

An Introduction to the Cosmic Microwave Background

Kaustuv Basu

kbasu@uni-bonn.de



eCampus | Lernplattform der Universität Bonn

astro8405: The Cosmic Microwave Background

Aktionen 🗸

This course intends to give you a modern and up-to-date introduction to the science and experimental techniques relating to the Cosmic Microwave Background. No prior knowledge of cosmology is necessary, your prerequisite are a basic understanding of electrodynamics and thermal physics and some familiarity with Python programming.

Lecture 09:

CMB Foregrounds and Component Separation

The "CMB sky"



The "CMB sky" ca. 90-s



The "CMB sky" ca. 2020



Planck collaboration (2015)

CMB foregrounds in temperature



CMB foregrounds in polarization



Components of the microwave sky



Components of the microwave sky



Credit: WebSky simulations (Stein et al. 2020)

SED of a typical galaxy



Bremmstrahlung (free-free) emission

Bremsstrahlung,, or 'braking-radiation', also known as free-free emission, is produced by collisions between particles in hot ionized plasmas. We generally study this in optically thin cases (e.g. in the X-rays), but in the radio regime the emission can be optically thick.



A hot plasma of ionized atoms emits radiation through the Coulomb collisions of the electrons and ions. The electrons experience large accelerations in the collisions and thus efficiently radiate photons, which escape the plasma if it is optically thin. The energy Q radiated in a single collision is obtained from Larmor's formula. The characteristic frequency of the emitted radiation is estimated from the duration of the collision, which, in turn, depends on the electron speed and its impact parameter (projected distance of closest approach to the ion). Multiplication of Q by the electron flux and ion density and integration over the range of speeds in the Maxwell–Boltzmann distribution yield the volume emissivity $j_{\nu}(\nu)$ (W m⁻³ Hz⁻¹), the power emitted from unit volume into unit frequency interval at frequency ν as a function of frequency. It is proportional to the product of the electron and ion densities and is approximately exponential with frequency. A slowly varying Gaunt factor modifies the spectral shape somewhat. Most of the power is emitted at frequencies near that specified by $h\nu \approx kT$.

Free-free emission from our Galaxy



Bremsstrahlung emission spectrum

For thermal plasma, the velocity distribution $\mathcal{P}(v) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right)$ is Maxwell–Boltzmann:

Leading to the well-known expression for free-free emission

$$j_{\nu}(\nu) \,\mathrm{d}\nu = g(\nu, T, Z) \,\frac{1}{(4\pi\varepsilon_0)^3} \frac{32}{3} \left(\frac{2}{3} \frac{\pi^3}{km^3}\right)^{1/2} \frac{Z^2 e^6}{c^3} n_\mathrm{e} n_\mathrm{i} \frac{\mathrm{e}^{-h\nu/kT}}{T^{1/2}} \mathrm{d}\nu.$$
(Volume emissivity; W/m³ at ν in $\mathrm{d}\nu$)



Thermal emission from interstellar dust

Dust grains follow Kirchoff's law for opaque bodies; their emission and absorption coefficients are equal when the grains are in LTE. We define a "grey body" as an emitter whose efficiency is less than 100% from that of a blackbody at any frequency.

In the Rayleigh–Jeans part of the spectrum, emission coefficient $\epsilon(v) \propto v^2$ would mean a blackbody radiation, so an inefficient emitter has $\epsilon(v) \propto v^{2+\beta}$, where $1 < \beta < 2$. Thus the RJ part is "bluer" than a blackbody and signifies optically thin emission.

$$S_{\nu} \sim \nu^{\beta} \cdot B_{\nu}(T_d)$$

 $\beta \sim 1.5$ for interstellar dust at $T_d \sim 20$ K.



Dust emission examples

Planck satellite imaging of the Galactic dust and its polarization pattern (Planck collaboration 2015)

Dust emission from the protoplanetary disk of HL Tau (ALMA data, Brogan et al.)





14

High-z starburst galaxies



Polarized dust emission in our Galaxy



Polarized dust emission originates due to (partial) alignment of the vibrating dust grains in an ambient magnetic field. This creates a preferred orientation of the radiation electric field vector.



This is similar to the polarization of starlight extinction, observed in the optical. It picks up the magnetic field component in the plane of the sky, where the direction of polarization is orthogonal to the magnetic field direction.



Multiple field alignment will also cause *depolarization*.

Images on left: Polarized dust emission in our Galaxy, as measured by the **Planck satellite**. Above: the BICEP-2 field in the south, below: the full sky.

Current knowledge on dust polarization



17

09: Foregrounds & Component Separation

The Cosmic Infrared Background

The Cosmic Infrared Background (CIB) is the thermal dust emission from all the extragalactic objects. Mainly, it is dominated by star-forming galaxies at redshifts $z \sim 1-2$. There are still significant uncertainties on the spectrum and its spatial variations (since we are approximating a very complex process with a simple modified blackbody spectrum of the form $I_v(dust) \sim v^{3+\beta}/(\exp(hv/kT_{dust}) - 1)$.

The separation of the CIB and Galactic dust is done by careful spatial templates.



The Cosmic Infrared Background



Unlike the Galactic dust, the CIB is practically unpolarized (polarization angles gets averaged out).

(Left) Location of the six fields used to study the Cosmic Infrared Background using Planck data

Simulations for the CIB at four of the Planck frequencies



Credit: ESA / Planck Consortium

Anomalous dust emission (spinning dust)



There is yet another type of dust emission, discovered relatively recently (from WMAP data, in the 1990s), that shows up as a continuum emission in the 30–90 GHz frequency range. The most common conjecture is that this is due to **spinning dust** particles! The signal is a weak "bump" of emission around 20–30 GHz.

 $\mathbf{E} \perp \mathbf{B}$

Anomalous dust emission often goes hand-in-hand with the regular dust emission







Anomalous dust emission (spinning dust)



Regions of very significant anomalous dust emission in the Galactic plane (marked by squared on the above figure) and some representative spectra (below). Credit: Planck collaboration (2014).



The exact nature of the AME sources and their abundance in the high Galactic latitudes are not well understood.

Thermal dust emission summary

The thermal emission from interstellar dust grains is a "grey body" or "modified blackbody" spectrum. The same spectrum is used to describe the cosmic infrared background (CIB), where it follows almost the same mean values as the Galactic dust.

$$I_{\nu}^{\text{dust}} = A \left(\frac{\nu}{\nu_0}\right)^{\beta} B_{\nu}(T_d) \qquad \begin{array}{l} A = \text{How much dust?} \\ T_d = \text{How hot is the dust?} \\ \beta = \text{What is the dust made of?} \end{array}$$

Thermal dust emission from the Galaxy is the main foreground at frequencies >100 GHz. This emission is partially polarized, due to the alignment of the dust grains in Galactic magnetic fields. At low frequencies (~30 GHz) there is an additional (and polarized) component of spinning dust emission, roughly confined to star-forming regions in the Galactic plane.

The alignment of dust grains in an ambient magnetic field is a new tool to probe the Galactic field structures, especially in regions far from the Galactic plane. But it is a major headache for cosmologists looking for the gravitational wave signature from inflation, expected to be imprinted on CMB B-modes.

Cyclotron and synchrotron radiation



Charged particles moving in a magnetic field experience an acceleration perpendicular to the direction of their motion.

In the non-relativistic limit ($\beta \approx 0$) we get cyclotron radiation, with symmetric angular dependence of radiation: $(1 - \sin^2\theta \cos^2\varphi)$.





Synchrotron spectrum of a single electron



The spectrum of a single electron is essentially a series of pulses. Each pulse has a finite time width, which is the "sweeping time" of the beam.

$$\Delta t_{\rm p} = \frac{2}{\gamma \omega_B 2\gamma^2} = \frac{1}{\gamma^3 \omega_B} = \frac{1}{\gamma^2 \omega_{\rm G}}$$

 $\omega_{\rm G} \equiv \frac{qB}{mc}$. (Cyclotron angular frequency; rad/s)

Therefore, the frequency spectrum (=Fourier transform of the time pulse) for single electron's emission is a continuum, with a maximum around the frequency



The power is relatively flat below this frequency, with a logarithmic slope of 1/3, and drops off rapidly above.

We can assume most of the radiation coming at this peak frequency v_{max} . Note the $\gamma^2 B_{\perp}$ dependence.

Spectrum from power-law electrons

The most interesting (and astrophysically relevant) application, of course, comes from the radiation when the electrons have a power-law distribution.

The spectrum can be calculated to good accuracy by assuming that each electron radiates all of its available energy, $P = -\frac{dE}{dt} = \frac{4}{3}\sigma_T \beta^2 \gamma^2 c U_B$ at the single frequency $v \approx \gamma^2 v_G$, (which is close to the critical frequency).

This leads to a power-law spectrum if the electron energies themselves follow a power-law distribution.



Synchrotron radiation is polarized

Synchrotron emission is always polarized, because the emitting electrons have a strong directionality forced upon them by the magnetic field.

Radiation from single electrons are elliptically polarized, but in an ensemble of particles, the transverse components cancel out and we get a linear polarization, roughly in a plane perpendicular to the magnetic field.

This fact is utilized also in measuring Faraday rotation, to determine the lineof-sight component of interstellar magnetic fields using a background source.



(a) Electrons emitting polarized synchrotron emission by spiralling along magnetic field lines. (b) Line of sight through a nebula with partially ordered magnetic field lines.

Synchrotron emission from the Galaxy



Synchrotron pathfinders for the B-mode



Preliminary map of all-sky polarization amplitude from the CBASS-North.

CBASS-North experiment in California





CCAT-prime for dust polarization



Figure 8. Effective resolution for SNR > 3 measurements of dust polarized intensity with *Planck* at 353 GHz (left) and FYST/Prime-Cam at 350 GHz (right), for the 4000 hour Wide Field Survey (WFS) at different levels of spatial resolution. Translucent contours indicate the boundaries of the FYST observable region.



CMB foregrounds summary



CMB foregrounds summary



The figure above (from Planck collaboration 2013) summarizes our knowledge of the CMB temperature foregrounds. Left: The frequency spectra of various foregrounds at fixed $\ell = 200$ (position of the first acoustic peak). Right: The angular power spectra of these foregrounds, at fixed v = 100 GHz.

Solid lines mark where the spectra are estimated from data, and dashed lines where they are extrapolated. For Galactic foregrounds, a 2-point Gaussian statistic (power spectrum) would be very far from the actual situation, anyway!

How to remove CMB foregrounds?



Component Separation: In general it's an inversion problem





 $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$ Invert for \mathbf{s}

However, in real life, it is extremely difficult to determine this matrix **A**. But we can make progress even without having as many channels as components or without having detailed information on all the foregrounds.

Component separation: The ILC method

The Internal Linear Combination (ILC) method aims to combine different frequency maps with specific weights, such that contributions from all the contaminating signals (plus instrument noise) are minimized in the sought-after signal (map). This is the principle of a minimum variance unbiased estimator.

This works especially well when we have poor knowledge of the foregrounds. But we must have a precise knowledge of the spectrum of the signal that we're interested in. Also, ILC should be used separately on different spatial scales for better results (using masks or wavelet filters).

The term "internal" refers to the fact that no prior information or auxiliary data from other observations are needed. The ILC method is one of the most assumption-free map making tools available! It only requires the following two assumptions:

- The observed maps represent a linear mixture of astrophysical components and noise.
- All components are uncorrelated.

ILC method formalism (I)

From the assumption that observed maps are a linear mixture of astrophysical components and noise, we can write

 $T_i(p) = a_i s(p) + n_i(p),$

where a_i are the components of a "mixing vector" which contains the spectrum of interest and has as many components as frequencies. $s_i(p)$ and $n_i(p)$ are the signal and noise components in each channel maps. In vector form:

$$\vec{T}(p) = \vec{a}s(p) + \vec{n}(p).$$

By forming a linear combination $s_{ILC}(p) = \vec{\omega}^T \vec{T}(p)$, the ILC method provides an estimation $s_{ILC}(p)$ of the desired signal s(p). Here ω_i are the "weights", or the desired ILC coefficients.

$$s_{\rm ILC}(p) = \vec{\omega}^{\rm T} \vec{a} s(p) + \vec{\omega}^{\rm T} \vec{n}(p).$$

The goal is to find these weights ω_i which will minimize the variance in the reconstructed map, $s_{ILC}(p)$. Following Eriksen (2004) et al., the map variance is calculated from

$$VAR(s_{\rm ILC}(p)) = VAR(s(p)) + VAR(\vec{\omega}^{\rm T}\vec{n}(p)) = \vec{\omega}^{\rm T}\hat{C}\vec{\omega},$$

where **C** is the frequency-to-frequency covariance matrix of the maps (computed empirically from the actual multi-frequency data).

ILC method formalism (II)

The condition of minimization of the variance means that

$$\frac{\partial}{\partial \omega_{\rm i}} \left[\vec{\omega}^{\rm T} \hat{C} \vec{\omega} \right] = 0.$$
 (1)

In addition, we have a normalization condition to preserve the sum total of the signal of interest:

$$\vec{\omega}^{\mathrm{T}}\vec{a} = \vec{a}^{\mathrm{T}}\vec{\omega} = 1.$$

It is straight-forward to solve equations 1 and 2 for the weights, in terms of the mixing vector (usually done by employing Langrange-multiplier technique). The solution is

$$\vec{\omega} = \frac{\hat{C}^{-1}\vec{a}}{\vec{a}^{\mathrm{T}}\hat{C}^{-1}\vec{a}}.$$

Therefore, the estimated map s_{ILC}(p) of our component of interest is given by

$$s_{\mathrm{ILC}}(p) = \frac{\vec{a}^{\mathrm{T}}\hat{C}^{-1}}{\vec{a}^{\mathrm{T}}\hat{C}^{-1}\vec{a}}\vec{T}(p).$$

ILC method formalism (the maths..)

Here are the maths of the intermediate steps. We start by writing each (observed) frequency maps as a linear combination of the signal and noise, written in vector form:

$$T_i(p) = \sum_j A_j s_j(p) + n_i(p), \qquad \vec{T}(p) = \vec{A}s(p) + \vec{n}$$

The ILC method gives an estimate of the component of interest, s(p), as an weighted sum of the observed frequency maps. The variance of this ILC-estimated map is computed as the following:

$$\begin{split} s_{\mathrm{ILC}}(p) &= \vec{\omega}^{\mathrm{T}} \vec{T}(p) \qquad VAR(s_{\mathrm{ILC}}(p)) = VAR(s(p)) + VAR(\vec{\omega}^{\mathrm{T}} \vec{n}(p)) = \vec{\omega}^{\mathrm{T}} \hat{C} \vec{\omega}, \\ \mathrm{Var}(s_{\mathrm{ILC}}(p)) &= \langle s_{\mathrm{ILC}}(p)^{2} \rangle - \langle s_{\mathrm{ILC}}(p) \rangle^{2} \\ &= \frac{1}{N_{pix}} \sum_{p=1}^{N_{pix}} \left[\sum_{i=1}^{N_{obs}} \omega_{i} T_{i}(p) \right]^{2} - \left\{ \frac{1}{N_{pix}} \sum_{p=1}^{N_{pix}} \left[\sum_{i=1}^{N_{obs}} \omega_{i} T_{i}(p) \right] \right\}^{2} \\ &= \sum_{i=1}^{N_{obs}} \sum_{j=1}^{N_{obs}} \omega_{i} \omega_{j} \left[\frac{1}{N_{pix}} \sum_{p=1}^{N_{pix}} T_{i}(p) T_{j}(p) \right] - \left\{ \sum_{i=1}^{N_{obs}} \omega_{i} \left[\frac{1}{N_{pix}} \sum_{p=1}^{N_{pix}} T_{i}(p) \right] \right\}^{2} \\ &= \sum_{i=1}^{N_{obs}} \sum_{j=1}^{N_{obs}} \omega_{i} \omega_{j} \left\{ \frac{1}{N_{pix}} \sum_{p=1}^{N_{pix}} T_{i}(p) T_{j}(p) - \left[\frac{1}{N_{pix}} \sum_{p=1}^{N_{pix}} T_{i}(p) \right]^{2} \right\} \\ &= \vec{\omega}^{\mathrm{T}} \hat{\mathbf{C}} \vec{\omega} \end{split}$$

The weights are computed by solving these two equations:

$$\frac{\delta}{\delta\omega_i} \left[\vec{\omega}^T \hat{C} \, \vec{\omega} \right] = 0, \quad \vec{\omega}^T \vec{A} = \vec{A}^T \vec{\omega} = 1.$$

Solution is obtained by using the Lagrange multiplier technique:

we:
$$\frac{\delta}{\delta\omega_i} \left[\vec{\omega}^T \hat{C} \, \vec{\omega} + \lambda (1 - \vec{\omega}^T \vec{A} \right] = 0$$

$$\begin{pmatrix} 2\hat{C} & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \begin{pmatrix} \vec{\omega} \\ \vec{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad \vec{\omega} = \frac{\hat{C}^{-1}\vec{d}}{\vec{d}^{\mathrm{T}}\hat{C}^{-1}\vec{d}}$$

An Introduction to the CMB

09: Foregrounds & Component Separation 37

Dumb ILC example: signal + noise

We have two maps, with signal and noise. The covariance matrix for observations is

$$\mathbf{C} = \begin{bmatrix} S + N_1 & S \\ S & S + N_2 \end{bmatrix} \qquad \mathbf{C}^{-1} = \frac{1}{\det(\mathbf{C})} \begin{bmatrix} S + N_2 & -S \\ -S & S + N_1 \end{bmatrix}$$

The ILC weights are simply: $\omega_1 = N_2 / (N_1 + N_2)$ and $\omega_2 = N_1 / (N_1 + N_2)$. This is same as weighting each map *I* proportionally to $1/N_i$.



Real ILC example: WMAP CMB all-sky



Planck CMB maps from the ILC method

Planck all-sky CMB maps from variants of the **ILC** method (Planck 2015 results, Diffuse component separation, A&A 594, A9)



Planck component separation techniques



Page Discussion

CMB and astrophysical component maps

NILC [edit]

Principle

The Needlet-ILC (hereafter NILC) CMB map is constructed from all Planck channels from 44 to 857 GHz and includes multipoles up to $\ell = 3200$. It is obtained by applying the Internal Linear Combination (ILC) technique in needlet space, that is, with combination weights which are allowed to vary over the sky and over the whole multipole range.

SMICA [edit]

Principle

SMICA produces a CMB map by linearly combining all Planck input channels (from 30 to 857 GHz) with weights which vary with the multipole. It includes multipoles up to $\ell = 4000$.

SEVEM [edit]

The aim of SEVEM is to produce clean CMB maps at one or several frequencies by using a procedure based on template fitting. The templates are internal, i.e., they are constructed from Planck data, avoiding the need for external data sets, which usually complicates the analyses and may introduce inconsistencies. The method has been successfully applied to Planck simulations^[2] and to WMAP polarisation data^[3]. In the cleaning process, no assumptions about the foregrounds or noise levels are needed, rendering the technique very robust.

COMMANDER-Ruler [edit]

COMMANDER-Ruler is the Planck software implementing a pixel based parametric component separation. Amplitude of CMB and the main diffuse foregrounds along with the relevant spectral parameters for those (see below in the Astrophysical Foreground Section for the latter) are parametrized and fitted in single MCMC chains conducted at $N_{\rm rm}(side)=256$ using COMMANDER, implementing a Gibbs Sampling. The CMB amplitude which is obtained in these runs corresponds to the delivered low resolution CMB component from COMMANDER-Ruler which has a FWHM of 40 arcminutes.

(Example for Coma cluster's SZ signal) **ILC** implementation





(c) 217 GHz



(e) 545 GHz





(a) $\omega_{100} \cdot T_{100}$



(c) \cdots + $\omega_{217} \cdot T_{217}$







(b) \cdots + $\omega_{143} \cdot T_{143}$



 $(d) \cdots + \omega_{353} \cdot T_{353}$



(e) \cdots + $\omega_{545} \cdot T_{545}$

(f) $\cdots + \omega_{857} \cdot T_{857}$

ILC summary: Example using CMB and SZ

$$d(\nu, \vec{\theta}) = a(\nu) \quad s_{CMB}(\vec{\theta}) + b(\nu) \quad s_{TSZ}(\vec{\theta}) + N(\nu, \vec{\theta})$$

$$\underbrace{Hanck frequency}_{maps} \quad \underbrace{CMB}_{fluctuations} \quad Thermal SZ \quad foregrounds + noise$$

ILC = weighted linear combination of frequency maps: •

$$\hat{s}(\vec{\theta}) = \sum_{\nu} w(\nu) d(\nu, \vec{\theta})$$

such that
$$\begin{cases} \langle \hat{s}^2 \rangle = w^t \langle dd^t \rangle w \text{ minimum (1)} \\ \sum_{\nu} w(\nu) a(\nu) = 1 \end{cases}$$
 (2)

• ILC weights :
$$w^t = \frac{a^t C^{-1}}{a^t C^{-1} a}$$
 ($C \equiv \langle dd^t \rangle$)

$$\Rightarrow \hat{s}(\vec{\theta}) = \underbrace{s_{CMB}(\vec{\theta})}_{recovered} + \underbrace{w^t b s_{TSZ} + w^t N}_{recovered}$$

$$\underbrace{signal}_{residuals}$$

$$\underbrace{signal}_{residuals}$$

$$\underbrace{signal}_{residuals}$$

$$\underbrace{signal}_{residuals}$$





Slide credit: M. Remazeilles

An Introduction to the CMB

Remazeilles et al 2011, 2013

Eriksen et al,

ILC summary: Example using CMB and SZ

$$d(\nu, \vec{\theta}) = a(\nu) \quad s_{CMB}(\vec{\theta}) + b(\nu) \quad s_{TSZ}(\vec{\theta}) + N(\nu, \vec{\theta})$$

$$\underbrace{Hanck frequency}_{maps} \quad \underbrace{CMB}_{fluctuations} \quad \underbrace{Thermal SZ}_{fluctuations} \quad foregrounds + noise$$

• ILC = weighted linear combination of frequency maps:

$$\hat{s}(\vec{\theta}) = \sum_{\nu} w(\nu) d(\nu, \vec{\theta})$$

such that
$$\begin{cases} \langle \hat{s}^2 \rangle = w^t \langle dd^t \rangle w & \text{minimum} (1) \\ \sum_{\nu} w(\nu) b(\nu) = 1 \end{cases}$$
 (2)

• ILC weights :
$$w^t = \frac{b^t C^{-1}}{b^t C^{-1} b}$$
 ($C \equiv \langle dd^t \rangle$)

$$\Rightarrow \hat{s}(\vec{\theta}) = \underbrace{s_{TSZ}(\vec{\theta})}_{W^{t}as_{CMB}} + w^{t}\Lambda$$

Benett et al, 2003 Tegmark et al, 2003 Eriksen et al, 2004 Delabrouille et al, 2009 Remazeilles et al 2011, 2013

· · ·	
recovered	minimized
signal	residuals
(2)	(1)





Slide credit: M. Remazeilles

Questions?



Feel free to email me or ask questions in our eCampus Forum