

Here, the property $\frac{\partial (x^t B x)}{\partial x} = 2Bx$ has been used where B is a symmetric square matrix and x is a column vector. The solution to the above equation is

$$\hat{\mathrm{T}} = \left(A^t \cdot N^{-1} \cdot A\right)^{-1} \cdot A^t \cdot N^{-1} \cdot d \tag{7}$$

Therefore, to determine $\hat{\mathrm{T}}$ via maximum likelihood method, one needs to find the inverse matrices $\left(A^t \cdot N^{-1} \cdot A\right)^{-1}$ and N^{-1}, both of size $n_t \times n_t$. These inverses are difficult to calculate computationally, particularly for large matrices, since the complexity of matrix inversion operation scales as $O(n^3)$ for a $n \times n$ matrix. Two methods can be adopted to calculate the inverses: a brute force inversion of the matrices using large parallel computers, or an iterative approach to the solution, using processes like Jacobi iterator. For large data sets such as WMAP and Planck, iterative methods are preferred and have been applied successfully using codes such as MapCUMBA, ROMA and Mirage etc.

3.3 Destriping methods

The aforementioned method of optimal map-making is complex in terms of the computational resources required to analyze large data sets. Destriping techniques are used to remove noise easily and efficiently by dividing it into a low-frequency correlated noise component and an uncorrelated white noise component. The TOD is divided into n_b segments of equal length n_{base}, such that $n_t = n_b n_{base}$. The baselines model the correlated noise, also known as "1/f noise" because it is inversely proportional to frequency of observation. The model now becomes

$$d = A \cdot T + \Gamma \cdot x + w \tag{8}$$

where $\mathbf{x}$ is a vector of length n_b containing the baseline amplitudes, and $\mathbf{\Gamma}$ is a matrix of size $n_t \times n_b$ that spreads the baseline amplitudes over TOD vector. The vector $\mathbf{w}$ represents uncorrelated white Gaussian noise. The white noise covariance matrix $C_w =< w \cdot w^t >$, is thus diagonal. One can note a major simplification compared to the previous method of maximum likelihood, where the noise covariance matrix is non-diagonal in general.
As in the previous case, the problem is to find maximum likelihood map $\mathbf{T}$. Applying Bayes' theorem as before (equation (5)), given the map $\mathbf{T}$ and baseline amplitudes $\mathbf{x}$, the probability of data $\mathbf{d}$ is proportional to n_t dimensional Gaussian probability density function:

$$P(d|T, x) \propto \frac{1}{|(2\pi)^{n_t} C_w|^{1/2}} exp\left[-\frac{1}{2} w^t \cdot C_w^{-1} \cdot w\right] \tag{9}$$

Maximizing this probability is equivalent to minimizing the function

$$\chi^2 = (d - A \cdot T - \Gamma \cdot x)^t \cdot C_w^{-1} \cdot (d - A \cdot T - \Gamma \cdot x) \tag{10}$$

Similar to the analysis done in the previous section, minimization with respect to $\mathbf{T}$ (i.e., $\frac{\partial \chi^2}{\partial T} = 0$) gives the maximum likelihood map

$$\hat{T} = \left(A^t \cdot C_w^{-1} \cdot A\right)^{-1} \cdot A^t \cdot C_w^{-1} \cdot (d - \Gamma \cdot x) \tag{11}$$

$$M = A^t \cdot C_w^{-1} \cdot A \tag{12}$$

M is a symmetric matrix. Substituting equation (11) back into equation (10) we get

$$\chi^2 = (d - \Gamma \cdot x)^t \cdot Z^t \cdot C_w^{-1} \cdot Z \cdot (d - \Gamma \cdot x) \tag{13}$$

where

$$Z = I - A \cdot M^{-1} \cdot A^t \cdot C_w^{-1} \tag{14}$$

Here, $\mathbf{Z}$ is a projection matrix ($Z^2 = Z$). In general, $\mathbf{Z}$ is symmetric, and so is $C_w^{-1}Z$ (since C_w is a diagonal matrix), so that

$$C_w^{-1}Z = Z^t C_w^{-1} = Z^t C_w^{-1} Z \tag{15}$$

We minimize equation (13) with respect to $\mathbf{x}$ (i.e., $\frac{\partial \chi^2}{\partial x} = 0$) to obtain (using equation (14) and (15)) an equation of the form

$$(\Gamma^t \cdot C_w^{-1} \cdot Z \cdot \Gamma) \cdot x = \Gamma^t \cdot C_w^{-1} \cdot Z \cdot d \tag{16}$$

and define

$$D = \Gamma^t \cdot C_w^{-1} \cdot Z \cdot \Gamma \tag{17}$$

Again, D is a symmetric matrix. The maximum likelihood estimate of $\mathbf{x}$ thus becomes

$$\hat{x} = D^{-1} \cdot \Gamma^t \cdot C_w^{-1} \cdot Z \cdot d \tag{18}$$

and that for $\mathbf{T}$ becomes

$$\hat{T} = M^{-1} \cdot A^t \cdot C_w^{-1} \cdot (d - \Gamma \cdot \hat{x}) \tag{19}$$

Equations (18) and (19) summarize the method. This method also requires calculation of two inverses: M^{-1} and D^{-1}, where M is a $n_p \times n_p$ matrix and D is a $n_b \times n_b$ matrix (C_w is a diagonal matrix, calculating its inverse is simply finding reciprocal of all diagonal elements and a trivial problem). However, because both n_p and n_b are smaller than n_t, and because the computational complexity of matrix inversion scales as $O(n^3)$, inverting M and D is faster than inverting $A^t \cdot N^{-1} \cdot A$ and N (which had to be computed in the previous method). This method is therefore preferred because it is faster and more accurate compared to the previously discussed method of maximum likelihood analysis. Moreover, this method is better in removing 1/f noise, which heavily contaminates low-frequency maps. This is only the algebra behind basic destriping; actual implementation involves more analysis and information about the noise spectrum and has been omitted. The following figure shows simulated full sky CMB map before and after destriping. Before destriping, the map is contaminated by stripes due to 1/f noise.

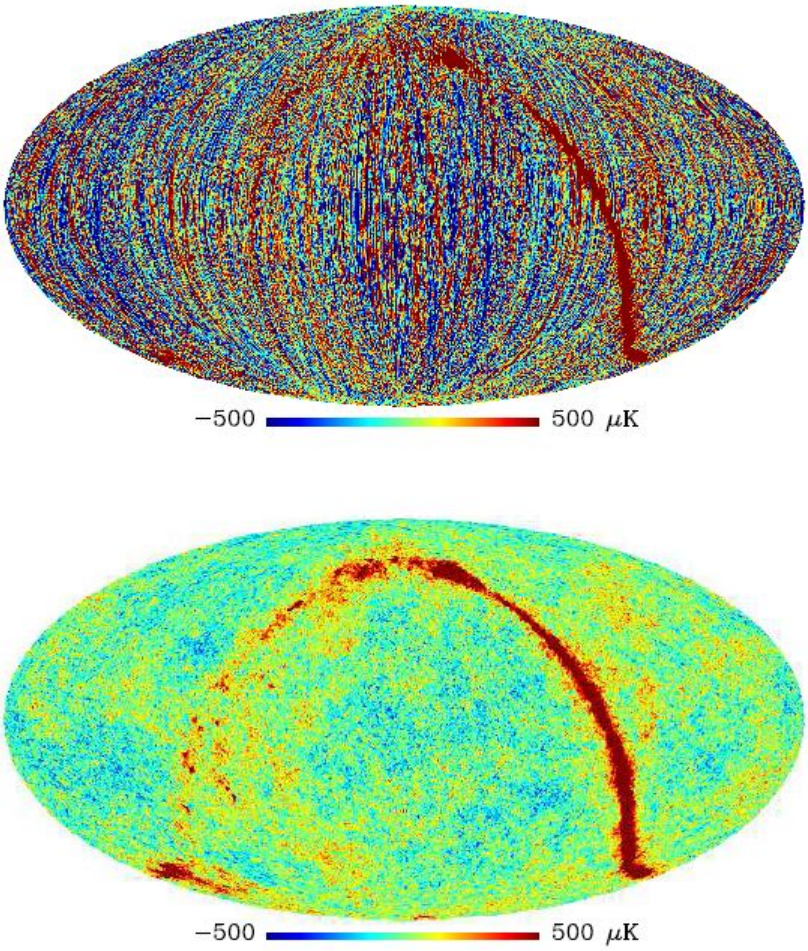

Figure 4: CMB map before (top) and after (bottom) destriping. Source: [8]

4 Foreground Removal

The map ($\hat{T}$) obtained from the aforementioned methods contains many other signals apart from the CMB itself. While measuring the anisotropies in CMB, the most prominent of the foreground effects is the CMB dipole, caused by the Doppler shift of CMB photons due to motion of the Sun relative to the rest frame of the CMB. The motion of the Earth around the Sun also contributes to the foreground significantly. However, these dynamics are well understood and therefore these effects can be removed from the map with high degree of precision. Other astrophysical signals can be major contaminants to CMB, for example:

Synchrotron emission. Synchrotron radiation is produced by relativistic electrons accelerated by a magnetic field. This effect can be produced by the magnetic field of our galaxy, apart from other sources such as pulsars. The emission spectrum is approximated by power law of the form $\sim \nu^{-3}$, making synchrotron radiation a major contaminant to CMB observations at low frequencies.

Bremsstrahlung (free-free) radiation. Also known as breaking radiation, Bremsstrahlung radiation is produced when electrons are decelerated in the presence of ions. Major sources are the ionized H II regions in our neighborhood. Again, the emission spectrum is approximated by a power law of the form $\sim \nu^{-2.1}$, and contaminates low frequency observations.

Galactic dust emission. Cold dust within our galaxy can also absorb and emit electromagnetic radi-

ation. Their spectrum is modeled as $\sim \nu^2$, and dominates at high frequencies

The following figure shows these foreground spectra compared to CMB.

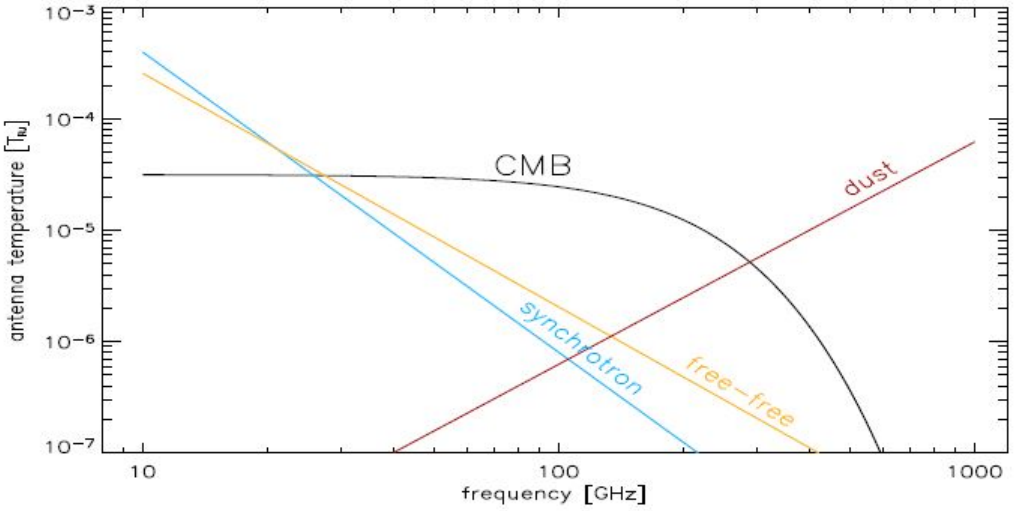

Figure 5: Foreground spectra. Source: [7]

In case of ground based experiments to measure CMB temperature, regions with excessive foreground emission are avoided while scanning the sky. In case of satellite based experiments which cover the whole sky, unavoidable foreground signals are removed by making use of the fact that these signals have quite different spatial and spectral distributions compared to the CMB.

4.1 Linear combination

Since the CMB temperature remains constant over a range of frequencies, multi-frequency maps can be linearly combined to remove some foreground signals (such as galactic signals) while preserving the CMB. This method is easy and efficient but results in a single CMB map with complicated noise properties which are difficult to analyze.

4.2 Template removal

Templates for known foreground signals are constructed from observations and removed from the CMB. The observations are conducted at a particular frequency and extrapolated over a range of frequencies. However, extrapolation in frequency leads to residual foreground uncertainties that remain after these templates have been fit and subtracted. Despite this limitation, template fitted maps have been widely used in experiments such as WMAP.

This form of component separation takes place in the pixel domain, in the sense that we model the foreground signals in each pixel from observation and remove them. Another form of component separation is possible, one which takes place in spherical harmonics domain, and relies on the fact that CMB and foregrounds have quite different electromagnetic spectra. The use of a particular technique for foreground removal depends on the type of emission to be removed, as well as the computational resources available at hand.

5 Power spectrum estimation

Once we have removed the foreground, the map can be used to calculate the power spectrum (C_l). The map can be decomposed as

$$T = s + n \tag{20}$$

where **s** is the signal and **n** is the noise on map pixels. The spherical harmonics, Y_{lm}, form a orthogonal basis on a sphere. The signal in a pixel p can be expanded in spherical coordinates as

$$s_p = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} B_l Y_{lm}(\theta_p, \phi_p) \tag{21}$$

where B_l stands for the Legendre transform of the instrumental beam assuming it is symmetric. Assuming that the CMB signal is Gaussian, the variance of the spherical harmonic coefficients a_{lm}, defined as angular power spectrum and denoted by C_l, contains all the cosmological information.

$$< a_{lm} a^*_{l'm'} >= C_l \delta_{ll'} \delta_{mm'} \tag{22}$$

The covariance matrix of the map is given by

$$M =< T \cdot T^t >=< s \cdot s^t > + < n \cdot n^t >= S + N \tag{23}$$

The signal part is given by

$$S_{pp'} =< s_p s_{p'} >= \sum_l \frac{2l+1}{4\pi} C_l B_l^2 P_l(cos\theta_{pp'}) \tag{24}$$

where $\theta_{pp'}$ is the angle on the sphere between pixels p and $p^{'}$. Equation (23) now becomes

$$M_{pp'} = \sum_l \frac{2l+1}{4\pi} C_l B_l^2 P_l(cos\theta_{pp'}) + N_{pp'} \tag{25}$$

The problem of power spectrum estimation thus reduces to finding C_l from T and N using equation (25). Solving this, the power spectrum is obtained as shown in the following plot.

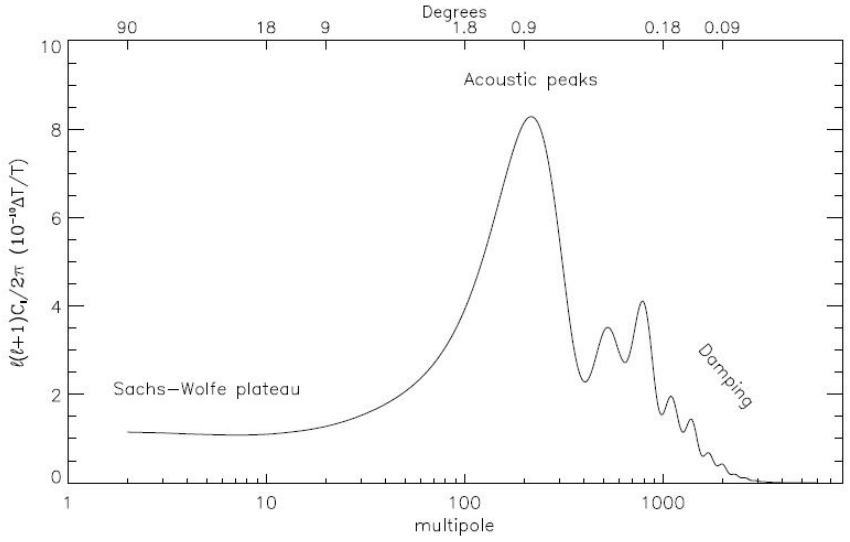

Figure 6: Temperature angular power spectrum. Source: [7]

Finally, one can use the power spectrum to estimate cosmological parameters, such as baryon density (Ω_b), matter density (Ω_m), Hubble Constant (H_0) etc. We skip this discussion for now.

6 Conclusions

The preceding chapters detail several techniques for making temperature maps of the CMB, which are used for estimating cosmological parameters. We see that of the three methods commonly used to make temperature maps of the CMB, destriping method is the most preferred because it is not only faster but it also uses a better model of noise, leading to more precise maps. The map-making process is followed by foreground removal to remove all astrophysical signals other than the CMB. The choice of foreground removal technique largely depends on the signal to be removed. However, the present discussion includes only the theoretical background at each step. The development of algorithms to carry out the calculations and their implementation on large parallel computers is a different topic in its own right, and has been skipped from this report. Once a map of the sky is obtained, power spectrum can be estimated, which is used to finally estimate the cosmological parameters. This report focuses only on map-making process, and these steps have been omitted.

Acknowledgments

I wish to extend my sincerest gratitude to Dr. Kinjal Banerjee, Department of Physics, BITS, Pilani - Goa Campus, for providing me this opportunity to work on this topic and for his constant guidance and support throughout the course of this project.

References

[1] P. Paykari and J.-L. Starck, *CMB Data Analysis*, 2011.

[2] J.-Ch. Hamilton, *CMB map-making and power spectrum estimation*, 2003.

[3] WMAP Observatory: Scan Strategy, *https://wmap.gsfc.nasa.gov/mission/observatory_scan.html.*

[4] Survey scanning and performance, *https://wiki.cosmos.esa.int/planckpla/index.php/Survey_scanning_and_performa*

[5] Journey/Planck/SpaceScience/Our Activities/ESA, *http://www.esa.int/Our_Activities/Space_Science/Planck/Jour*

[6] Planck scanning the Sky | Planck Mission, *http://planck.cf.ac.uk/mission/scanning.*

[7] M. Tristram and K. Ganga, *Data Analysis Methods for the Cosmic Microwave Background*, 2008.

[8] Keihänen et. al., *Making CMB temperature and polarization maps with Madam*, 2018.