Role of polarization in 21cm surveys

Zheng Zhang

Université Paris Cité

École doctorale des Sciences de la Terre et de l'Environnement et Physique de l'Univers - ED560

Laboratoire AstroParticules et Cosmologie (APC) - Groupe Cosmologie

Role of polarization in 21cm surveys

Par Zheng ZHANG

Thèse de doctorat de Physique de l'univers

Dirigée par Martin BUCHER
et par Nicholas KAISER

Présentée et soutenue publiquement le 12 Octobre 2023

Devant un jury composé de:

Rapporteur	Philip BULL	Reader, University of Manchester
Rapporteur	Marta SPINELLI	Astronome adjoint, Observatoire de la Côte d'Azur
Examinateur	Alessandra TONAZZO	Professor, Université Paris Cité, Présidente du jury
Examinateur	Kavilan MOODLEY	Professor, University of KwaZulu-Natal
Examinateur	Ken GANGA	DR, CNRS
Directeur de thèse	Martin BUCHER	DR, CNRS

Abstract (français)

Titre: Le rôle de la polarisation dans les relevés à 21 cm

Résumé: La détection du signal à 21 cm est cependant difficile, car le signal extragalactique recherché est masqué par des avant-plans galactiques de trois à quatre ordres de grandeur plus brillants. L'avant-plan galactique étant spectralement lisse, il est théoriquement possible d'extraire le signal extragalactique recherché. Cependant, les systématiques instrumentales peuvent mélanger une partie de cet avant-plan spectralement lisse avec les composantes variant rapidement avec la fréquence pour qu'il soit confondu avec le signal extragalactique. La fuite de polarisation est l'un des défis les plus difficiles à relever car l'avant-plan fortement polarisé subit une rotation de Faraday lorsqu'il se propage dans le milieu interstellaire en présence du champ magnétique galactique. Bien que le signal cosmologique souhaité soit non polarisé, les observations radio étant basées sur la polarimétrie, une représentation inexacte des mesures de polarisation entraîne une fuite de l'avant-plan polarisé dans le signal. Les avant-plans polarisés avec une structure spectrale complexe compliquent l'élimination des avant-plans basée sur la régularité spectrale de l'avant-plan galactique.

Dans cette thèse, nous considérons plusieurs aspects du rôle de la polarisation dans la cartographie d'intensité à 21 cm. Nous passons en revue les mesures de polarisation du ciel radio, nous formalisons la mesure du signal de tension en termes de la figure du champ lointain de l'antenne, et nous présentons un modèle de mesures de puissance polarisée. En utilisant l'approximation linéaire, nous décrivons le processus de mesure et d'analyse des données. Nous définissons des opérateurs de projection de données qui réalisent un compromis optimal entre plusieurs objectifs de réduction ou d'extraction de données, y compris le rejet de la polarisation, la réduction du bruit et la sélection et le masquaas de degrés de liberté spécifiques du faisceau. Nous proposons également une stratégie optimale d'extraction des composante I de Stokes. En introduisant des spectres de puissance *a priori* sur le ciel polarisé, nous montrons que trouver les modes appropriés dominés par Stokes I est une question d'intégration de la polarisation. Enfin, nous développons une figure de mérite pour quantifier l'impact d'une caractérisation inexacte et imprécise du faisceau et de sa complexité sur l'extraction de la carte d'intensité. En introduisant une description statistique des erreurs systématiques du faisceau, nous modélisons l'impact de la fuite de polarisation et du mélange de modes dans la reconstruction des composante I de Stokes. En utilisant des modèles statistiques du ciel et du bruit, nous évaluons le rapport signal/bruit (SNR) de chaque mode de Fourier avant et après la prise en compte de l'incertitude du faisceau. Une figure de mérite (FoM) est ensuite définie pour établis un classement les antennes en fonction du facteur par lequel le SNR total pour la détection d'un signal cosmologique est réduit après la prise en compte de l'erreur d'incertitude du faisceau.

Mots clefs: radioastronomie, polarisation, cartographie d'intensité HI

Abstract

Title: Role of polarization in 21cm surveys

Abstract: The last decade has seen a surge of interest in using 21cm intensity mapping (IM) as a probe of the large-scale structure of the Universe. Detection of the 21cm signal is challenging, however, because the desired extragalactic signal is obscured by three to four orders of magnitude brighter Galactic foregrounds. Because the galactic foreground is spectrally smooth, it is theoretically possible to extract the desired extragalactic signal. However instrumental systematics can cause part of this spectrally smooth foreground to leak into components varying rapidly with frequency and thus be mistaken for an extragalactic signal. Polarization leakage is one of the most difficult challenges because the highly polarized foreground undergoes Faraday rotation as it propagates through the ISM in the presence of the galactic magnetic field. Although the desired cosmological signal itself is assumed unpolarized, since radio observations are based on polarimetry, an inaccurate representation of the polarization measurements lead to a leakage of the polarized foreground into the signal. Polarized foregrounds with a complex spectral structure would complicate foreground removal relying on the spectral smoothness of the galactic foreground.

In this thesis we consider several aspects of the role of polarization in 21cm intensity mapping. We review polarization measurements of the radio sky, formalize the measurement of the voltage signal in terms of the antenna far-field pattern, and present a model of polarized power measurements. Using the linear approximation, we describe the measurement and data analysis process. We define data projection operators that achieve the optimal compromise among multiple goals for data reduction or extraction, including polarization rejection, noise reduction and selection and avoidance of specific degrees of freedom of the beam. We also propose an optimal Stokes I extraction strategy. By introducing *a priori* power spectra over the polarized sky, we show that finding the appropriate Stokes I dominated modes is a matter of integrating out the polarization. Finally, we develop a figure-of-merit to quantify the impact of inaccurate and imprecise beam characterization and beam complexity on intensity map extraction. By introducing a statistical description of the beam systematics, we model the impact of polarization leakage and mode mixing in the Stokes I reconstruction. Using statistical models of the sky and the noise, we evaluate the signal-to-noise ratio (SNR) of each Fourier mode before and after accounting for beam uncertainty. A figure-of-merit (FoM) is then defined to rank reflector-feed systems according to the factor by which the total SNR for detecting a cosmological signal is reduced after accounting for beam uncertainty error.

Keywords: radio astronomy, polarization, 21cm intensity mapping

Dedication

This is dedicated to Nick Kaiser, my teacher and friend who I will always miss.

Acknowledgements

I would like to thank the entire HIRAX team for their warm hospitality. I would also like to thank the DIM ACAV$^+$ for the financial support of my PhD scholarship. Special thanks to my advisor Martin Bucher and co-advisor Nick Kaiser for making my three years in Paris so enjoyable. The knowledge that they have taught me and the support that they have given me have meant a great deal to me. Finally, I would like to give a special thank you to my wife, Yuyan, who has eased many of the anxieties that I have faced. It is because of her love and companionship that I have been able to continue to move forward at my own pace.

Preface

Part of the work in this thesis was collaborative, and here I state explicitly what is my contribution and what is the contribution of my collaborators. This thesis will result in the following publications:

1. Optimal Stokes-I Extraction for 21cm Intensity Mapping experiments (Zheng Zhang and Martin Bucher)

2. Toward a Beam Quality Figure-of-Merit for Radio Intensity Mapping (Martin Bucher, Carla Pieterse, Kavilan Moodley, Zheng Zhang, and Dirk de Villiers)

3. Polarization Rejection Beam Quality Figure-of-Merit for Radio Intensity Mapping (Zheng Zhang, Martin Bucher, Carla Pieterse, and Dirk de Villiers)

In Paper 1 (Zhang and Bucher, in prep.), the contribution of Martin Bucher was proposing the idea on using sky priors to project out polarization. My contribution was the formalism, analysis, and discussion. In Paper 2 (Bucher et al., in prep.), the contribution of Martin Bucher was the idea and the initial formalism of the isolated figure-of-merit (FoM) beam evaluation. Kavilan Moodley contributed the input power spectra both for the signal and the foregrounds as well as the apodization in frequency scheme. Carla Pieterse computed the Hirax beam data using CST and beam error simulations. Prof Dirk de Villiers provided guidance on the antenna modelling aspects of the project. My contribution was to improve of the formalism and the numerical realisation of the analysis. In Paper 3 (Zhang et al., in prep.), the contribution of Martin Bucher was the idea of generalizing FoM beam evaluation to fully polarized measurement. Carla Pieterse assisted by Dirk de Villiers provided the Hirax beam data and beam error simulations. My contribution was the polarized FoM formalism, input power spectra of polarization, the numerical realisations, and analysis and discussion.

I have also contributed to the simulation and data analysis pipeline for HIRAX, specifically a likelihood-based power spectrum estimator and an improved analytical covariance model for the diffuse polarized foreground.

After defending this thesis, I intend to work with my advisor on finalizing and submitting these papers. We will also work on a REACH paper applying the formalism to clarifying the impact of polarization on global 21 cm measurements.

Contents

List of Figures

List of Tables

Part I

Cosmology

Chapter 1

Cosmic expansion and acceleration

The fate of the universe is an ancient question, perhaps as old as human civilization. One can imagine that the earliest humans came out of their caves on sentimental nights and watched the twinkling stars decorating the sky in an unchanging pattern from year to year, filling their minds with infinite curiosity and imagination. "Will the universe continue like this? What was it like a long time ago? What will it be like in the distant future?" If they needed a motive when they first created words like "eternity," there could be no more inspiring scenario than contemplating the fate of the universe.

After a long process of civilization, and especially with the development of modern science, the fate of the universe gradually changed from a purely philosophical question into a scientific one that can be systematically discussed through theory and observation. We now know that, among the fundamental interactions, gravity dominates the dynamics of the universe on large scales because of its long-range nature and its universality to all objects. However, after the first theory of gravity was developed under Newton's apple tree, mankind did not immediately embrace its first cosmological model. It was not until 1917, two years after the publication of Einstein's General Relativity (GR), that the first cosmological model was proposed by Einstein. He was probably motivated by the generalized Copernican principle[1] that the universe should be homogeneous not only in space but also in time. To allow for static (non-expanding) solutions, Einstein modified his field equations by introducing a cosmological constant Λ with units of inverse length squared so that

$$\mathbf{G} - \Lambda \mathbf{g} = 8\pi\kappa \mathbf{T}, \tag{1.1}$$

where $\kappa \equiv G_N/c^4$ with G_N the Newtonian gravitational constant. Although Einstein himself did not interpret it this way, later cosmologists saw Λ more as a constant energy density of the vacuum, or a matter with strong tension or, say, repulsive gravity. To see this point, we could move $\Lambda\mathbf{g}$ to the right side of the field equation and see that the term would correspond to a matter with $T^{\mu\nu} = (\Lambda/8\pi\kappa)\mathrm{diag}(1, -1, -1, -1)$. If Λ

[1] Also known as the perfect cosmological principle.

is positive, this would be matter with positive energy density $\rho_\Lambda = T^{00}/c^2$ but strong negative pressure $P = -\rho_\Lambda c^2$.

However this static model was overturned soon after it was proposed. In the late 1920s, Edwin Hubble (Hubble, 1929) observed Cepheid variable stars far outside the Milky Way and found that the local universe is expanding. With this discovery, Einstein lost the original motivation[2] for introducing Λ and thus removed it from his field equations. After removing this repulsive-gravity stuff, however, the question of the fate of the universe became more concrete and even a little scary: will the expansion of the universe slow down because gravity attracts all the stuff? Will it finally collapse into a Big Crunch?

These daunting questions drove the rapid development of cosmic observations. In the 1930s, supernova astronomers Walter Baade and Fritz Zwicky realized that the superbrightness of supernovae (Baade, 1938) could be used to trace the expansion of the universe in the past, and thus had the potential to reveal the future fate of the universe. In principle, both the brightness and the color of a supernova can be used as indicators of how far away it is. The farther away a supernova is from us, the earlier the signal was emitted, the fainter the observed brightness, and the redder the color. If we can observe enough supernovae, we can extend the "Hubble diagram" (magnitude vs $\log(z)$) to obtain the "deceleration" of the expansion of the universe. Although this seemed like a good idea, supernovae were not good enough as standard candles at the time, since the brightness of supernovae varies by a factor of two or three. Therefore, the idea of using supernovae to measure the deceleration did not become a reality immediately after it was proposed.

It wasn't until the mid-1980s that the "Type Ia" subclass of the supernovae was identified and was then considered as a good standard candle[3]. In 1987, Saul Perlmutter and Carl Pennypacker came up with the idea of using the Type Ia supernovae to revive Baade and Zwicky's idea of measuring the deceleration of the expansion of the Universe. This was not an easy task at the time and was fraught with uncertainty: They were not sure if they could find enough supernovae. Even if they could, would they be bright enough to confirm the subtype [4]? How standard were these supernovae? How to correct for the brightness of supernovae dimmed by dust? Fortunately, these problems have been solved one by one. This history and more details of the story are detailed in the Perlmutter et al. (2011). Observations then succeeded in accumulating a number of supernova measurements, and the history of cosmic expansion slowly emerged. As a milestone, Perlmutter et al. (1998) provided the first evidence for a

[2]However, loss of motivation for it is not the same thing as rejection of the possibility.

[3]To be precise, the peak brightness of Type Ia supernovae was what astrophysicists knew to be standard.

[4]A distant supernova was seen in the faint tail of the spectrum. They were concerned about the uncertainties in the K correction used to compare the observed faint tail with a supernova seen at the peak of the spectrum.

Universe with a cosmological constant, using a very well measured supernova. (Note that the "evidence" is based on the assumptions that (1) GR is correct on cosmological scales, and (2) except for the possible cosmological constant, there are no unknown forms of energy.) Soon after, two independent teams (Knop et al., 1999; Riess et al., 1998) provided even stronger evidence for an accelerating universe and a cosmological constant, respectively, using more supernovae. Today an accelerating universe has been convincingly established by more precise cosmological measurements. In retrospect, the discovery of cosmic acceleration certainly came as a huge shock to supernova astronomers, who had expected to measure the deceleration of the expanding universe.

Chapter 2

More evidence of a low density universe

In the previous section we have reviewed observations of cosmic acceleration as direct evidence for the existence of some energy form or gravity theory with uncertain details. However, in the history of cosmology, supernova observations were only one of many factors that led to the cosmological constant returning to the centre of cosmologists' attention.

After the discovery of cosmic expansion, Einstein and de Sitter appealed to an expanding, homogeneous and isotropic, spatially flat, matter-dominated universe, the so-called Einstein-de Sitter model, which was the standard cosmological model until the 1990s. By the mid-1990s, the Einstein-de Sitter model was showing numerous cracks. There was an abundance of evidence in favour of a low-density universe. Below we briefly present the "combined onslaught" on the model. A more detailed discussion can be found in the conference proceedings (Turok, 1997).

2.1 Cluster baryon fraction

The discovery of the Hubble expansion at the time encouraged the hypothesis that the universe could have evolved from a hot, dense fireball. Since temperatures in the earliest moments may have been so high that nuclei could not exist, the hypothesis predicts a period of nucleosynthesis during the expansion and cooling of the universe, known as the Big Bang nucleosynthesis (BBN). The discussion of Big Bang nucleosynthesis was initiated by Alpher, Herman, and Gamow (1948), and they believed that all elements could be produced in the hot, dense state of the early universe. The shortcomings of this idea were soon recognized: on the one hand, the lack of stable nuclei of mass number $A = 5$ and $A = 8$, and the Coulomb repulsion between highly charged nuclei, prevent significant nucleosynthesis beyond Li^7. On the other hand, many of the details of the observed abundances of elements such as Tc were explained in terms of stellar processes. However, the 'stellar synthesis' seems deficient to explain the high He^4 abundances observed in various primitive objects. Therefore, the possibility of a

'universal synthesis' was revived to explain the origin of the abundant He^4 (Hoyle and Tayler, 1964).

Soon after the identification of the problem of the origin of helium and the discovery of the cosmic background radiation by Penzias and Wilson in 1965, Peebles (1966) and Wagoner et al. (1967) carried out detailed calculations of BBN. Their pioneering work established a standard BBN model of the early universe in which only significant amounts of D, He^3, He^4, and Li^7 can be produced in the universal synthesis. The standard BBN model has extraordinarily simple physics. It is characterized by a single parameter, the ratio η (or $\eta_{10} \equiv \eta/10^{-10}$) of the number of baryons to the number of photons. Although this parameter has little effect on the evolution of the early universe, it does determine the relative abundance of light nuclei, which is an important observational relic of the early universe. Since the number of photons is derived from the CMB temperature $n_\gamma = 411$ cm^{-3}, η gives the present-day baryon density:

$$\Omega_b h^2 = \frac{\rho_b}{\rho_c/h^2} = \frac{\eta n_\gamma m_N}{\rho_c/h^2} = 3.65 \times 10^{-3} \eta_{10}, \tag{2.1}$$

where m_N is the neutrino mass, $H_0 = 100h$ (km/s)/Mpc and ρ_c is the critical density defined by equation (3.10).

Using numerical integration of reaction networks (Wagoner et al., 1967), the light nuclei abundances are predicted. A useful fitting formula for the predicted mass fraction of He_4 is (Hogan, 1997; Sarkar, 1996)

$$Y_P(\mathrm{He}^4) = 0.235 + 0.012 \ln\left(\frac{\eta_{10}}{2}\right)\left(\frac{\eta_{10}}{2}\right)^{-0.2} + 0.011\left[1 - \left(\frac{\eta_{10}}{2}\right)^{-0.2}\right] \pm 0.0006, \tag{2.2}$$

where the subscript P denotes 'primordial'. The abundances of other light elements by number are given by (Sarkar, 1996)

$$\begin{aligned}
\left(\frac{D}{H}\right)_P &= 3.6 \times 10^{-5\pm0.06}\left(\frac{\eta_{10}}{5}\right)^{-1.6} \\
\left(\frac{He^3}{H}\right)_P &= 1.2 \times 10^{-5\pm0.06}\left(\frac{\eta_{10}}{5}\right)^{-0.63} \\
\left(\frac{Li^7}{H}\right)_P &= 1.2 \times 10^{-11\pm0.2}\left[\left(\frac{\eta_{10}}{5}\right)^{-2.38} + 21.7\left(\frac{\eta_{10}}{5}\right)^{2.38}\right].
\end{aligned} \tag{2.3}$$

These predictions as functions of η are shown graphically in Figure 2.1.

Despite the beauty of the theory, the measurement of actual primordial abundances has been a considerable challenge. Because D, He^3, and Li^7 are produced in such small quantities, the establishment of their cosmic origin is more difficult and takes longer time than that of He^4. With a lot of work in the 1970s and 80s, it was found that observations of He^4 in extragalactic HII regions, D in quasar absorbers, and Li^7 in old, metal-poor halo stars (Spite and Spite, 1982) can be used to constrain the abundances

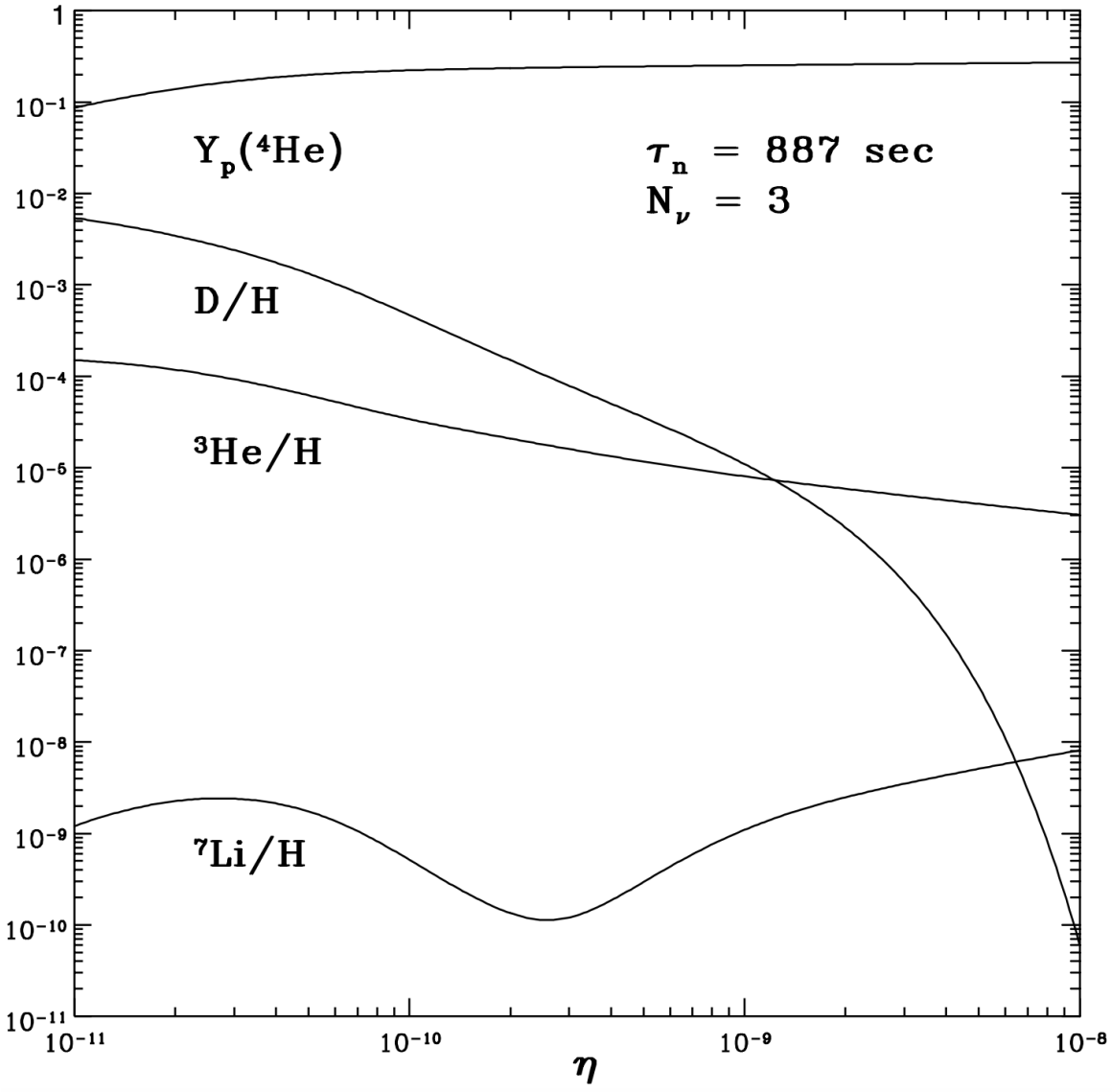

Figure 2.1: Dependence of primordially synthesized light element abundances on the baryon-photon fraction. τ_n is the lifetime of neutron. This figure is taken from Sarkar (1996).

of primordial light nuclei. In the 1990s, it was already possible to establish a concordance between theory and observation. The concordance showed the consistency of the predictions of the Standard Model and led to limits on the nucleon-to-photon ratio, e.g., $2.5 < \eta_{10} < 6$ calculated by Copi et al. (1995). Here we mention a recent analysis from Riemer-Sørensen and Jenssen (2017); Cooke et al. (2018), which gives

$$5.8 < \eta_{10} < 6.5. \tag{2.4}$$

Then the present-day baryon density defined by equation (2.1) is $\Omega_b h^2 \simeq 0.022$.

Since rich clusters of galaxies provide a 'fair sample' for studying the universal baryon fraction, the cluster baryon fraction can be used to estimate the fraction of the universal baryon density to the total matter density. The mass of the baryons in clusters is made up of at least two parts: the hot intracluster gas and the luminous parts of the galaxies. Thus, the cluster baryon fraction is characterized by

$$\frac{\Omega_b}{\Omega_m} \gtrsim \frac{(\text{mass of gas + stars})}{(\text{cluster mass})} \simeq 0.07h^{-1.5} + 0.05. \tag{2.5}$$

The matter density of the universe is

$$\Omega_m \lesssim \frac{\Omega_b}{0.07h^{-1.5} + 0.05} \simeq 0.3. \tag{2.6}$$

Therefore, the baryon density given by the standard BBN model and the high baryon content observed in clusters appeal to a low density universe with $\Omega_m \simeq 0.3$.

2.2 Age of universe

In the 1980s and early 90s, another controversy about the age of the universe was also widely discussed. Independent determination techniques seemed to have conflicting estimates. For some favored values of the Hubble constant at that time (e.g. $H_0 = 73 \pm 7 \pm 8$ in Freedman and Feng (1999)), the cosmic age given by the dynamics of the Einstein-de Sitter universe is found to be smaller than the lower bound of the age of the oldest stars. Below is a brief summary of these cosmic age techniques.

- Dynamical argument

  In the Einstein-de Sitter model, the scale factor $a(t)$ can be expressed explicitly as

  $$a(t) = \left(\frac{t}{t_0}\right)^{\frac{2}{3}}, \tag{2.7}$$

  where t_0 is the present age of the universe. This implies that t_0 can be

determined by the Hubble constant

$$t_0 = \frac{2}{3} H_0^{-1}. \tag{2.8}$$

If we assume the modern observed Hubble constant value $H_0 \approx 70\,(\mathrm{km/s})/\mathrm{Mpc}$, the universe has an age of $t_0 \approx 9$ billion years in this model.

- Nuclear arguments
  - Oldest stars (globular clusters)

    The basic age of a globular cluster is determined by the time required for low-mass stars to burn up their core hydrogen and leave the main sequence.

    It is shown in Schramm (1990) that the age of globular clusters gives an age estimate of about $14 \pm 2 \pm 2$ billion years. A firm lower bound is $\gtrsim 10$ billion years.
  - Radioactive dating (nucleocosmochronology)

    Nucleocosmochronology is a technique that uses the abundances and production ratios of radioactive nuclides, in conjunction with information about the chemical evolution of the Galaxy, to derive information about the time scales over which the elements of the Solar System were formed (Schramm, 1990). Meyer and Schramm (1986) gave a typical estimate for the age of the galaxies, which is of the order of 9.6 billion years.

As we later learned, the introduction of the Λ was the end of this controversy.

2.3 Large scale velocity flows

Because galaxies generally have peculiar velocities in the Hubble flow, the observed redshift in the rest frame of the Local Group is

$$cz = H_0 r + \hat{r} \cdot (\mathbf{v}(\mathbf{r}) - \mathbf{v}(0)) \tag{2.9}$$

where $\mathbf{v}(0)$ represents the peculiar velocity of the Local Group, and $\hat{r}$ is the unit vector toward the galaxy of interest. On large scales, since the gravitational instability equations can be linearized, we have direct proportionality between late-time velocity and density field divergence Peebles (1993):

$$\nabla \cdot \mathbf{v}(\mathbf{r}) = -\Omega^{0.6} \delta(\mathbf{r}), \tag{2.10}$$

which can be rewritten in Fourier coordinates as

$$i\mathbf{k}\cdot\mathbf{v}(\mathbf{k}) = -\Omega^{0.6}\delta(\mathbf{k}). \tag{2.11}$$

Defining the velocity power spectrum $P_v(k) \sim \langle \mathbf{v}(\mathbf{k})^2 \rangle$, we have

$$P_v(k) = \Omega^{1.2} k^{-2} P(k). \tag{2.12}$$

Then the expected amplitude of the bulk flow on a scale R is given by (Strauss, 1997)

$$\langle \mathbf{v}(R)^2 \rangle = \frac{\Omega^{1.2}}{2\pi^2} \int \mathrm{d}k P(k) W^2(kR). \tag{2.13}$$

where W is the smoothing window in Fourier coordinates. If the phases of the Fourier modes of the density field are random, then each component of the velocity field is Gaussian distributed, which means that $v(R)$ is Maxwellian distributed.

Bahcall et al. (1994); Bahcall and Oh (1996) investigated the distributions of the velocities of rich clusters of galaxies for different cosmological models and compared these model predictions with observations. As can be seen in Figure 2.2, the observed cluster velocity function is most consistent with a low–mass-density, $\Omega \simeq 0.3$, CDM model. The absence of a high-velocity tail shows inconsistency with the $\Omega = 1$ CDM.

2.4 Evolution of cluster abundance

White et al. (1993) showed how the rms linear fluctuation in the mass distribution on scales of $8\mathrm{h}^{-1}$ Mpc, denoted by σ_8, can be constrained by mass and abundance of clusters. By modelling clusters as "objects" above a mass threshold in a density fluctuation field that was initially Gaussian, they used the cluster mass function to constrain $\sigma_8\Omega_m^{0.6}$:

$$\sigma_8\Omega_m^{0.6} \simeq 0.5 - 0.6. \tag{2.14}$$

This result generally contradicts the high amplitude of σ_8 inferred for a $\Omega = 1$ cold dark matter model. For example, one would expect to measure $\sigma_8 \sim 1.1$ for $h = 5$ if the COBE anisotropies arise from primordial scale-invariant perturbations in an $\Omega = 1$, cold dark matter universe (Wright et al., 1992; Efstathiou et al., 1992). Therefore, if COBE fluctuations was not overestimated, the standard CDM model at the time was rejected at the 2.5σ confidence level (White et al., 1993).

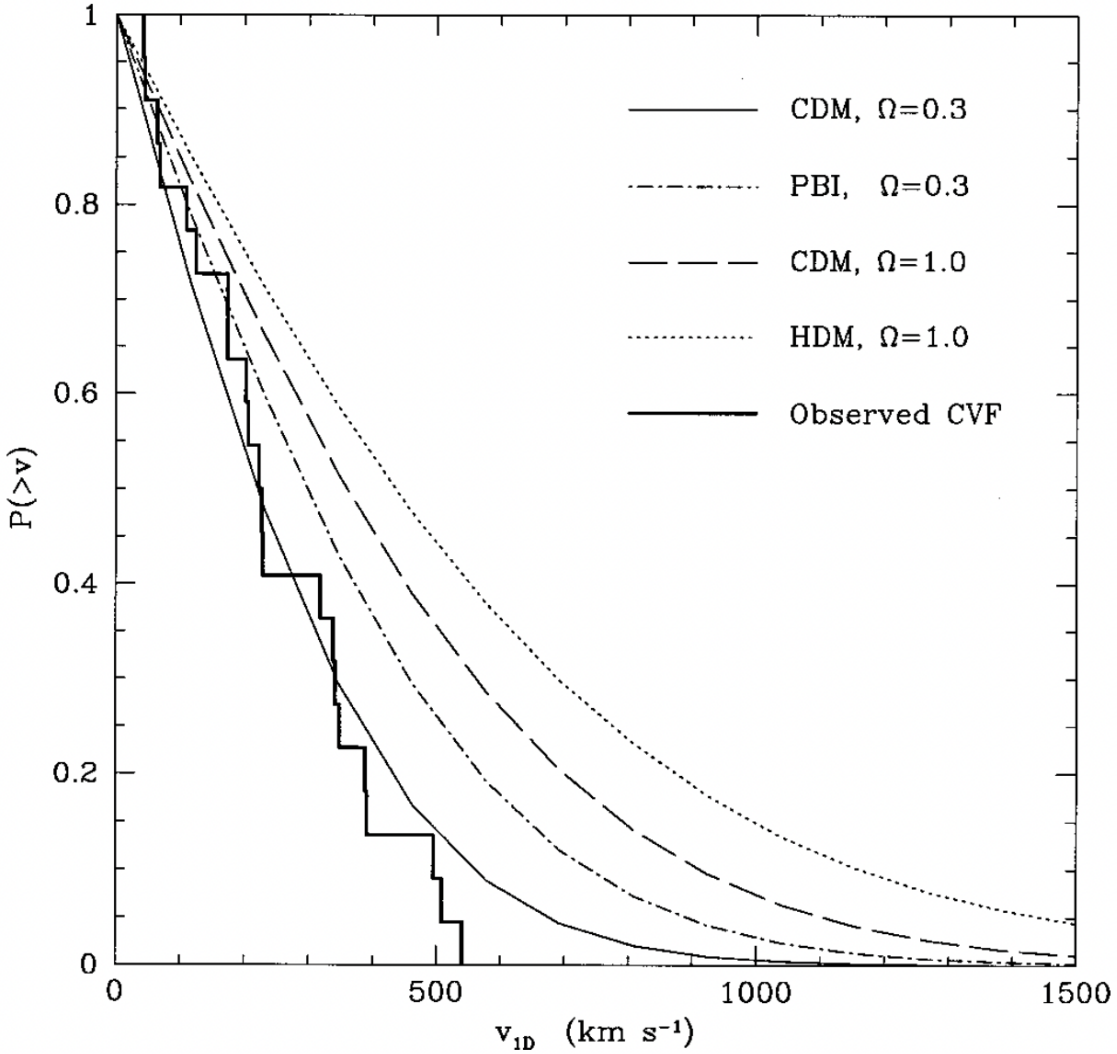

Figure 2.2: Observed vs model predicted cluster velocity functions (CVF). $P(> v)$ is the ratio of the number of clusters with velocities greater than v to the total number of samples. This figure is taken from Bahcall and Oh (1996).

Chapter 3

Dark energy or modified gravity?

Observations, in particular of the Cosmic Microwave Background and of large-scale galaxy clustering, have shown that our Universe is very isotropic and homogeneous on large spatial scales, which is also known as the cosmological principle. If one assumes that GR is the correct description of gravity on cosmological scales, then the SNIa surveys derive the cosmic acceleration as a natural consequence of any spatially isotropic and homogeneous model of the Universe.

Although the supernova teams that first discovered the cosmic acceleration used the cosmological constant to frame their analysis, which is probably the simplest mathematical form for their goal, it is clear that Λ is not the only candidate to explain the acceleration. If we drop the assumption that the vacuum energy density is constant, an alternative explanation could be that the cosmic acceleration is driven by an unknown form of energy, namely "dark energy," with negative pressure. Instead of, or probably in addition to, the introduction of new energy forms, there are also theories that explain the acceleration of the universe by modifying GR (see review articles, e.g., (Joyce et al., 2015)). One possibility is to replace the Ricci scalar R by a function $R + f(R)$ in the gravitational action (Carroll et al., 2004). Other options can be more radical. For example, one could introduce extra dimensions and allow gravitons to "leak" from the brane that represents the observable universe (Dvali et al., 2000).

The two handles for experimental tests of dark energy and modified gravity theories are to measure the history of cosmic expansion and the history of matter clustering with high precision over a wide range of redshifts. The equation (3.12) will show that within GR, the present values for $\Omega_{m,0}$, $\Omega_{r,0}$ and H_0 together with the measurement of $H(z)$ determine the equation of state of dark energy $\omega(z)$. In addition to Hubble measurements, distance measurements [equations (3.20) and (3.16)] using standard rulers or candles also constrain $\omega(z)$, though they require an integral over $H(z)$ which is itself an integral over $\omega(z)$. In the linear perturbation theory based on GR, $H(z)$ and $G(z)$ are connected in a consistent way. The growth rate of gravitational clustering may differ from the GR prediction in modified gravity models. Thus, comparing the

history of cosmic structure growth with the history of cosmic expansion can provide strong constraints on modified gravity theories.

3.1 Homogeneous expanding universe models

The cosmological principle allows us to characterize the overall geometry and evolution of the Universe with two parameters: the universal scale factor $a(\tau)$ accounting for the expansion (or contraction) and the spatial curvature constant. The general space-time metric, known as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, is written in terms of these two parameters:

$$\mathrm{d}s^2 = -c^2\,\mathrm{d}\tau^2 + a^2(\tau)\left[\mathrm{d}r^2 + f^2(r)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right], \tag{3.1}$$

where

$$f(r) = \begin{cases} R\sin\left(r/R\right), & \text{for positive curvature,} \\ r, & \text{for zero curvature,} \\ R\sinh\left(r/R\right), & \text{for negative curvature.} \end{cases} \tag{3.2}$$

If we define the conformal distance, $\chi \equiv r/R$, and absorb R into the scale factor, we obtain the more convenient form for the FLRW metric

$$\mathrm{d}s^2 = -c^2\,\mathrm{d}\tau^2 + a^2(\tau)\left[\mathrm{d}\chi^2 + S_k(\chi)^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right] \tag{3.3}$$

where $k = 0, \pm 1$ is the curvature constant and

$$S_k(\chi) = \begin{cases} \sin\chi, & \text{for } k = 1, \\ \chi, & \text{for } k = 0, \\ \sinh\chi, & \text{for } k = -1. \end{cases} \tag{3.4}$$

Assuming the matter content of the Universe can be mimicked by an ideal fluid, Alexander Friedmann (Friedmann, 1922) found the dynamical equations describing the FLRW Universe. For a fundamental observer comoving with the cosmological fluid, Einstein field equations lead to the Friedmann equations, which are the energy equation

$$\dot{a}^2 = \frac{8\pi}{3}G_N\rho a^2 - c^2k, \tag{3.5}$$

the acceleration equation

$$\ddot{a} = -\frac{4\pi}{3}G_N(\rho + 3P/c^2)a, \tag{3.6}$$

and the continuity equation

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P/c^2). \tag{3.7}$$

The energy equation is commonly called the *Friedmann equation*. It has a simple Newtonian analog where when comparing to the expansion of the dust shell in Newtonian gravity, we see that the term $(8\pi/3)G_N\rho a^2$ can be interpreted as potential energy competing with the kinetic energy term $\dot{a}^2$.

These three equations about $a(\tau)$, $\rho(\tau)$ and $P(\tau)$ are not independent; each can be obtained from the other two. To obtain solutions we still need an '*equation of state*' (EoS) relating the pressure and energy density $P = P(\rho)$. $P_m = 0$ for pressureless matter,[1] while for radiation $P_r = \rho c^2/3$. As for the dark energy, our ignorance about its EoS is encapsulated in an unknown time dependent function

$$\omega(\tau) \equiv P_{DE}(\tau)/\rho(\tau)c^2. \tag{3.8}$$

3.1.1 Cosmological parameters

As a result of the Hubble expansion, a local observer detecting light from a distant emitter sees a *cosmological redshift*, defined as

$$1 + z = \frac{\nu_1}{\nu_2} = \frac{a(\tau_2)}{a(\tau_1)} \tag{3.9}$$

where ν_1 and ν_2 are the emitted and observed frequencies, respectively, and τ_1 and τ_2 are the corresponding cosmic times. In the following we use z instead of τ because it is more convenient for discussing observations. It is also common to use the Hubble parameter $H \equiv \dot{a}/a$ and the critical density ρ_c defined by

$$\rho_c \equiv \frac{3H^2}{8\pi G_N}, \tag{3.10}$$

which effectively represents the the kinetic term about $\dot{a}^2$ in the form of potential energy.

Substituting the equations of state into equation (3.7), one obtains evolutions of different energy components. For pressureless matter, $\rho_m(z)/\rho_m(z=0) = (1+z)^3$, while for radiation, $\rho_r(z)/\rho_r(z=0) = (1+z)^4$. The evolution for dark energy is given by

$$\frac{\rho_{DE}(z)}{\rho_{DE}(z=0)} = \exp\left[\int_0^z \mathrm{d}z\, \frac{3(1+\omega)}{1+z}\right]. \tag{3.11}$$

If ω is constant, this reduces to $\rho_{DE}(z)/\rho_{DE}(z=0) = (1+z)^{3(1+\omega)}$.

Putting the above results and conventions together, and using the subscript '0' to denote quantities evaluated in the present universe, we can rewrite the Friedmann

[1] Here the pressure is the quantity that would be measured by the local co-moving observer. Note that the dynamical pressure of a gas is included in ρ instead of the overall pressure P.

equation as

$$H^2(z) = H_0^2\left(\frac{\rho}{\rho_{c,0}}\right) - c^2\frac{k}{a^2}$$
$$= H_0^2\left[\Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{DE,0}\,\exp\left(\int_0^z \mathrm{d}z\,\frac{3(1+\omega)}{1+z}\right)\right] - c^2\frac{k}{a^2}, \tag{3.12}$$

where $\Omega_{i,0} \equiv \rho_{i,0}/\rho_{c,0}$ for $i = m$, r, DE. It is also common to define $\Omega_k \equiv -c^2k/H_0^2$, although the curvature term should not be understood as an energy form.

Much effort in modern cosmology has been devoted to constraining $\omega(z)$. Equation (3.12) shows that if the present values of the fractional energy densities and the measurement of $H(z)$ pin down $\omega(z)$. Apart from measuring $H(z)$, others measure the distance-redshift relationship. In the next section, we briefly review distance measurements in cosmology.

3.1.2 Distance measurements

The comoving distance between two objects in the Universe is the distance that would be measured along a path defined between the two spatial points in the current Universe. Assuming $a_0 = 1$, so that $a(z) = (1+z)^{-1}$, then the line-of-sight comoving distance is given by

$$D_C = \int_{\tau(z)}^{\tau_0} \frac{c\,\mathrm{d}\tau'}{a} = \int_z^0 \frac{c\,\mathrm{d}z'}{a\dot{z}'} = \int_0^z \frac{c\,\mathrm{d}z'}{H}. \tag{3.13}$$

The transverse comoving distance D_M is defined so that the comoving distance between two events at the same cosmological redshift but separated on the sky by $\delta\theta$ is $D_M\delta\theta$. The relation between the transverse and line-of-sight comoving distances is given by (Hogg, 2000)

$$D_M = \begin{cases} D_H\frac{1}{\sqrt{\Omega_k}}\sinh\left[\sqrt{\Omega_k}D_C/D_H\right], & \text{for } \Omega_k > 0, \\ D_C, & \text{for } \Omega_k = 0, \\ D_H\frac{1}{\sqrt{|\Omega_k|}}\sin\left[\sqrt{|\Omega_k|}D_C/D_H\right], & \text{for } \Omega_k < 0, \end{cases} \tag{3.14}$$

where $D_H \equiv c/H_0$ is the Hubble distance.

The foregoing comoving distances are calculated entirely from the cosmological model. In actual observations, however, the key quantities are two apparent distances, D_A and D_L. The *angular diameter distance*, D_A, is calculated using the angular size obtained by observing the standard ruler. The observed angular size $\delta\theta$ is related to

δl, the proper transverse size of the object when the observed signal was emitted, by

$$(\text{Comoving transverse size}) = D_M \delta\theta \sim (1+z)\delta l. \tag{3.15}$$

D_A is then defined as the ratio of δl to $\delta\theta$

$$D_A \equiv \frac{\delta l}{\delta\theta} = \frac{D_M}{1+z}. \tag{3.16}$$

Similarly, the *luminosity distance* D_L is obtained from the apparent flux density S_ν of the standard candle with luminosity L_ν. Physically, S_ν and L_ν are related to D_M by the conservation of photons, so that

$$(\text{Number of photons}) = \frac{4\pi D_M^2 S_\nu(\nu_o)\,\mathrm{d}\nu_o\,\mathrm{d}\tau_o}{h\nu_o} \sim \frac{L_\nu(\nu_e)\,\mathrm{d}\nu_e\,\mathrm{d}\tau_e}{h\nu_e} \tag{3.17}$$

where ν_o is the observed frequency and $\nu_e = (1+z)\nu_o$ is the emission frequency. The time scales for emitting and receiving the photons are estimated as $\mathrm{d}\tau_e = \nu_e^{-1}$ and $\mathrm{d}\tau_o = \nu_o^{-1}$. The conservation of photons gives

$$4\pi D_M^2 S_\nu(\nu_o) = \frac{L_\nu(\nu_e)}{1+z}. \tag{3.18}$$

The luminosity distance is then defined as the apparent distance such that

$$\begin{aligned} 4\pi D_L^2 S_\nu(\nu_o)\,\mathrm{d}\nu_o &\equiv L_\nu(\nu_e)\,\mathrm{d}\nu_e \\ &= (1+z)L_\nu(\nu_e)\,\mathrm{d}\nu_o. \end{aligned} \tag{3.19}$$

As a result, D_L in terms of D_M is

$$D_L = (1+z)D_M. \tag{3.20}$$

3.2 The inhomogeneous universe

In the conventional Big Bang model, the very early universe underwent a period of accelerated expansion known as inflation. Inflationary cosmology was developed in a series of papers in the 1980s, including Brout et al. (1978), Starobinsky (1980), Kazanas (1980), Sato (1981), Guth and Pi (1982), Linde (1982) and Albrecht et al. (1982). A brief story in which quantum fluctuations generate cosmological fluctuations can be understood simply as follows: The inflationary field dominates the energy density of the universe during inflation, which ends when the potential steepens. According to the uncertainty principle, this field, as the clock of inflation, will have spatially varying fluctuations. As a result, inflation will end at different times in different places. Because of this local difference in inflation history, the amount of expansion varies from region

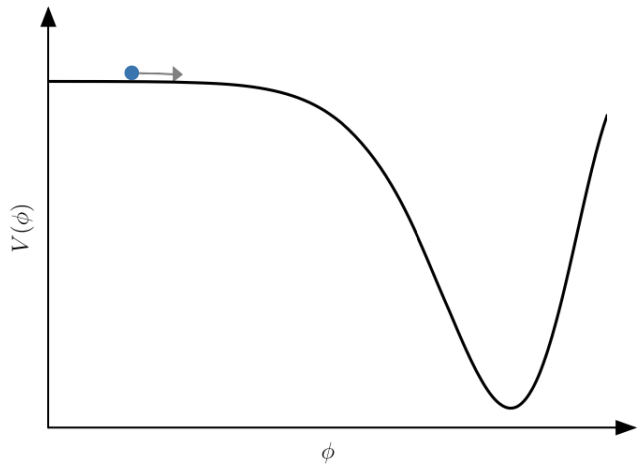

Figure 3.1: A scalar field slowly rolling down a potential. This is taken from Dodelson and Schmidt (2020).

to region, which inevitably causes the spatial variations in energy density. Thus it is often said that the seeds of structure were planted during inflation. In other words, the inflation provided the initial conditions for the formation and evolution of structure.

Most models of inflation are slow-roll models and the simplest candidate to drive the accelerated expansion is the potential energy of a canonical scalar field $\phi(\mathbf{x}, t)$, which can have negative $\rho + 3P$. The energy momentum tensor for ϕ is given by

$$T^{\alpha}{}_{\beta} = g^{\alpha\mu}\frac{\partial\phi}{\partial x^{\nu}}\frac{\partial\phi}{\partial x^{\beta}} - \delta^{\alpha}{}_{\beta}\left[\frac{1}{2}g^{\mu\nu}\frac{\partial\phi}{\partial x^{\mu}}\frac{\partial\phi}{\partial x^{\nu}} + V(\phi)\right]. \tag{3.21}$$

For a homogeneous field, the energy density ρ and the pressure P are

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi), \qquad P = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{3.22}$$

A slowly rolling scalar field with little kinetic energy hence has negative pressure (see Figure 3.1). The inflationary period comes to an end when the field has reached the minimum of its potential. At this point, however, the field is no longer slowly rolling, but oscillating around the minimum with significant kinetic energy. The kinetic energy must then be transferred to the particles of the Standard Model, which is still a speculative process called *reheating*. After this period with highly uncertain details, the universe was in a hot and dense state. The interactions between the Standard Model particles were so strong that thermal equilibrium was reached between different species. As the universe expands and cools, the interaction rate of the particles decreases. When the interaction rate of a species falls below the expansion rate, the local thermal equilibrium of that species is no longer well maintained, and it therefore decouples from the thermal equilibrium fluid. On the other hand, as the temperature decreases, the net direction of some of the particle reactions is reversed, and even

some lighter elements are allowed to appear: Shortly after the decoupling of neutrinos, electron-positron annihilation occurred, and neutrons thermodynamically decoupled from protons; three minutes after the end of the Big Bang, the so-called Big Bang Nucleosynthesis (BBN) created the first light elements; and about 380,000 years after the Big Bang, the temperature of the universe dropped to about 1eV, low enough for free electrons and protons to form stable atomic hydrogen. This event is also known as "recombination"; prior to recombination, the plasma of protons and electrons was effectively coupled to photons by Thomson scattering. Since a large number of electrons were trapped after the recombination, the photons decoupled from the baryons and began to be a free stream, becoming what is now known as the cosmic microwave background radiation.

The primordial inhomogeneities induced by inflation and the formation of new forms of energy during the cooling of the Universe make today's Universe full of fascinating structures. In this chapter we present a schematic discussion of structure growth using linear perturbation theory. The goal is to introduce the Baryon Acoustic Oscillations, the frozen relics left over from the pre-decoupling universe, as a standard statistical ruler used in HI intensity mapping.

3.2.1 Gravity and matter perturbations

The first order perturbation of the Einstein equation is

$$\delta G^{\mu}{}_{\nu} = 8\pi G_N \delta T^{\mu}{}_{\nu}, \tag{3.23}$$

where $\delta G^{\mu}{}_{\nu}$ is the perturbed Einstein tensor. Using the scalar-vector-tensor (SVT) decomposition, a general metric perturbation is represented by $4+4+2$ SVT degrees of freedom, that is, four scalar fields plus two divergence-less vector fields plus one symmetric, divergence-less, trace-free tensor field.[2] Since a general coordinate transformation is parameterized by $2+2+0$ SVT functions of space and time, the general metric perturbation has only $2+2+2$ physical SVT degrees of freedom. Here we consider only the scalar perturbation in the Newtonian gauge

$$\mathrm{d}s^2 = a^2(\eta)\left[-(1+2\Psi)\,\mathrm{d}\eta^2 + (1-2\Phi)\delta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\right], \tag{3.24}$$

where $\Psi(\mathbf{x},\eta)$ and $\Phi(\mathbf{x},\eta)$ are the two scalar degrees of freedom, and we have defined a dimensionless time coordinate, the conformal time η, so that $c\,\mathrm{d}\tau = a(\tau)\,\mathrm{d}\eta$.

The stress-energy perturbations can be formulated in two ways. One is to perturb the phase-space distribution function of each species to first order, which in turn produces the fully perturbed stress-energy tensor; the other is to represent and perturb the stress-energy tensor using macroscopic physical parameters, including energy density,

[2] These are 3D spatial fields.

pressure, bulk velocity, and shear stress. The latter additionally requires conservation equations and equations of state to constrain these physical parameters. The former approach, on the other hand, already implies these constraints since it is given from first principles. In this schematic discussion, we follow the second path and expand the matter stress energy tensor around the homogeneous background density $\bar{\rho}$ and pressure $\bar{P}$:

$$T^0{}_0 = -\bar{\rho}(\eta) - \delta\rho \tag{3.25}$$

$$T^i{}_0 = -[\bar{\rho}(\eta) + \bar{P}(\eta)](v^i/c) \tag{3.26}$$

$$T^i{}_j = [\bar{P}(\eta) + \delta P]\delta^i{}_j + \Pi^i{}_j \tag{3.27}$$

where v^i is the bulk velocity of the fluid and $\Pi^i{}_j$ is a symmetric, transverse and traceless tensor describing the shear stress. For fluids with frequent collisions, as is the case for the limited scope of the next section, the pressure is kept isotropic and $\Pi = 0$. It is common to define the dimensionless density contrast $\delta \equiv \delta\rho/\rho$ and the momentum density $q^i = (\bar{\rho} + \bar{P})v^i/c$ for convenience. For scalar fluctuations, we can write the Fourier modes of v^i and $\Pi^i{}_j$ as (Mukhanov et al., 1992; Kodama and Sasaki, 1984; Baumann, 2022)

$$v^i = i\hat{\mathbf{k}}^i v \tag{3.28}$$

$$\Pi^{ij} = (\bar{\rho} + \bar{P})\hat{\mathbf{k}}^{\langle i}\hat{\mathbf{k}}^{j\rangle}\sigma \tag{3.29}$$

where $\hat{\mathbf{k}}^{\langle i}\hat{\mathbf{k}}^{j\rangle} \equiv \hat{\mathbf{k}}^i\hat{\mathbf{k}}^j - (1/3)\delta^{ij}$ stands for the longitudinal, traceless component. The total stress energy tensor is the sum of those of different species, $T_{\mu\nu} = \sum_s T^{(s)}_{\mu\nu}$, which implies

$$\delta\rho = \sum_s \delta\rho_{(s)}, \qquad \delta P = \sum_s \delta P_{(s)}, \qquad \Pi^{ij} = \sum_s \Pi^{ij}_{(s)}, \qquad q^i = \sum_s q^i_{(s)}. \tag{3.30}$$

The equations of motion of the perturbation parameters $\{\delta\rho, \delta P, v^i, \Pi^{ij}\}$ follow the conservation equations of the total stress-energy tensor, $\nabla_\mu T^\mu{}_\nu = 0$. For species that do not transfer momentum and energy to other fluids, $\{\delta\rho_{(s)}, \delta P_{(s)}, v^i, \Pi^{ij}_{(s)}\}$ also satisfy the conservation equations. For $\nu = 0$ we find the continuity equation

$$\delta'_{(s)} = -\frac{\bar{\rho}_{(s)} + \bar{P}_{(s)}}{\bar{\rho}_{(s)}}\left(\frac{\partial_i v^i_{(s)}}{c} - 3\Phi'\right) - 3\mathcal{H}\left(\frac{\delta P_{(s)}}{\bar{\rho}_{(s)}} - \frac{\bar{P}_{(s)}}{\bar{\rho}_{(s)}}\delta_{(s)}\right) \tag{3.31}$$

where '$\prime$' denotes the derivative $\partial/\partial\eta$ and $\mathcal{H} \equiv a'/a$. For $\nu = i$ we find the Euler equation

$$q^{i\prime}_{(s)} = -4\mathcal{H}q^i - (\bar{\rho}_{(s)} + \bar{P}_{(s)})\partial^i\Psi - \partial^i\delta P - \partial_j\Pi^{ij}. \tag{3.32}$$

To study the evolution of Ψ and Φ, we need two independent components of the

Einstein's equations. The first component we consider is the time-time component, which in Fourier coordinates is given by

$$\delta G^0{}_0 = \frac{2}{a^2}\left(-k^2\Phi + 3\mathcal{H}^2\Psi + 3\mathcal{H}\Phi'\right). \tag{3.33}$$

As a result, the 00-component of the Einstein equation is

$$k^2\Phi - 3\mathcal{H}(\mathcal{H}\Psi + \Phi') = 4\pi G_N a^2 \delta\rho. \tag{3.34}$$

The second component is usually conveniently chosen as the longitudinal, traceless part of the spatial part, $G^i{}_j$:

$$\left(\hat{\mathbf{k}}^j\hat{\mathbf{k}}_i - \frac{1}{3}\delta^j_i\right) G^i{}_j = \frac{2}{3a^2}k^2\left(\Phi - \Psi\right) \tag{3.35}$$

which should be equal to the traceless part of the stree-energy tensor

$$\left(\hat{\mathbf{k}}^j\hat{\mathbf{k}}_i - \frac{1}{3}\delta^j_i\right) T^i_j = \left(\hat{\mathbf{k}}^j\hat{\mathbf{k}}_i - \frac{1}{3}\delta^j_i\right) \Pi^i{}_j = \frac{2}{3}(\bar{\rho} + \bar{P})\sigma = \frac{2}{3}\sum_s (\bar{\rho}_{(s)} + \bar{P}_{(s)})\sigma_{(s)}. \tag{3.36}$$

Then the second component of the Einstein equation is given by

$$k^2(\Phi - \Psi) = 8\pi G_N a^2(\bar{\rho} + \bar{P})\sigma. \tag{3.37}$$

3.2.2 Primordial sound waves

During the hot and dense post-inflation period, inhomogeneity leads to overdensity growth in the cosmic plasma. On large scales, gravity dominates, so the fluctuations grow continuously. On small scales, however, the increase in energy density can cause the pressure and its gradient to increase to the point where the pressure exceeds gravity, triggering sound waves that travel outward. The characteristic length at which the pressure is comparable to gravity is called the Jeans length.

In the primordial plasma before recombination, fluctuations of photons, electrons, baryons and dark matter are coupled in a perturbed spacetime. Electrons are tightly coupled to baryons by Coulomb scattering, and photons are coupled to electrons, and hence to baryons, by Thomson scattering. We can thus consider a conserved, tight-coupling photon-electron-baryon fuild, or in short baryon-photon fluid, which can be constrained by the fluid equations $\nabla^\mu(T^{(b)}_{\mu\nu} + T^{(\gamma)}_{\mu\nu}) = 0$. This approximation works well until the recombination. Since the cold dark matter behaves as a pressure-less perfect fluid, we also have $\nabla^\mu T^{(c)}_{\mu\nu} = 0$. The baryon-photon fluid couples to the dark matter fluid by gravity.

We assume that the pressure of the fluid is isotropic so that $\Pi^{ij} = 0$. From the

equation (3.37) we see that the direct result of the assumption is

$$\Psi \approx \Phi. \tag{3.38}$$

The combined momentum density is conserved

$$q^i = \frac{4}{3}\bar{\rho}_\gamma v^i + \bar{\rho}_b v^i = \frac{4(1+R)}{3}\bar{\rho}_\gamma v_\gamma \tag{3.39}$$

where $R \equiv 3\bar{\rho}_b/(4\bar{\rho}_\gamma)$. This conservation is described by the Euler equation

$$[(1+R)\mathbf{v}]' = -\frac{1}{4}\nabla\delta_{(\gamma)} - (1+R)\nabla\Psi. \tag{3.40}$$

Since the total number of photons in the fluid can be considered conserved ($\omega_{(\gamma)} = 1/3$), the continuity equation for photons holds

$$\delta'_{(\gamma)} = -\frac{4}{3}\nabla\cdot\mathbf{v} + 4\Phi'. \tag{3.41}$$

Combining the time derivative of the continuity equation and the spatial derivatives of the Euler equation we get a second order equation only in $\delta_{(\gamma)}$:

$$\delta''_{(\gamma)} + \frac{\mathcal{H}R}{1+R}\delta'_{(\gamma)} + c_s^2 k^2 \delta_{(\gamma)} = -\frac{4}{3}k^2\Phi + 4\Phi'' + \frac{4R'}{1+R}\Phi' \tag{3.42}$$

where the *sound speed* of the baryon-photon fluid is defined as

$$c_s = \sqrt{\frac{1}{3(1+R)}}. \tag{3.43}$$

The physical interpretations of these equations are straightforward. Qualitatively, the coupling with baryons added energy and momentum to the fluid, but did not increase its pressure. Since radiation dominates density and pressure in this period, radiation pressure plays an important role in fluid dynamics. A gravitational over-aggregation of baryonic photonic fluids leads to a significant radiation pressure gradient, which triggers a spherical sound wave in the surrounding fluid medium. [3] The acoustic wave propagates at the speed of sound c_s until the baryons and photons are completely decoupled after recombination. A spherical wave of baryons without 'propellant' thus stops propagating. The longest possible journey of such a sound wave is from the end of inflation to the end of decoupling, and this distance is defined as the *sound horizon* at decoupling:

$$s = \int_0^{\tau_{dec}} \frac{c_s(\tau)}{a(\tau)}\,\mathrm{d}\tau. \tag{3.44}$$

[3] This simplified discussion of BAO considers a single initial overdensity. This is not the real scenario, but illustrates the preferred scale set by the decoupling of photons and baryons.

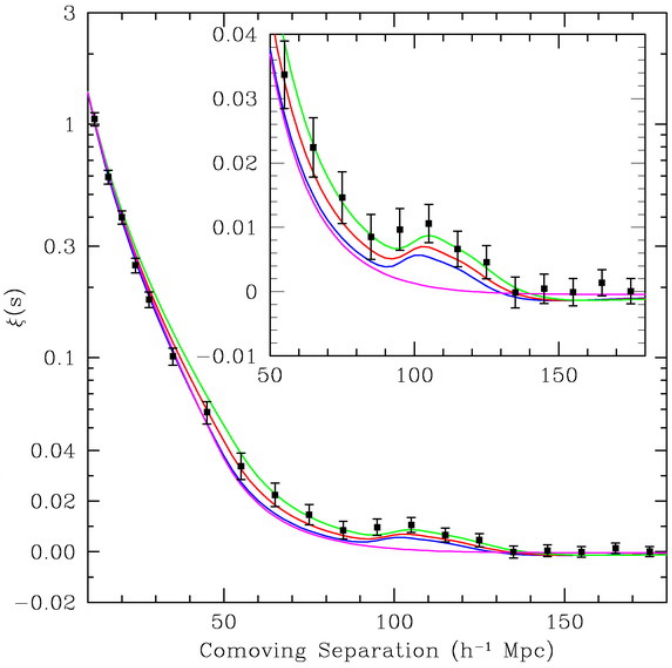

Figure 3.2: Large-scale redshift-space correlation function of the SDSS LRG sample. The inset shows a zoomed view with a linear vertical axis. $\Omega_m h^2 = 0.12$ (green line), 0.13 (red line) and 0.14 (blue line), all with $\Omega_b h^2 = 0.024$ and $n_s = 0.98$ (the spectral index of the initial potential power spectrum, as defined in equation (3.52)). The magenta line shows a pure CDM model ($\Omega_m h^2 = 0.105$) in which the acoustic peak is not present. This figure is taken from Eisenstein et al. (2005).

The sound horizon is a characteristic clustering scale of baryons. It is imprinted in the correlation function as a peak that appears in Fourier space as the harmonic oscillations, also called Baryon Acoustic Oscillations (BAO) (Eisenstein and White, 2004; Eisenstein et al., 2005; Bassett and Hlozek, 2009; Chang et al., 2008; Villaescusa-Navarro et al., 2017; Wyithe et al., 2008). This feature is weak, but very valuable because the CMB observations (Aghanim et al., 2020b) predict its scale very well.

After the decoupling, the baryons cluster with the dark matter. The matter fluid could be described by a pressureless non-relativistic fluid with the conservation equations given by

$$\begin{aligned} \delta'_{(m)} &= -\nabla \cdot \mathbf{v}_{(m)} + 3\Phi' \\ \mathbf{v}'_{(m)} &= -\mathcal{H}\mathbf{v}_{(m)} - \nabla\Psi \end{aligned} \tag{3.45}$$

which derives the equation of motion for δ:

$$\delta''_{(m)} + \mathcal{H}\delta'_{(m)} = \nabla^2\Psi + 3(\Phi'' + \mathcal{H}\Phi'). \tag{3.46}$$

On subhorizon scales with $k \gg \mathcal{H}$, we assume that the time derivatives $\{\Phi'', \mathcal{H}\Phi'\}$ are subdominant and can be dropped, i.e., $k^2\psi \gg \Phi'' + \mathcal{H}\Phi'$. Then the above equation can be rewritten as

$$\delta''_{(m)} + \mathcal{H}\delta'_{(m)} + 4\pi G_N a^2 \delta\rho \approx 0 \tag{3.47}$$

where we replaced $\nabla^2\Psi$ using the 00 component of the Einstein's field equation (eqn (3.34)). $\delta\rho$ is the fluctuation of the total energy density. However, since the dark

energy doesn't contribute to the fluctuations, both in the matter-dominated era or at late times we have

$$\delta\rho \approx \bar{\rho}_{(m)}\delta_{(m)}. \tag{3.48}$$

We can solve the equation of motion for $\delta_{(m)}$ by separating the variables $\delta_{(m)}(\mathbf{x}, z) = D(z)\delta_{\rm in}(\mathbf{x})$, where $\delta_{\rm in}$ is the initial overdensity field. The time component $D(z)$ can be seen as a linear combination of the two modes, the growing mode and the decaying mode. The growing mode is called the linear growth factor, denoted by $D_+(z)$, which is in general an integral with H. The linear growth factor in the full ΛCDM universe is given by (Peebles, 1980):

$$D_+(z) = \frac{5\Omega_{m,0}}{2}H(z)\int_z^\infty \mathrm{d}z' \frac{1+z'}{H(z')^3}. \tag{3.49}$$

In the linear perturbation theory, the evolution of the contrast density field is "self-similar": the overdense regions grow and the underdense regions decay. The small density fluctuations evolve in this way until the perturbation becomes of order unity.

3.2.3 Matter power spectrum as probe of BAO

As a frozen relic left over from the decoupling era, the BAO is an ideal statistical standard ruler in observational cosmology.

The advantage of the BAO as a standard ruler lies not only in the simplicity of linear physics or the scales that are well predicted by the CMB, but more importantly, the transverse and radial components of the BAO provide $d_A(z)$ and $H(z)$, respectively (Bassett and Hlozek, 2009):

$$s_\parallel(z) = \frac{c\,\mathrm{d}z}{H(z)}, \qquad\qquad s_\perp(z) = (1+z)D_A\delta\theta. \tag{3.50}$$

One method to extract the statistical scale is to measure the 2-point correlation function or the power spectrum $P_m(k)$ of the contrast overdensity field. The matter power spectrum is defined as

$$\langle\delta(\mathbf{k}_1)\delta^*(\mathbf{k}_2)\rangle \equiv (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2)P_m(k_1). \tag{3.51}$$

In section 3.2.2 we used the linear perturbation theory to study the evolution of an overdensity field $\delta(k)$, from which the evolution equation of the matter power spectrum is derived. To study the evolution history, we still need the initial condition of the overdensity field.

Inflation theories predict a nearly scale-invariant potential power spectrum $P_\Phi \propto k^{n_s-4}$, where n_s is the spectral index with a value close to 1.[4] Since $\delta(k) \propto k^2\Psi$, the

[4]The measurement from Planck 2018 (Aghanim et al., 2020b) is $n_s = 0.9649 \pm 0.0042$.

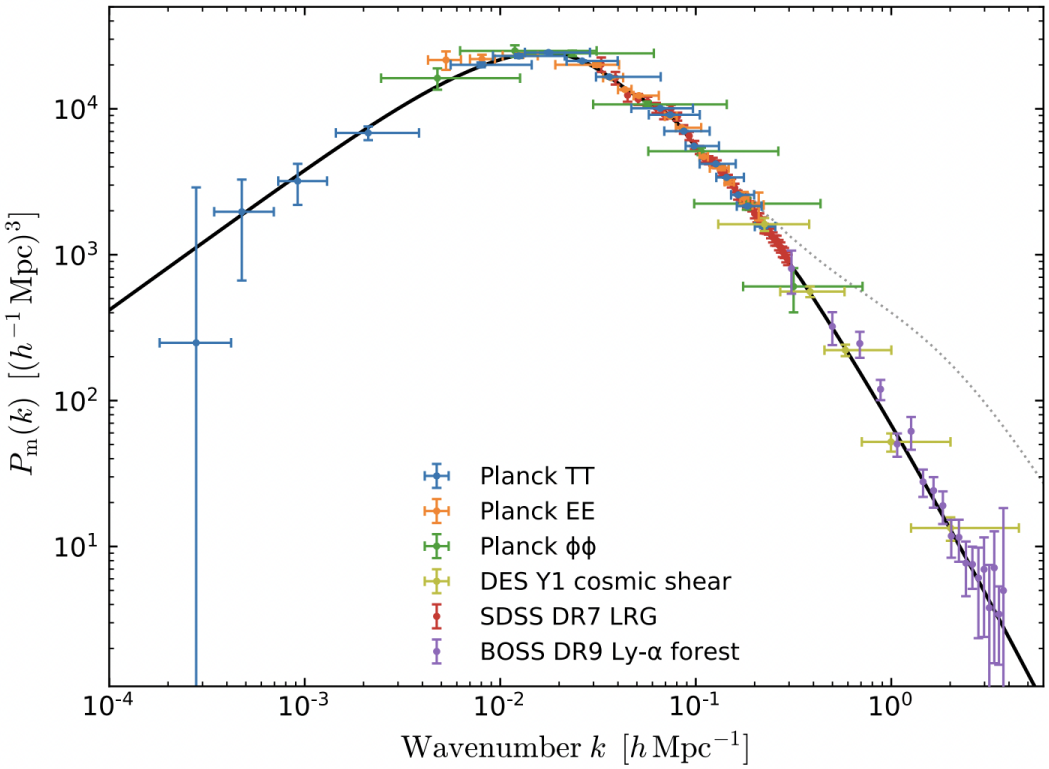

Figure 3.3: Linear matter power spectrum (at $z = 0$) inferred from several cosmological probes. This figure is taken from Aghanim et al. (2020a), where references to similar maps can also be found.

initial matter power spectrum satisfies

$$P_m^{(\rm init)}(k) \propto k^{n_s}. \tag{3.52}$$

As discussed in section 3.2.2, baryons and dark matter evolve together as a matter fluid after the baryon-photon decoupling. Since the Hubble horizon ($\sim H^{-1}$) grows with time, a mode starts to grow as soon as it enters the horizon. Since different scales enter the Hubble horizon at different times, the growth is scale dependent and the effect is usually expressed as the transfer function $T(k, z)$. The linear matter power spectrum can be written as

$$P_m(k, z) \propto D_+^2(z) T^2(k, z) k^{n_s} \tag{3.53}$$

where $D_+(z)$ is the linear growth factor. As shown in Figure 3.3 is the linear matter power spectrum at $z = 0$ inferred from several cosmological probes.

Chapter 4

21cm cosmology and HI intensity mapping

In recent years, particularly over the past decade, there has been a heightened interest in employing 21cm intensity mapping to investigate the large-scale structure of the Universe (Bull et al., 2015; Chang et al., 2010; Villaescusa-Navarro et al., 2018). HI intensity mapping is dedicated to "low redshift" or "late time" cosmology, probing the latter stages of the universe. This technique leverages the optically thin 21cm spectral line, which originates from the hyperfine transition of neutral atomic hydrogen (HI). Notably, atomic hydrogen was pervasive in the universe subsequent to recombination.

While reionization in the later stages of the universe eradicated much of the neutral hydrogen, within galaxies, overdense HI exhibits a recombination rate that is proportional to n^2. This rate is often sufficient to surpass the ionization rate, which is proportional to n and arises due to the pervasive background of ionizing radiation. (See e.g. Draine (2010) for detailed discussion.) Therefore, in the context of larger and denser clouds, it becomes imperative to consider the shielding effects.

When the spin temperature exceeds the CMB temperature, the phenomenon of overdense stimulated emission becomes dominant over absorption. This results in a net 21cm emission, which can be detected as the discrepancy between the observed temperature and the CMB temperature. By monitoring the redshifted 21cm emission line, one can trace the three-dimensional distribution of galaxies. This allows researchers to study the clustering of collapsed halos over a wide range of redshifts and spatial scales. Such observations are invaluable, providing a robust means to refine and constrain cosmological models (Bull et al., 2015).

In this chapter we begin with a detailed exploration of the fundamental and observational facets of 21cm physics in section 4.1. Special emphasis is placed on the spin statistics that drive the 21cm brightness temperature. We then discuss the utility of the HI intensity as a biased tracer of matter on cosmological scales, accompanied by a brief review of major advances in this field. Finally, section 4.3 serves as a general introduction to the challenge of polarization and as an outline for the rest of the thesis.

4.1 Basic 21cm physics

In this section, we will see how the physics of atomic hydrogen and the statistics of abundant hydrogen atoms tell us part of the cosmic history. We will focus on one macroscopic quantity, the ratio of the number densities n_i of hydrogen atoms in the 1s singlet $F = 0$ and 1s triplet $F = 1$ levels of a neutral hydrogen cloud. Here $F = I + S$ where I is the proton spin and S is the electron spin. (Throughout this part we denote 0 and 1 as singlet and triplet levels, respectively). As we will see in the following sections, it is this ratio, characterized by the spin temperature T_S defined in equation (4.3), that determines the detectability of the 21 cm signal.

In general, two types of processes cause the hyperfine transitions of hydrogen atoms and thus change the ratio n_1/n_0: collisional processes and radiative processes. We will first discuss the collisional processes and introduce T_S statistically. Then we will discuss the radiative processes.

4.1.1 Spin statistics of thermal HI gas

Let's first discuss an ideal neutral hydrogen cloud. It's ideal for: 1. There is no background radiation. 2. It's in a state of equilibrium. We need to further restrict what equilibrium means. Obviously it can't be just the kinetic equilibrium, since we're talking about the statistics of the hyperfine states. Equilibrium here also means "hyperfine state equilibrium". The degrees of freedom involved are the kinetic degrees of freedom and the hyperfine states deboted by 1 (aligned spins) and 0 (antiparallel spins). Other hyperfine states beyond 0 and 1 are reasonably not considered.

Collision and decay ($1 \to 0$) are two important regimes in this ideal HI cloud with respect to the 21 cm transition. However, only collision is a statistical behavior, which is the energy redistribution regime here in the ideal system, while decay is an atomic behavior. If the efficiency of collisions is greater than that of decay, one can say that this system would reach and then remain in a kinetic and hyperfine equilibrium state. To be exact, we need to solve a Boltzmann equation that couples the spin and velocity distributions (Hirata and Sigurdson, 2007). But here we just assume that a single ratio of number densities of 1 and 0 can hold for the whole HI gas.

To characterize the ratio n_1/n_0, a macroscopic parameter, one has to choose an appropriate statistical system. It's not a simple choice between the Fermi system, the Bose system, or the classical Boltzmann system. We have to start with the intrinsic statistical scheme of the detailed energy redistribution. In collision processes, the kinetic degrees of freedom are responsible for energy exchange not only among themselves but also with hyperfine degrees of freedom (via H-e^- collisions and H-H collisions). However, the hyperfine degrees of freedom could only exchange energy with the kinetic degrees of freedom, not among themselves, which leads us to consider each hyperfine

degree of freedom separately. For this reason, one can safely imagine the environment of a hyperfine degree of freedom as: a single hyperfine degree of freedom contacting a reservoir with kinetic temperature T_k. For any given hyperfine state, a single frozen degree of freedom, the abundance of possible microstates of the reservoir gives the statistical weight of the hyperfine state,

$$n_i \propto g_i e^{-E_i/k_B T_k}. \tag{4.1}$$

So the ratio n_1/n_0 is

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-E_{10}/k_B T_k}. \tag{4.2}$$

As we can see, the ratio is determined by the kinetic temperature in our ideal system. In general, we can define a parameter, the spin temperature T_S, to characterize n_1/n_0 in any system, even beyond hyperfine equilibrium:

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-E_{10}/k_B T_S}. \tag{4.3}$$

Now we have the generalized definition of T_S. In our toy model, T_S is simply T_K. We might as well always keep this model in mind, where T_S has both statistical and thermal meaning.

Although we have extrapolated a statistical way to characterize the distribution of hyperfine states in a HI cloud, we still need to learn about microscopic collisional processes. Details of the atomic behavior allow us to discuss a real system where all kinds of processes take place and one needs to figure out the contributions of each source. Usually one mainly considers (i) H-H collisions; (ii) H-e^- collisions; (iii) other species. Let C_{10} and C_{01} be the de-excitation ($1 \to 0$) and excitation ($0 \to 1$) rates per atom from collisions, respectively. For each type of collision, the rates are proportional to the number of colliding particles times the cross-section, some function of the kinetic temperature. Therefore, C_{10}/C_{01}, which determines the direction of the hyperfine transition in collisional processes, is also a function of the kinetic temperature.

4.1.2 Spin statistics of HI gas in radiation fields

If we place our ideal hydrogen cloud in a radiation field, further considerations should be made to determine the spin temperature. For astrophysical systems, the radiation field is the CMB and ionized photons in the ISM or IGM. There are two main types of photons involved in the hyperfine transitions between 0 and 1: the 21 cm photons and the Lyα photons. The former are responsible for the direct transitions between 1 and 0, while the latter are responsible for Lyα scattering processes, some of which significantly affect the spin temperature.

We will first describe a general radiative process and then characterize these two

corresponding processes. In the following sections we use the *specific intensity* or *brightness* I_ν to describe the energy carried by rays passing through a given direction, per unit area, per unit frequency, per unit solid angle, and per unit time.

Atomic emission and absorption

Before discussing radiation propagation in general, we should first characterize absorption and emission per atom. Einstein first discovered the relationship between the atomic emission and absorption by proposing three processes: spontaneous emission, stimulated emission, and absorption. They are respectively characterized by the Einstein coefficients A_{21}, B_{21}, B_{12}, where 1 and 2 generally denote two discrete energy levels, E_1 and E_2. A transition from 1 to 2 (2 to 1) occurs by absorption (emission) of a photon of energy $h\nu_0$. Each Einstein coefficient ($\sec^{-1}$) describes the transition probability per unit time for its corresponding process. In fact, there are processes, such as the Doppler effect, that can cause the true energy difference between E_1 and E_2 to differ from $h\nu_0$. This spectral structure peaked at ν_0 is described by a line profile function $\phi(\nu)$. We conventionally use the normalized line profile in our discussion

$$\int_0^\infty \phi(\nu) d\nu = 1. \tag{4.4}$$

In general, $\phi(\nu)$ includes natural, thermal, pressure broadening, etc (Furlanetto et al., 2006; Rybicki and Lightman, 2008).

These Einstein coefficients are essentially atomic properties that don't depend on the macroscopic state of the system. But we can assume that local emission and absorption in a beam of radiation are in detailed balance, which is a reasonable assumption in many interested systems, and this will provide us with an equation to constrain these coefficients. Phenomenologically, the spontaneous emission is independent of brightness I_ν, but stimulated emission and absorption are proportional to I_ν. Thus, the detailed balance describes the equality of the number of transitions per unit time per unit volume from and to state 1:

$$A_{21} n_2 + B_{21} n_2 I_\nu - B_{12} n_1 I_\nu = 0, \tag{4.5}$$

which derives the brightness

$$I_\nu = \frac{A_{21}}{B_{12}(n_1/n_2) - B_{21}}. \tag{4.6}$$

For further discussion of the relations between the Einstein coefficients, we need the distribution functions for n_i, as well as the analytical form of I_ν. The detailed balance implies the hyperfine equilibrium, for which the kinetic equilibrium is a prerequisite.

Substituting n_1/n_0 using equation (4.3) we get

$$I_\nu = \frac{A_{21}}{B_{12}(g_1/g_2)e^{-E_{12}/k_BT} - B_{21}} \tag{4.7}$$

What's more, in thermal equilibrium we also have that the intensity follows Plank's law, which is expressed as

$$B_\nu(T) = \frac{2h\nu^3/c^2}{e^{h\nu/k_BT} - 1}. \tag{4.8}$$

Rendering $I_\nu = B_\nu(T)$ gives the Einstein relations:

$$g_1B_{12} = g_2B_{21}, \quad A_{21} = \frac{2h\nu^3}{c^2}B_{21}. \tag{4.9}$$

Based on the atomic properties discussed above, we can now move on to the macroscopic absorption, emission, and other radiative transfer properties as a beam of radiation travels through matter. We denote the emission coefficient and the absorption coefficient as j_ν and α_ν respectively. To obtain j_ν, one needs to know the frequency distribution of the emitted radiation during spontaneous decay. A simple assumption is that the emission is distributed in accordance with the same line profile $\phi(\nu)$ that describes the absorption (Rybicki and Lightman, 2008).

Assuming that each spontaneous emission is statistically isotropic, the emission coefficient, which describes the amount of energy emitted per unit volume, per unit solid angle, per unit frequency, and per unit time, can be related to A_{21} by

$$j_\nu dV d\Omega d\nu dt = (h\nu_0/4\pi)\phi(\nu)n_2A_{21}dV d\Omega d\nu dt, \tag{4.10}$$

from which we get

$$j_\nu = \frac{h\nu_0}{4\pi}n_2A_{21}\phi(\nu). \tag{4.11}$$

Similarly, we can express α_ν, by definition $dI_\nu = -\alpha_\nu I_\nu ds$, in terms of Einstein's coefficients

$$\alpha_\nu = \frac{h\nu}{4\pi}\phi(\nu)(n_1B_{12} - n_2B_{21}). \tag{4.12}$$

Note that the emission coefficient only takes into account spontaneous emission, while stimulated emission is considered as negative absorption, which is included in the absorption coefficient, simply because these two processes both depend on the brightness I_ν.

The radiative transfer function in terms of emission and absorption coefficients is given by

$$\frac{dI_\nu}{ds} = -\alpha_\nu I_\nu + j_\nu. \tag{4.13}$$

We may define the optical depth

$$\begin{aligned}\tau_\nu(s) &= \int_{s_0}^{s} \alpha_{s'}\, \mathrm{d}s' = \int \mathrm{d}s\, \frac{h\nu}{4\pi}\phi(\nu)(n_1 B_{12} - n_2 B_{21}) \\ &= \int \mathrm{d}s\, \frac{3c^2 A_{21}}{8\pi\nu_{21}^2}\left(n_1\frac{g_2}{g_1} - n_2\right)\phi(\nu) = \int \mathrm{d}s\, \sigma_{21}\left(n_1\frac{g_2}{g_1} - n_2\right)\phi(\nu),\end{aligned} \tag{4.14}$$

where $\sigma_{21} = 3c^2 A_{21}/8\pi\nu_{21}^2$ is the cross section of the atomic transition. The optical depth is given by integrating the absorption coefficient along the path of a traveling beam of radiation. $\tau_\nu > 1$ corresponds to a so-called *optically thick* or *opaque* medium. If $\tau_\nu < 1$, the medium is *optically thin*. Then the radiative transfer equation in terms of τ_ν is

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu, \tag{4.15}$$

where $S_\nu \equiv j_\nu/\alpha_\nu$ is the rescaled source function. For a constant source function

$$I_\nu(\tau_\nu) = I_\nu(0)e^{-\tau_\nu} + S_\nu(1 - e^{-\tau_\nu}). \tag{4.16}$$

Radiative transfer of the 21cm line in the CMB

We now look at the distortion of the CMB passing through a HI cloud as a consequence of the 21cm transition. To characterize this radiative transfer, we need the values of the Einstein coefficients. One can compute A_{10} and then use the Einstein relations to get all these coefficients so that the radiative transfer of the 21 cm line is determined. We assume that $I_\gamma = I_{CMB}$, since almost all 21 cm photons come from the CMB.

J. P. Wild first calculated A_{10} in Wild (1952). The transition probability of the 21 cm line is $A_{10} = 2.85 \times 10^{-15}$ sec^{-1}. Therefore, the optical depth of the 21 cm line can be calculated in terms of A_{10}:

$$\tau_\nu = \int ds\sigma_{10}(1 - e^{-E_{10}/k_B T_S})\phi(\nu)n_0 \approx \sigma_{10}\left(\frac{h\nu}{k_B T_S}\right)\left(\frac{N_{HI}}{4}\right)\phi(\nu), \tag{4.17}$$

where N_{HI} is the column density of atomic hydrogen and the factor 1/4 accounts for the fraction of atoms in the hyperfine 0 state, since all applications have $T_S \gg E_{10}/k_B$.

We can see $\tau_\nu \ll 1$, which implies the transparency of the 21 cm line for the transfer in the neutral hydrogen cloud. This is undoubtedly a non-trivial merit. For a more precise expression, we use a simple assumption, often a good one in astrophysics, that the IGM gas expands uniformly with the Hubble flow. Then the velocity broadening of a line segment s will be $\Delta V \sim sH(z)$ and the line profile function will be $\phi(\nu) \sim c/[sH(z)\nu]$ (Furlanetto et al., 2006). The column density along s can be written as $N_{HI} = x_{HI} n_H(z) s$, where x_{HI} is the neutral fraction of hydrogen. Then an exact

expression for the 21 cm optical depth is

$$\begin{aligned}\tau_{\nu_0} &= \frac{3}{32\pi}\frac{hc^3A_{10}}{k_BT_S\nu_0^2}\frac{x_{HI}n_H}{(1+z)(dv_\parallel/dr_\parallel)} \\ &\approx 0.0092(1+\delta)(1+z)^{3/2}\frac{x_{HI}}{T_S}\left[\frac{H(z)/(1+z)}{dv_\parallel/dr_\parallel}\right].\end{aligned} \tag{4.18}$$

Lyman lines and Wouthuysen-Field effect

In a HI cloud, not only the 21cm photons play an important role in the redistribution of the hyperfine 0 and 1 states, but also the Lyman photons. A ground state atom can absorb a Lyman photon and jump to a higher energy level. After a while, the atom will leave the unstable state and emit a photon. These atoms can decay directly, nP→1S, so that the atom goes back to the ground *fine* state and a Lyman photon is produced, changing the hyperfine states if the initial and final hyperfine states are different. This process is called Lyman resonance, UV scattering, or other combinations of these words. Lyman excited atoms can also cascade through intermediate levels to produce different photons.

A simple comparison between Lyα and higher Lyman-n levels tells us that only Lyα lines have a large influence on the statistics of hyperfine states: The possibilities of direct decay from higher Lyman-n levels to the ground state are $P_{nP\to1S}\sim0.8$, so typically a Lyman-n photon can scatter $1/(1-P_{nP\to1S})\sim5$ times before undergoing a decay cascade (Loeb and Furlanetto, 2013). Lyα photons, on the other hand, can typically scatter hundreds of thousands of times. Thus, higher Lyman-n coupling is subdominant compared to Lyα scattering. Siegfried Wouthuysen and George Field were the first to systematically study Lyα scattering. In their honor, the *Wouthuysen-Field effect*, shown in Figure 4.1, describes how atoms can change hyperfine states through a series of Lyα absorption and spontaneous reemission. Quantum selection rules allow transitions of $\Delta F=0,\pm1$, except $0\to0$. An atomic hydrogen can jump up and down between two fine levels, 1S and 2P, but the hyperfine levels it locates are not unique. So some of this scattering contributes to the spin flip.

To understand the impact of the Wouthuysen-Field effect on the spin temperature, we further discuss the radiative transfer properties of the Lyα photon. Atomic physics tells us that the Lyα decay rate is $A_\alpha=8\pi^2e^2f_\alpha/3m_ec\lambda_\alpha^2=6.25\times10^8\ \mathrm{s}^{-1}$, where $f_\alpha=0.4162$ is the oscillator strength, and the frequency of the Lyα line is $\nu_\alpha=2.47\times10^{15}$ Hz. It is safe to assume that the stimulated emission can be neglected. Therefore, we can express the optical depth of Lyα photons as (Loeb and Furlanetto, 2013)

$$\begin{aligned}\tau_\alpha &= \int dr\sigma_\alpha\phi_\alpha(\nu)n_{\rm HI} = \frac{3A_\alpha\lambda_\alpha^3}{8\pi}\frac{x_{\rm HI}n_H(z)}{H(z)} \\ &\approx 1.6\times10^5x_{HI}(1+\delta)\left(\frac{1+z}{4}\right)^{3/2}.\end{aligned} \tag{4.19}$$

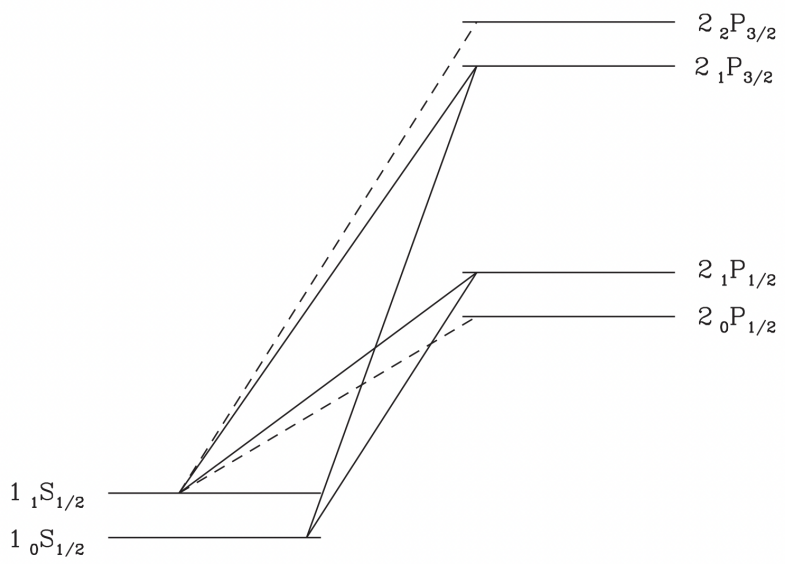

Figure 4.1: Level diagram illustrating the Wouthuysen-Field effect. The solid lines indicate transitions that cause spin flip, and the dashed lines indicate transitions that are not associated with spin flip. This figure is courtesy of Furlanetto et al. (2006)

.

Qualitatively, we can conclude that the neutral hydrogen cloud is optically thick for Lyα lines ($\tau_\alpha \gg 1$). These properties help us to get a picture of the scattering processes: Lyα photons traveling in the HI cloud - not a long trip, then it will be absorbed, but soon after another Lyα photon will emerge. After another short trip, the same process happens again. It can scatter in this way a large number of times, redistributing energy frequently, which led Wouthuysen to propose the a thermal scenario in Wouthuysen (1952): "You can take the gas in a large container with perfectly reflecting walls. Let the gas be in equilibrium at temperature T, together with Planck radiation of the same temperature. The scattering processes will not affect the radiation spectrum", and he further claimed that "after a finite but large number of scattering processes, the photons will obtain a statistical distribution in the vicinity of the initial frequency, proportional to the Planck radiation spectrum of temperature T". This is innovative, although a bit imprecise[1]. This phenomenological explanation is based on the fact that the system is a black body in thermal equilibrium, regardless of whether there is gas in the container, suggests that $T_\alpha \approx T_K \approx T_S$ in thermal equilibrium. In the next section we give a more rigorous discussion of the coupling between these temperatures.

4.1.3 Spin temperature in Tug-of-War

As we have discussed, T_S is determined by several competing processes, collisional processes and radiative processes. Let's use temperatures to characterize all these

[1]Either the specific intensity or the shape of the spectrum can be used to determine the temperature of a blackbody radiation. So not 'proportional' but 'equal'.

forms of energy. Note that a temperature T is essentially a statistical description. Thermal temperature can be transformed into another macroscopic parameter, β:

$$1/k_B T = \beta(E) \equiv \frac{\partial \ln \Omega}{\partial E}, \tag{4.20}$$

where Ω is the number of possible microstates when the system has energy E. By the principle of equal "a priori" probabilities,[2] the β of two systems are equal ($\beta_1 = \beta_2$) when the two systems are in thermal equilibrium. This is the direct result of maximizing the likelihood, or say the entropy, or say the number of microstates of the combined system. Two thermal temperatures are of interest: the kinetic temperature of the baryons, T_K, and the brightness temperature of the background 21cm photons, T_γ.

A peculiar notion of temperature, the temperature of Lyα photons T_α, is also involved. This temperature is not a thermal temperature and is defined by atomic absorption and emission behavior (see equation (4.31)). For later convenience, we also define $T_* \equiv E_{10}/k_B$.

In general, the steady state of hyperfine states when all three processes are involved can be described as follows

$$n_1(A_{10} + B_{10} I_\nu + P_{10} + C_{10}) = n_0(B_{01} I_\nu + P_{01} + C_{01}), \tag{4.21}$$

where $I_\nu = B_\nu(T_\gamma) \approx I_{\rm CMB}$ is the specific intensity of the background 21cm photons. P_{10} and P_{01} are the rates of hyperfine transitions caused by scattering of UV photons, in our case Lyα photons.

The coefficients in the equation (4.21) are not all independent, and the relation between any pair of coefficients that are inverse processes of each other can be determined. From the equation (4.9) we get

$$\frac{B_{01}}{B_{10}} = 3, \qquad \frac{A_{10}}{B_{10}} = \frac{2h\nu^3}{c^2} = \frac{2k_B\nu^2}{c^2}\frac{h\nu}{k_B} = \frac{\partial B_\nu(T)}{\partial T} T_*. \tag{4.22}$$

where we have applied the Rayleigh Jeans limit in the last equality. The ratio C_{01}/C_{10} can be obtained by considering a thermal equilibrium gas with only the collisional processes. Since $T_* \ll T_K$ in the collisional regime, the equation (4.2) gives

$$\frac{C_{01}}{C_{10}} \approx 3\left(1 - \frac{T_*}{T_K}\right). \tag{4.23}$$

The ratio of P_{01} and P_{10} is a bit more complicated. Labeling the 1S and 2P hyperfine

[2] This is not Bayesian statistics. However, the terminology is commonly used as the first postulate of statistical mechanics.

levels 0-5 from lowest to highest energy, we have

$$P_{01} \propto B_{03} I_{\nu_{03}} \frac{A_{31}}{A_{30}+A_{31}} + B_{04} I_{\nu_{04}} \frac{A_{41}}{A_{40}+A_{41}} \tag{4.24}$$

and

$$P_{10} \propto B_{13} I_{\nu_{13}} \frac{A_{30}}{A_{30}+A_{31}} + B_{14} I_{\nu_{14}} \frac{A_{40}}{A_{40}+A_{41}} \tag{4.25}$$

Using the relations between the Einstein coefficients, we have

$$B_{03} I_{\nu_{03}} = \frac{g_3 B_{30} I_{\nu_{03}}}{g_0} = A_{30} \frac{g_3}{g_0} \frac{T_b(\nu_{03})}{T_{03}} = \frac{g_3}{g_0} A_{30} \mathcal{N}(\nu_{03}), \tag{4.26}$$

where $T_{03} \equiv h\nu_{03}/k_B$ and $T_b(\nu_{03})$ is the brightness temperature of $I_{\nu_{03}}$ in the Rayleigh-Jeans limit. The statistical weight $g = 2F+1$ for an nFJ state. The ratio $\mathcal{N}(\nu_{03}) \equiv T_b(\nu_{03})/T_{03}$ can be understood as the photon occupation number at ν_{03}. Similar results can be found for B_{03}, B_{13}, and B_{14}. Using a sum rule for hyperfine transitions, which states that the sum of all transitions from given nFJ to a given $n'J'$ but arbitrary F' is proportional to $2F+1$, we can get the ratios

$$A_{30} : A_{31} : A_{40} : A_{41} = 1 : 2 : 2 : 1. \tag{4.27}$$

Then the ratio between P_{01} and P_{10} can be written as

$$\begin{aligned} \frac{P_{01}}{P_{10}} &= \frac{\mathcal{N}(\nu_{03}) \frac{g_3}{g_0} \frac{A_{30}A_{31}}{A_{30}+A_{31}} + \mathcal{N}(\nu_{04}) \frac{g_4}{g_0} \frac{A_{40}A_{41}}{A_{40}+A_{41}}}{\mathcal{N}(\nu_{13}) \frac{g_3}{g_1} \frac{A_{30}A_{31}}{A_{30}+A_{31}} + \mathcal{N}(\nu_{14}) \frac{g_4}{g_1} \frac{A_{40}A_{41}}{A_{40}+A_{41}}} \\ &= 3 \frac{\mathcal{N}(\nu_{03}) + \mathcal{N}(\nu_{04})}{\mathcal{N}(\nu_{13}) + \mathcal{N}(\nu_{14})}. \end{aligned} \tag{4.28}$$

Since $\nu_{0a} = \nu_{1a} + \nu_{01}$ and $\nu_{01} \ll \nu_{1a}$ for $a = \{3, 4\}$, the above expression can be expanded to the first order as

$$\begin{aligned} \frac{P_{01}}{P_{10}} &= 3 \frac{\mathcal{N}(\nu_{03}) + \mathcal{N}(\nu_{04})}{\mathcal{N}(\nu_{13}) + \mathcal{N}(\nu_{14})} \\ &= 3 \left(1 + \frac{\partial_\nu \mathcal{N}(\nu_{13}) + \partial_\nu \mathcal{N}(\nu_{14})}{\mathcal{N}(\nu_{13}) + \mathcal{N}(\nu_{14})} \nu_{01} \right) \end{aligned} \tag{4.29}$$

For later convenience we define the effective color temperature T_α so that

$$\frac{P_{01}}{P_{10}} \equiv 3 \left(1 - \frac{T_*}{T_\alpha} \right). \tag{4.30}$$

The equation (4.29) gives the explicit expression for T_α:

$$T_\alpha = -\frac{h}{k_B} \frac{\mathcal{N}(\nu_{13}) + \mathcal{N}(\nu_{14})}{\partial_\nu \mathcal{N}(\nu_{13}) + \partial_\nu \mathcal{N}(\nu_{14})}. \tag{4.31}$$

We are now ready to express the spin temperature more explicitly as a combination of other temperatures. If the populations of the hyperfine states 0 and 1 do not change, we have

$$3\left(1-\frac{T_*}{T_S}\right)\approx\frac{n_1}{n_0}=\frac{B_{01}I_\nu+P_{01}+C_{01}}{A_{10}+B_{10}I_\nu+P_{10}+C_{10}}, \tag{4.32}$$

which gives

$$\begin{aligned} T_S^{-1} &\approx \frac{A_{10}T_*^{-1}+P_{10}T_\alpha^{-1}+C_{10}T_K^{-1}}{A_{10}(T_\gamma/T_*)+P_{10}+C_{10}} \\ &= \frac{T_\gamma^{-1}+x_\alpha T_\alpha^{-1}+x_c T_k^{-1}}{1+x_\alpha+x_c}, \end{aligned} \tag{4.33}$$

where

$$x_c=\frac{T_*}{T_\gamma}\frac{C_{10}}{A_{10}},\ \text{and}\ \ x_\alpha=\frac{T_*}{T_\gamma}\frac{P_{10}}{A_{10}}. \tag{4.34}$$

Conventionally, we relate the scattering rate P_{01} to the total Lyα scattering rate P_α, the total scattering rate per atom of Lyα photons (Furlanetto et al., 2006)

$$P_\alpha=4\pi\chi_\alpha\int d\nu J_\nu(\nu)\phi_\alpha(\nu), \tag{4.35}$$

where $\chi_\alpha = 0.4162\pi e^2/m_e c$ is the cross section, $J_\nu(\nu)$ is the angle-averaged specific intensity of the radiation field, and $\phi_\alpha(\nu)$ is the Lyα line profile. The sum rule gives $P_{01} = 4P_\alpha/27$ if the radiation field is assumed to be constant (Meiksin, 2000). The coupling coefficient can now be rewritten as (Furlanetto et al., 2006)

$$x_\alpha=\frac{16\pi^2T_*e^2f_\alpha}{27A_{10}T_\gamma m_e c}S_\alpha J_\alpha, \tag{4.36}$$

where S_α is a correction of order unity and J_α is the specific flux evaluated at Lyα frequency.

4.1.4 21cm signal

We quantify I_ν by the equivalent brightness temperature, $T_b(\nu)$, such that $I_\nu = B_\nu(T_b)$. In the Rayleigh-Jeans limit where

$$I_\nu=\frac{2\nu^2}{c^2}kT_b, \tag{4.37}$$

the solution of the radiative transfer equation can now be written in terms of $T_b(\nu)$

$$T_b'(\nu)=T_{ex}(1-e^{-\tau_\nu})+T_R'(\nu)e^{-\tau_\nu}, \tag{4.38}$$

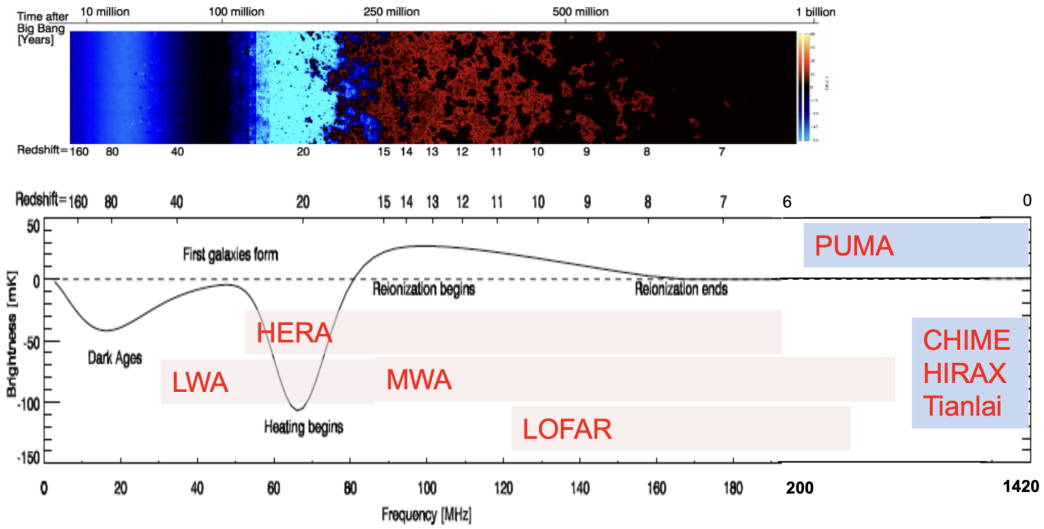

Figure 4.2: Global history of the 21 cm signal. This is taken from Pritchard and Loeb (2012).

where the excitation temperature T_{ex} is just the spin temperature T_S, and $T'_R = T_{CMB}$. The desired 21cm signal is then

$$\begin{aligned}\delta T_b &= \frac{T_S - T_\gamma}{1+z}(1 - e^{-\tau_\nu}) \approx \frac{T_S - T_\gamma}{1+z}\tau_\nu \\ &\approx 27 x_{HI}(1+\delta)\left(\frac{\Omega_b h^2}{0.023}\right)\left(\frac{0.15}{\Omega_m h^2}\frac{1+z}{10}\right)^{1/2}\left(\frac{T_S - T_\gamma}{T_S}\right),\end{aligned} \tag{4.39}$$

where x_{HI} is the neutral fraction of hydrogen, δ is the fractional overdensity in baryons (Pritchard and Loeb, 2012).

History of 21cm signal

In this section we look at different epochs when different mechanisms dominate the spin temperature. As shown in Figure 4.2, there are several important points in time. They are at redshifts: $z \approx 200$, when gas and photons are about to decouple; $z \approx 30$, the beginning of the Dark Ages; $z_\star$, when the first galaxies form; z_α, when the gas is everywhere strongly coupled to T_K; z_h, when the heating is significant and T_K is again equal to T_γ; z_T, the time when the 21 cm signal saturates; and z_R, about the time when reionization is complete. These epochs are not exactly determined, and even the order of some events is not certain (see e.g. Pritchard and Loeb (2012) for a more detailed discussion).

- When $200 \lesssim z \lesssim 1100$, the Compton scattering dominates the thermal equilibrium, setting $T_S = T_K = T_\gamma$. So there is no detectable 21 cm signal. After that, photons and baryons decouple but the baryon fluid is still dense and the collisional coupling dominates, so we still have $T_S \approx T_K$.

- Since the photons cool with the Hubble flow with $T_\gamma \propto (1+z)$ and the baryons cool with $T_K \propto (1+z)^2$, there is a time, about $40 \lesssim z \lesssim 200$, when T_K is less than T_γ. This epoch has a global 21cm absorption signal.

- When $z_\star \lesssim z \lesssim 40$, as the gas density decreases, the radiative coupling gradually plays an more important role than the collisional coupling, which sets $T_S = T_\gamma$. Thus, there is no detectable 21 cm global signal in this epoch. $z_\star$ is an important critical point when the first sources appear, emitting both Lyα photons and x-rays.

- In regions around the first sources appeared at $z = z_\star$, the Wouthuysen-Field effect becomes important and sets $T_S \approx T_K$. The Lyα coupling may not be as efficient as other couplings at the beginning. As more and more stars form, the Lyα coupling becomes stronger until saturated. Thus, the spin temperature in this epoch provides an absorption signal as $\bar{T}_K \lesssim \bar{T}_S \lesssim T_\gamma$. At the critical point, $z = z_\alpha$, the Lyα coupling dominates and $\bar{T}_S = \bar{T}_K$.

- The absorption signal in the previous epoch exists until $\bar{T}_K$ increases to be equal to T_γ, which happens at $z = z_h$ due to the heating regime in the structure formation. Immediately after z_h, $T_S \sim T_K > T_\gamma$ and we can see emission signal. The 21cm emission signal saturates and becomes independent of T_S when $T_K \gg T_\gamma$. This can be seen from equation (4.39) in the limit $(T_S - T_\gamma)/T_S \to 1$.

4.2 HI intensity mapping

4.2.1 21cm power spectrum

Radio intensity mapping, the focus of this thesis, which maps large-scale structure in the redshifted 21 cm emission of atomic hydrogen without resolving individual galaxies, offers a promising and economical way to measure BAO over large volumes (Bull et al., 2015) . As a redshift survey, the observed HI power spectrum also suffers from the redshift space distortion (RSD) (Kaiser, 1987) effect like traditional galaxy surveys. The distortion can be empirically written as

$$P_{\rm HI}(k,z) \propto (b_{\rm HI} + f\mu^2)^2 \exp\left(-k^2\mu^2\sigma_{\rm NL}^2\right) {D_+}^2(z) T^2(k,z) k^{n_s}, \tag{4.40}$$

where $\mu \equiv k_\parallel/k$ is the cosine of the line-of-sight angle and $b_{\rm HI}$ is the bias due to the use of HI as a tracer of matter. The linear growth rate f is related to the linear growth factor $D_+(z)$ by $f = d\ln D_+/d\ln a$. The first term $b_{\rm HI} + f\mu^2$ is the linear regime (Percival and White, 2009) of the RSD of the HI overdensity. While the exponential term accounts for the 'Fingers of God' effect, where an uncorrelated velocity dispersion

growing on small scales washes out the radial fluctuations beyond the cutoff set by the nonlinear dispersion, $\sigma_{\rm NL}$.

4.2.2 Evidence for HI in galaxies

As introduced at the chapter's outset, galaxies can host overdense HI regions with elevated spin temperatures, primarily due to shielding effects. In this section, we provide a concise overview of the presence of HI regions within galaxies.

The space that separates the stars and fills the galaxies is diluted but far from empty, called the interstellar medium (ISM), a mixture of gas, high energy particles and dust grains. The gas is mainly hydrogen, which makes up about 70% of the mass, followed by helium, which makes up 28% of the mass. The remaining 2% is made up of heavier elements, commonly known as metals (Draine, 2010).

In the 1920s and 30s, the first evidence for a ubiquitous ISM came from observations of absorption lines at visible wavelengths (Heger, 1922; Merrill, 1934). By the 1940s, many visible resonance lines from ions and molecules had been discovered, including NaI, CaII, TiII, CaI, KI, LiI, CH, NH, CN, CH^+, and C_2. More absorption lines were found in UV observations, e.g., Copernicus, IUE and HST, because the typical excitation energies of the ground state resonance transitions are a few eV. In addition to UV and visible observations, longer wavelength millimetre and shorter X-ray bands have been used to probe the ISM: millimetre observations reveal the presence of Giant Molecular Clouds at temperatures below 100 K, and X-ray observations reveal the presence of gas at temperatures above 10^5K. The existence of the 21cm line of atomic hydrogen was first predicted in Van de Hulst (1945) and, with the development of radio astronomy, was later detected from our Galaxy (Ewen and Purcell, 1951) and the Magellanic Clouds (Kerr et al., 1954). In general, all these observations at different wavelengths show that the ISM is highly inhomogeneous and can be divided into phases with different temperatures, densities and ionisation ratios (Draine, 2010). In Table 4.1, we summarise a typical classification of different phases of interstellar gas. The constituents of the ISM are inevitably present between galaxies, the so-called intergalactic medium (IGM), since there is no well-defined boundary to a galaxy. The classification of the interstellar gas is of course extended to the case of the intergalactic medium.

21cm intensity observations provide information on the abundance and distribution of HI, which is the dominant species in two neutral thermal phases in the ISM:

Warm Neutral Medium (WNM), composed of the diffuse, relatively low-density ($n_H \sim 0.6\text{cm}^{-3}$) atomic neutral hydrogen gas with a temperature $T \sim 8000$K. The WNM is thought to occupy $\sim 30-40\%$ of the volume of the Galactic disc and to carry a significant fraction of the total mass of neutral hydrogen.

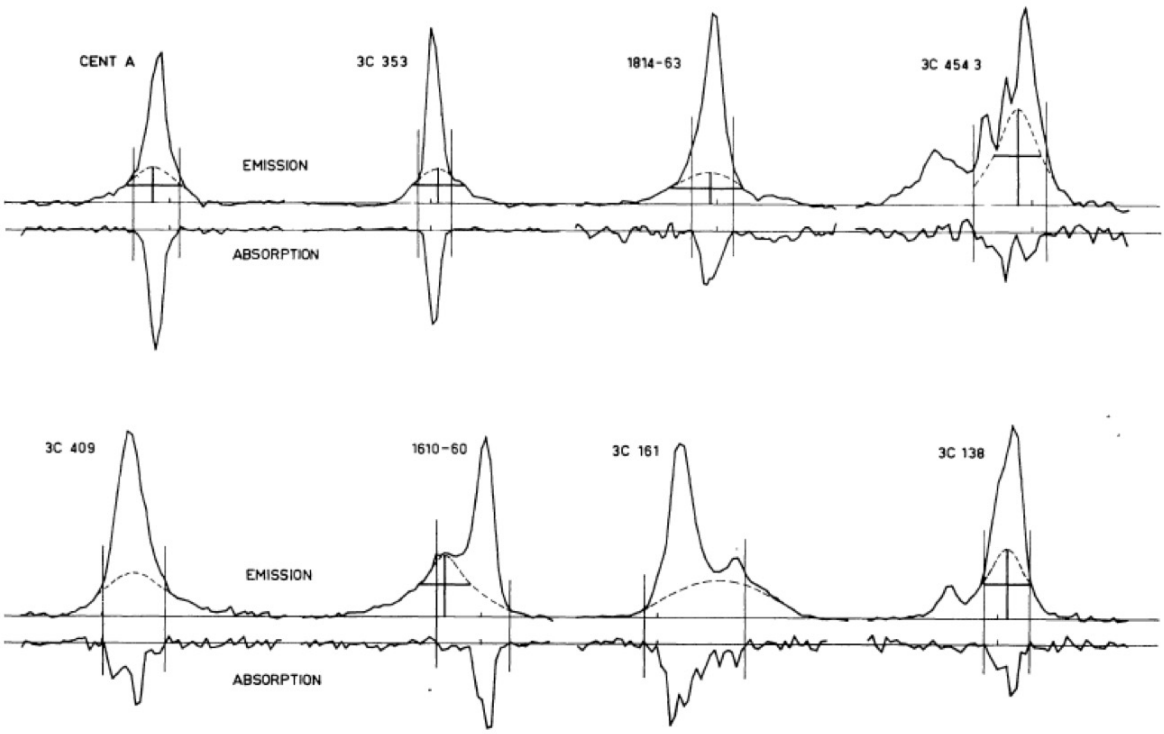

Figure 4.3: HI 21cm emission and absorption signals of 8 extragalactic radio sources from Radhakrishnan et al. (1972). This figure is taken from Ryden and Pogge (2021). The vertical axis is the contrast brightness temperature against the neighbourhood radio background. Dashed lines are fits to optically thin emission line components where there is no corresponding absorption component. The velocity limits of the optically thick absorption components are indicated by the vertical lines. The narrow absorption lines correspond to discrete CNM regions. The broad, optically thin emission is diffuse WNM.

Cold Neutral Medium (CNM), composed of the cold ($T < 100$K), dense ($n = 20-60\text{cm}^{-3}$) clouds and filaments with a low filling factor ($\sim 1-4\%$). It occupies a very small percentage of the volume of the Galactic disc. However, it is so dense that it accounts for $30-60\%$ of the mass of the ISM. The dense molecular clouds are a colder ($T \sim 10-20$K) and denser ($n_H \sim 103-106\text{cm}^{-3}$) environment within the CNM, bounded by self-gravity. It provides a home for complex molecules, a playground for dust chemistry, and a place for star formation.

The integrated 21cm emission and absorption of HI from CNM and WNM in galaxies was observed and studied in the 1970s. As an illustrative case, Figure 4.3 shows line profiles of 21cm emission and absorption towards 8 extragalactic radio sources (Radhakrishnan et al., 1972). The narrow absorption lines correspond to discrete CNM regions, and the broad, optically thin emission is diffuse WNM (Ryden and Pogge, 2021).

4.2.3 Experimental progress

In 2010, Chang et al. made the pioneering detection of the 21 cm signal within the intensity mapping (IM) domain using the Green Bank Telescope (GBT). Their work showcased a 3D 21 cm intensity field spanning redshifts $z = 0.53$ to 1.12, overlapping with approximately 10,000 galaxies from the DEEP2 galaxy survey (Davis et al., 2003). This crucial observation confirmed the 21 cm intensity field around $z \approx 1$ as an apt reflection of galaxy distribution.

Progressing on this foundation, Masui et al. (2013) identified a cross-power spectrum around $z \approx 0.8$, bridging 21 cm IM with galaxies from the WiggleZ Dark Energy Survey. Switzer et al. (2013) set a novel benchmark by identifying the upper limit of the 21 cm auto-power spectrum. Additionally, Anderson et al. (2018) correlated a cross-power spectrum between Parkes telescope's 21 cm intensity maps and the 2dF galaxy maps at $z \approx 0.08$.

In a significant stride, Paul et al. (2023) reported the first direct detection of the cosmological power spectrum using the 21-cm emission of neutral hydrogen (HI) with the new MeerKAT radio telescope. This groundbreaking work, focused on intensity mapping, has painted a vivid three-dimensional portrait of the universe's matter distribution at radio frequencies. Remarkably, unlike previous efforts which relied on cross-correlations with galaxy surveys, this study offers independent measurements of the HI power spectrum at redshifts 0.32 and 0.44, boasting high statistical significance (8.0σ and 11.5σ respectively). Their exploration also presents constraints on the fluctuations of the HI distribution, further elucidating the parameters of the HI mass function and HI halo model. Undoubtedly, this represents a monumental advancement towards precision cosmology with HI intensity mapping using modern radio telescopes (Paul et al., 2023).

Despite these strides, the full auto-correlated 21 cm power spectrum remains a tantalizing goal. Here we highlight several trailblazing experiments targeting the redshift range $0 < z < 3$: BINGO (Dickinson, 2014), FAST (Smoot and Debono, 2017), SKA (Santos et al., 2015; Bacon et al., 2020), HIRAX (Newburgh et al., 2016), CHIME (Amiri et al., 2022), Tianlai (Chen et al., 2015), and the aforementioned MeerKAT. Collectively, these initiatives aim to elucidate the potential of low-redshift 21 cm IM experiments in enhancing our grasp of cosmological parameters.

Although we have mentioned that HI intensity mapping is dedicated to "low redshift" or "late time" cosmology, probing the latter stages of the universe, specifically within the $0 < z < 3$ bracket, the expansive capabilities of the higher redshifts (e.g., the SKA Low Array) are recognized and hold the promise of extending 21cm IM studies even beyond $z = 3$ (Barry et al., 2022).

4.3 Polarization challenge

Detecting the 21 cm intensity (or Stokes I)[3] presents a formidable challenge. The desired extragalactic signal is consistently overshadowed by galactic foregrounds that are three to four orders of magnitude brighter. However, the inherent spectral smoothness of the total intensity of the Galactic foreground offers a glimmer of hope. Theoretically, this smoothness allows the extraction of the cosmological signal with its complex spectral behavior. Depending on the specific analysis technique, one can either fit the foreground with models of smooth spectral functions, or extract data at higher wavenumbers.

Since the desired cosmological signal is inherently unpolarized, polarization should, in theory, not affect unpolarized intensity measurements. But the reality of radio observations, even unpolarized ones, is revealed by polarimetry. Given the nuanced nature of the cosmological 21cm signal, imperfections are hard to avoid. Instrumental systematics typically manifests itself in two dominant problems: **mode mixing** and **polarization leakage**. The former, which has been studied in the context of beam chromaticity in several papers (see, e.g., Thyagarajan et al. (2016)), which allows some low wavenumber (conjugate to frequency, foreground-dominated) modes to leak into components that vary rapidly with frequency. In part V of this paper, we will see that general beam uncertainties can also cause the mixing of transverse spatial modes. Polarization leakage, on the other hand, captures polarimetric imperfections, whether due to misaligned feeds, dish manufacturing anomalies, or beam uncertainties. All of these challenges can essentially be characterized as beam discrepancies in sky coordinates.

The scenario becomes even more complicated when considering the spectral structure of polarized foregrounds. As these highly polarized foregrounds traverse the ISM under the influence of the galactic magnetic field, they undergo Faraday rotation. This interaction gives the polarized foreground a tangled spectral characteristic that confounds foreground mitigation efforts that are based on its spectral smoothness. Due to these complexities, polarization leakage stands out as a significant obstacle in several 21cm experiments, including HI intensity mapping.

The goal of this thesis is to deconstruct the role of polarization in HI intensity mapping and, more generally, in 21cm surveys. Our investigation will focus on

1. Understanding antenna measurements based on radio polarimetry (part II).
2. Understanding the polarization of the diffuse galactic foreground (part III).
3. Discussing the Stokes I extraction mechanism within polarization measurements (part IV).

[3] See section 5.1.1 for a comprehensive discussion of polarization.

4. Evaluating the effects of polarized beam uncertainties in intensity mapping experiments (part V).

Table 4.1: Phases of interstellar gas

Phase	T (K)	n_H (cm^{-3})	Observed by
hot ionized medium (HIM)	$\gtrsim 10^{5.5}$	~ 0.004	• UV and X-ray emission • Radio synchrotron emission
HII gas	10^4	$0.3 - 10^4$	• Optical line emission • Thermal radio continuum
Warm HI (WNM)	~ 5000	0.6	• **HI 21cm emission, absorption** • Optical, UV absorption lines
Cool HI (CNM)	~ 100	30	• **HI 21cm emission, absorption** • Optical, UV absorption lines
Diffuse H_2	~ 50	~ 100	• HI 21cm emission, absorption • Optical, UV absorption lines • CO 2.6mm emission
Dense H_2	10-50	$10^3 - 10^6$	• CO 2.6mm emission • Dust FIR emission
Cool stellar outflows	$50 - 10^3$	$1 - 10^6$	• Optical, UV absorption lines • Dust IR emission • HI, CO, OH radio emission

* n_H is the density of hydrogen in the phase of gas.
* The information in this table is taken from Draine (2010).

Part II

Polarization measurement for radio cosmology

Chapter 5

The polarized radio sky

In this Part II of this thesis, we introduce the foundational principles of polarization measurements in radio cosmology. We begin by establishing the mechanisms that describe the radio signal. The directional signal received by the antenna is simplified as a time-varying electric field signal on the celestial sphere, and the portion of its radio frequency band (from 3kHz to 3GHz) is referred to as the radio sky. In this chapter, we provide an overview of the description of the polarization field in the radio sky.

5.1 Polarization of electromagnetic waves and Stokes parameters

5.1.1 Monochromatic waves

Classical free electromagnetic waves can be described as a superposition of plane-wave solutions of Maxwell equations. The time evolution of a single chromatic wave is given by the temporal frequency ω, while the spatial configuration can be described by the wave vector $\mathbf{k}$, which characterizes the spatial frequency and the direction of the phase advance, and the orientation of the EM vectors in the transverse directions. The latter is called the *polarization* of the EM wave. We will discuss only the polarization of the electric vector $\mathbf{E}$, since the magnetic vector simply remains perpendicular to $\mathbf{E}$.

The most general state of polarization for a single free monochromatic wave of given $\mathbf{k}$ and ω can be described by a pair of orthogonal unit vectors, denoted as $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$, in the plane perpendicular to $\mathbf{k}$. Then the electric vector can be written as the real part of

$$\mathbf{E} = \mathcal{E}^1 \hat{\mathbf{e}}_1 + \mathcal{E}^2 \hat{\mathbf{e}}_2 = (\mathcal{A}_1 e^{i\phi_1} \hat{\mathbf{e}}_1 + \mathcal{A}_2 e^{i\phi_2} \hat{\mathbf{e}}_2) e^{i\omega t}. \tag{5.1}$$

It is obvious that in order to determine the magnitude and orientation of $\mathbf{E}$ at a given moment, we have to determine four quantities: the amplitudes $\mathcal{A}_1$ and $\mathcal{A}_2$ and the absolute phases ϕ_1 and ϕ_2. However, the orientation of $\mathbf{E}$ at a given moment may not be of primary interest to astrophysicists, who are typically interested in the

local[1] directional behavior of EM waves over time. In other words, the absolute values of ϕ_1 and ϕ_2 are neither directly observable nor of interest in most radio astronomy applications, but their relative phase $(\phi_1 - \phi_2)$ is usually a desired quantity.

Therefore, to operationally define the *polarization state* of a monochromatic wave at a given time t and location $\mathbf{x}$, we can consider the local directional behavior of $\mathbf{E}$ within a time period centered at $(t - \tau/2, t + \tau/2)$, where $\tau \gg 1/\omega$. This involves determining three degrees of freedom within the time period: the amplitude $\mathcal{A}_1$, which can be determined by $\langle \mathcal{E}^1 \mathcal{E}^{1*} \rangle$; the amplitude $\mathcal{A}_2$, which can be determined by $\langle \mathcal{E}^2 \mathcal{E}^{2*} \rangle$; and the phase difference $\phi_1 - \phi_2$, which can be determined by $\langle \mathcal{E}^1 \mathcal{E}^{2*} \rangle$ once $\mathcal{A}_1$ and $\mathcal{A}_2$ are known. Here $\langle \dots \rangle$ means averaging over the time period τ. Since the endpoint of $\mathbf{E}$ traces out an ellipse, a monochromatic wave is said to be *elliptically polarized*. Thus, the state of polarization can be described in terms of the parameters of the polarization ellipse.

To facilitate later discussion, we introduce the *polarization tensor* $\mathcal{P}$ to characterize the polarization of any electromagnetic wave:

$$\mathcal{P} \equiv \langle \mathbf{E} \otimes \mathbf{E}^* \rangle = \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \, \langle \mathcal{E}^i \mathcal{E}^{j*} \rangle. \tag{5.2}$$

where the repeated indices are summed using the Einstein summation convention. Since the tensor components $\langle \mathcal{E}^i \mathcal{E}^{j*} \rangle$ are a 2×2 Hermitian matrix, $\mathcal{P}$ can always be uniquely expanded in terms of conventional Pauli spin matrices as (see e.g. Tinbergen (2005))

$$\mathcal{P} = \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \, \left[\sigma^I_{ij} \, I + \sigma^Q_{ij} \, Q + \sigma^U_{ij} \, U + \sigma^V_{ij} \, V \right] \tag{5.3}$$

where $\sigma^I = \sigma_0$, $\sigma^Q = \sigma_3$, $\sigma^U = \sigma_1$, and $\sigma^V = \sigma_2$.[2] Equations (5.2) and (5.3) give

$$\begin{aligned} I &\equiv \langle \mathcal{E}^1 \mathcal{E}^{1*} \rangle + \langle \mathcal{E}^2 \mathcal{E}^{2*} \rangle, \;\; Q \equiv \langle \mathcal{E}^1 \mathcal{E}^{1*} \rangle - \langle \mathcal{E}^2 \mathcal{E}^{2*} \rangle, \\ U &\equiv \langle \mathcal{E}^1 \mathcal{E}^{2*} \rangle + \langle \mathcal{E}^2 \mathcal{E}^{1*} \rangle, \;\; V \equiv -i(\langle \mathcal{E}^1 \mathcal{E}^{2*} \rangle - \langle \mathcal{E}^2 \mathcal{E}^{1*} \rangle). \end{aligned} \tag{5.4}$$

An important observation from the above expressions is that I, Q, U, V correspond exactly to the Stokes parameters, as defined based on the polarization ellipse. Since $\mathcal{A}_1$, $\mathcal{A}_2$, ϕ_1, and ϕ_2 are constant, the relation $I^2 = Q^2 + U^2 + V^2$ holds for all monochromatic waves.

A monochromatic wave with a given $\mathbf{k}$ and ω is in general purely elliptically polarized. The polarization state is defined over a time interval and can be specified by the polarization ellipse, which in turn can be described by the Stokes parameters. These parameters can be determined operationally using $\langle \mathcal{E}^i \mathcal{E}^{j*} \rangle$.

[1] By the term "local" the author refers specifically to the time-varying orientation of $\mathbf{E}$ at a given position, rather than its orientation at a wavefront.

[2] Pauli spin matrices are $\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

5.1.2 Quasi-monochromatic waves

In section 5.1.1, we discussed the directional behavior of the electric vector of freely propagating monochromatic waves. This nonrandom behavior is assumed by forcing constant values of $\mathcal{A}_1, \mathcal{A}_2, \phi_1$, and ϕ_2 in equation (5.1). This assumption is based on the coherent nature of radio emissions, which can be described as $t_{\rm co} > \tau \gg 1/\omega$, where $t_{\rm co}$ is the coherent time of the EM wave. However, this assumption does not always hold: some degree of incoherence in the observed signal is inevitable due to changes in the physical conditions of the radiation source itself or the modifications of the EM waves along the propagation path. In such cases, the incoherence must be taken into account, and the electric vector components should be expressed as

$$\mathcal{E}^i(\omega, t) = \mathcal{A}_i(t) e^{i\phi_i(t)} e^{i\omega t} \tag{5.5}$$

where $\mathcal{A}_i$, and ϕ_i are now time-dependent variables. Fortunately, most radio astrophysical signals are coherent on time scales much larger than $1/\omega$, so we can easily find an integration time τ over which $\mathcal{A}_i$ and ϕ_i are approximately constant.

Another problem is more important than the coherence of the radio source: In practice, we always observe a superposition of many frequency components rather than a single monochromatic component, each with its own polarization. Assuming that the combination of the frequency components is described by a normalized narrow top-hat filter, and that each component is coherent so that $\mathcal{A}_i$ and ϕ_i are considered constants over the time period τ, the polarization components of the electric vector of this set of "quasi-chromatic" waves can be expressed as (Rybicki and Lightman, 2008)

$$\mathcal{E}^i(\omega, t) = \int_{\omega - \frac{\Delta\omega}{2}}^{\omega + \frac{\Delta\omega}{2}} \frac{\mathrm{d}\omega'}{\Delta\omega} \mathcal{A}_i(\omega') e^{i\phi_i(\omega')} e^{i\omega' t} \tag{5.6}$$

The corresponding correlation is given by

$$\langle \mathcal{E}^i \mathcal{E}^{j*} \rangle(\omega, t) = \int_{t - \frac{\tau}{2}}^{t + \frac{\tau}{2}} \int_{\omega - \frac{\Delta\omega}{2}}^{\omega + \frac{\Delta\omega}{2}} \int_{\omega - \frac{\Delta\omega}{2}}^{\omega + \frac{\Delta\omega}{2}} \frac{\mathrm{d}t'}{\tau} \frac{\mathrm{d}\omega'}{\Delta\omega} \frac{\mathrm{d}\omega''}{\Delta\omega} \mathcal{A}_i(\omega') \mathcal{A}_j^*(\omega'') e^{i[\phi_i(\omega') - \phi_j(\omega'')]} e^{i(\omega' - \omega'')t'} \tag{5.7}$$

It is easy to see from the Cauchy-Schwarz inequality and equation (5.4) that

$$I^2 \geq Q^2 + U^2 + V^2. \tag{5.8}$$

5.1.3 Polarization of the superposition of independent waves

Now let's consider a superposition of different waves,

$$\mathbf{E} = \sum_k \mathbf{E}^{(k)}, \tag{5.9}$$

each of which can be decomposed as

$$\mathbf{E}^{(k)} = \mathcal{E}^{1,(k)}\hat{\mathbf{e}}_{1,(k)} + \mathcal{E}^{2,(k)}\hat{\mathbf{e}}_{2,(k)}. \tag{5.10}$$

These different waves are assumed to be "independent", which means that there are no permanent phase relations between these waves so that

$$\langle \mathcal{E}^{i,(k)}\mathcal{E}^{j,(l)} \rangle = 0, \quad \text{if } k \neq l. \tag{5.11}$$

Then the correlations in the definition of the Stokes parameters (Equation (5.4)) are given by

$$\langle \mathcal{E}^{i}\mathcal{E}^{j*} \rangle = \sum_k \sum_l \langle \mathcal{E}^{i,(k)}\mathcal{E}^{j,(l)*} \rangle = \sum_k \langle \mathcal{E}^{i,(k)}\mathcal{E}^{j,(k)*} \rangle \tag{5.12}$$

It proves the *superposition principle* of the Stokes parameters that

$$I = \sum I^{(k)}, \qquad Q = \sum Q^{(k)}, \qquad U = \sum U^{(k)}, \qquad V = \sum V^{(k)}. \tag{5.13}$$

The additivity also implies that an arbitrary set of Stokes parameters can be decomposed into a completely elliptically polarized wave of intensity $I_{pol} = \sqrt{Q^2 + U^2 + V^2}$ and a completely unpolarized wave of intensity $I - I_{pol}$. An extended concept is the *degree of polarization*, which defines the ratio of the polarized intensity to the total intensity

$$\Pi \equiv \frac{I_{pol}}{I}. \tag{5.14}$$

5.1.4 Polarization under rotation

We have just discussed two equivalent ways (Equations (5.2) and (5.3)) of characterizing the polarization of monochromatic waves defined by equation (5.1), but these descriptions are based on a chosen basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ of the polarization. Usually, the only constraint on the polarization bases is that they must be perpendicular to $\mathbf{k}$, although one may also require that $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ and $\mathbf{k}$ to follow a particular chirality, such as the right-hand convention. Therefore, our description of polarization usually has a rotational degree of freedom about $\mathbf{k}$. Since our discussion will involve the representation of the same polarization state in different bases, we need to understand how the polarization description based on the Stokes parameters changes with the basis.

We will use the rotation matrix $\mathcal{R}(\psi)$ to express the transformation of the polarization basis so that

$$\begin{pmatrix} \hat{\mathbf{e}}'_1 \\ \hat{\mathbf{e}}'_2 \end{pmatrix} = \mathcal{R}(\psi) \begin{pmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{pmatrix} = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{pmatrix}, \tag{5.15}$$

then $\mathcal{P}$ could be rewritten with the new basis as

$$\begin{aligned}
\mathcal{P} &\equiv \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \left[\sigma^I_{ij} I + \sigma^Q_{ij} Q + \sigma^U_{ij} U + \sigma^V_{ij} V\right] \\
&= (\mathcal{R}^T)_{ia}(\mathcal{R}^T)_{jb}\hat{\mathbf{e}}'_a \otimes \hat{\mathbf{e}}'_b \left[\sigma^I_{ij} I + \sigma^Q_{ij} Q + \sigma^U_{ij} U + \sigma^V_{ij} V\right] \\
&= \hat{\mathbf{e}}'_a \otimes \hat{\mathbf{e}}'_b \left[\sigma^I_{ab} I + \sigma^Q_{ab} \left(\cos(2\psi)Q - \sin(2\psi)U\right) + \sigma^U_{ab} \left(\sin(2\psi)Q + \cos(2\psi)U\right) + \sigma^V_{ab} V\right] \\
&\equiv \hat{\mathbf{e}}'_a \otimes \hat{\mathbf{e}}'_b \left[\sigma^I_{ab} I' + \sigma^Q_{ab} Q' + \sigma^U_{ab} U' + \sigma^V_{ab} V'\right]
\end{aligned} \tag{5.16}$$

Thus, after the rotation of the basis vectors, Stokes I and V remain unchanged, while Stokes Q and U are transformed such that

$$\begin{aligned}
Q' &= \cos(2\psi)Q - \sin(2\psi)U \\
U' &= \sin(2\psi)Q + \cos(2\psi)U
\end{aligned} \tag{5.17}$$

An equivalent description using complex exponents is also useful:

$$Q' + iU' = e^{2i\psi}(Q + iU), \tag{5.18}$$

or similarly,

$$Q' - iU' = e^{-2i\psi}(Q + iU). \tag{5.19}$$

5.2 Polarization field of radio sky

In this section, we discuss the representations of polarization for the entire radio sky, which refers to the radiation field observed by the telescope. Since the size of a telescope receiver is negligible compared to its distances from astrophysical sources, the radiation field can be approximated as a time-varying vector field on a 2-sphere manifold.[3]

To describe the polarization of the electric vector field, we define the polarization tensor field on the sphere. This requires us to specify the choice of polarization basis vectors at each point on the sphere. We use the natural choice of local spherical coordinate basis vectors, with $\hat{\mathbf{e}}_1 = \hat{\theta}$ and $\hat{\mathbf{e}}_2 = \hat{\phi}$ for each point. The description of the polarization is then given by a polarization tensor field using the spherical coordinate basis vectors, expressed in terms of Stokes parameters as follows (See e.g. Smirnov (2011)):

$$\mathcal{P}(\hat{\mathbf{n}}) = \hat{\mathbf{e}}_i(\hat{\mathbf{n}}) \otimes \hat{\mathbf{e}}_j(\hat{\mathbf{n}}) \left[\sigma^I_{ij} I(\hat{\mathbf{n}}) + \sigma^Q_{ij} Q(\hat{\mathbf{n}}) + \sigma^U_{ij} U(\hat{\mathbf{n}}) + \sigma^V_{ij} V(\hat{\mathbf{n}})\right] \tag{5.20}$$

Thus, given the coordinate basis vectors and the integration time, a complete descrip-

[3]By "manifold", the author means that the linear space of the vector is the local tangent space of the sphere.

tion of the polarization of the radiation field corresponds to four real sky maps, each corresponding to a Stokes parameter.

Going one step further, and continuing the previous discussion on the rotation of the polarization basis vector, we can decompose the polarization tensor field, which has four degrees of freedom, into two real scalar fields, Stokes I and V[4], and a complex field with spin ± 2, $Q \pm iU$.

5.2.1 Multipole expansion of the polarized radio sky

In this section, we look into the widespread use of multipole expansion in intensity mapping. It proves invaluable not only for analyses related to the angular power spectrum but also for interpolation and coordinate transformation tasks.

Spherical harmonic expansion of the scalar fields

Any squre-integrable scalar field $T(\hat{\mathbf{n}})$ on the unit sphere S^2 can be expanded with the Laplace spherical harmonics $Y_{lm}(\theta, \phi)$:

$$T(\theta, \phi) = \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \phi), \tag{5.21}$$

where

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta) e^{im\phi} \tag{5.22}$$

and

$$a_{lm} = \int_0^{2\pi} \mathrm{d}\phi \int_{-1}^{1} \mathrm{d}\,(\cos\theta)\ T(\theta, \phi) Y_{lm}^*(\theta, \phi). \tag{5.23}$$

Here $P_l^m(\cos\theta)$ are the associated Legendre polynomials with the Condon-Shortley phase.

When T is a real scalar field, such as the temperature field, the expansion coefficients a_{lm} satisfy the *reality condition*,

$$a_{lm}^* = (-1)^m a_{l,-m} \tag{5.24}$$

which is a direct consequence of

$$Y_{lm}(\theta, \phi) = (-1)^m Y_{l,-m}^*(\theta, \phi). \tag{5.25}$$

Below we summarize some other useful properties of spherical harmonics, which will be used later in our discussion:

[4]We do not consider parity inversion here. Otherwise, the parity odd Stokes V sky should be considered as a pseudo-scalar field.

- The completeness condition:

$$\sum_{l=0}^{+\infty}\sum_{m=-l}^{+l} Y^*_{lm}(\theta,\phi)Y_{lm}(\theta',\phi') = \delta(\phi-\phi')\delta(\cos\theta-\cos\theta') \tag{5.26}$$

- The orthogonality relation:

$$\int_0^{2\pi} \mathrm{d}\phi \int_{-1}^{1} \mathrm{d}\left(\cos\theta\right)\ Y^*_{l'm'}(\theta,\phi)Y_{lm}(\theta,\phi) = \delta_{l'l}\delta_{m'm}. \tag{5.27}$$

- Transformation under rotation $R(\alpha,\beta,\gamma)$, where α,β,γ are Euler angles in the $z-y-z$ right hand convention:

$$Y_{lm}(R\hat{\mathbf{n}}) = \sum_{m'=-l}^{+l} [D^{(l)}_{mm'}(R)]^* Y_{lm'}(\hat{\mathbf{n}}) \tag{5.28}$$

  where $D^{(l)}_{mm'}$ is the Wigner D-matrix.

  One type of rotation of particular interest to ground observers is the ϕ rotation of the azimuth.[5] The rotation matrix then becomes $R(\phi,0,0)$ and the Wigner D-matrix degenerates to

$$D^{(l)}_{mm'}(\phi,0,0) = e^{-im'\phi}\delta_{m'm} \tag{5.29}$$

  and the spherical harmonics transform as

$$Y_{lm}(R\hat{\mathbf{n}}) = Y_{lm}(\hat{\mathbf{n}})e^{im\phi}. \tag{5.30}$$

  This compelling property has inspired the use of m-mode analysis in several drift-scan surveys within this domain (Shaw et al., 2014).

Expansion of the Stokes Q and U skies

There are also sets of basis functions that can be used to expand spin-s functions. These are the so-called spin-s spherical harmonics, ${}_sY_{lm}(\theta,\phi)$. In particular, the spin-weighted spherical harmonics of $s=0$ are simply the ordinary spherical harmonics, ${}_0Y_{lm}(\theta,\phi) = Y_{lm}(\theta,\phi)$.

Given a field η with spin weight s,[6] one can construct a spin-$(s+1)$ field $\eth\eta$ where $\eth$ is a covariant derivative operator in the sphere:

$$\eth\eta \equiv -(\sin\theta)^s \left\{\frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right\}[(\sin\theta)^{-s}\eta]. \tag{5.31}$$

[5] Say the sky srift caused by the Earth rotation.
[6] That means $\eta' = e^{is\phi}\eta$ under a conventional rotation with angle ϕ.

Similarly, there is an operator $\bar{\eth}$ that can decrease the spin weight of the function acted on by 1 (Torres del Castillo, 2007):

$$\bar{\eth}\eta \equiv -(\sin\theta)^{-s}\left\{\frac{\partial}{\partial\theta} - \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right\}[(\sin\theta)^s\eta]. \tag{5.32}$$

The spin-weighted spherical harmonics ${}_sY_{lm}(\theta,\phi)$ are then related to the ordinary spherical harmonics as

$${}_sY_{lm} = \begin{cases} \sqrt{\frac{(l-s)!}{(l+s)!}}\,\eth^s Y_{lm}, & 0 \le s \le l; \\ \sqrt{\frac{(l+s)!}{(l-s)!}}\,(-1)^s\,\bar{\eth}^{-s}Y_{lm}, & -l \le s \le 0; \\ 0, & l < |s|. \end{cases} \tag{5.33}$$

Just like ordinary spherical harmonics, the spin-s spherical harmonics also satisfy the completeness and orthogonality relations by simply replacing Y_{lm} in equations (5.26) and (5.27) with ${}_sY_{lm}$. The formalism for rotation transformations is somewhat different. Under a rotation of the spherical coordinate system $\hat{\mathbf{n}} \to R\hat{\mathbf{n}}$, the spin-$s$ spherical harmonics become

$${}_sY_{lm}(R\hat{\mathbf{n}}) = e^{is\psi(R,\hat{\mathbf{n}})}\sum_{m'}[D^{(l)}_{mm'}(R)]^*\,{}_sY_{lm'}(\hat{\mathbf{n}}) \tag{5.34}$$

where ψ is the local rotation angle from the old local coordinate basis vectors to the new ones. It's easy to see that the spin-weighted harmonics satisfy

$$[{}_sY_{lm}(\theta,\phi)]^* = (-1)^{s+m}\,{}_{-s}Y_{l,-m}(\theta,\phi). \tag{5.35}$$

To expand the Stokes Q and U sky, we define the complex polarized intensity fields of the spin weights of ± 2:

$$I_p^{\pm}(\hat{\mathbf{n}}) = (Q \pm iU)(\hat{\mathbf{n}}),$$

which can be expanded as

$$I_p^{\pm}(\hat{\mathbf{n}}) = \sum_{lm} a_{\pm 2,lm}\;{}_{\pm 2}Y_{lm}(\hat{\mathbf{n}}) \tag{5.36}$$

where the expansion coefficients $a_{\pm 2,lm}$ are

$$a_{\pm 2,lm} = \int_0^{2\pi}\mathrm{d}\phi\int_{-1}^{1}\mathrm{d}\cos\theta\;I_p^{\pm}(\theta,\phi)\;{}_{\pm 2}Y^*_{lm}(\theta,\phi). \tag{5.37}$$

Substituting equation (5.35) into the conjugate form of equation (5.37) we get

$$a^*_{-2,lm} = (-1)^m a_{2,l,-m} \tag{5.38}$$

Since the basis functions ${}_{\pm2}Y_{lm}$ themselves contain the properties of the spin ±2, the expansion coefficients have the advantage of being rotationally invariant.

5.2.2 Scalar representations with fixed polarization basis

In the previous discussion we expressed the polarization tensor field as two real scalar fields plus a spin-weighted complex field. However, after a coordinate transformation, we still need to locally correct for the extra phase of the spin-weighted field caused by the rotation of the polarization basis. This will complicate the actual analysis, e.g. if one wants to simulate time series measurements in an antenna static frame, which is preferred by a beam simulation program. Therefore, we will want to make a polarization convention that is independent of coordinate transformations.

A natural idea is to obtain an equivalent set of coefficients $\{a_{E/B,lm}\}$ via a one-to-one and onto mapping of $\{a_{\pm2,lm}\}$. These coefficients satisfy the reality condition and thus define the equivalent real scalar fields. This idea is not difficult to implement: observing equation (5.35), one can always find such a mapping that

$$a_{E,lm} = -(a_{2,lm} + a_{-2,lm})/2, \qquad a_{B,lm} = i(a_{2,lm} - a_{-2,lm})/2, \tag{5.39}$$

and we get two real scalar fields

$$E(\hat{\mathbf{n}}) = \sum_{lm} a_{E,lm}\, Y_{lm}(\hat{\mathbf{n}}), \qquad B(\hat{\mathbf{n}}) = \sum_{lm} a_{B,lm}\, Y_{lm}(\hat{\mathbf{n}}). \tag{5.40}$$

Here we briefly summarize this section to better understand the point that the E and B fields are coordinate invariant descriptions of the Stokes Q and U fields. In a coordinate system, O, one can conventionally define the Stokes Q and U fields using the coordinate basis vectors and obtain the corresponding spin-weighted spherical harmonic coefficients, $a_{\pm2,lm}$. We can obtain a linearly transformed set of these coefficients, $a_{E/B,lm}$, which satisfy the reality condition, i.e., equation (5.24), and which equivalently describe the two spherical scalar fields, the E and B fields. If we consider a new coordinate system, O', then the fields E and B are invariant under coordinate transformations, simply because they are defined associated with the coordinate system O. In other words, changing the coordinates does not change the polarization basis. Thus, from this point of view, the fields E and B are not scalar fields in the strict sense of the word, they are just descriptions with a frozen convention of the polarization basis vectors.

In section 18.1, we also present a scalar representation in the flat-sky approximation.

Chapter 6

Antenna measurement

In the previous chapter, we introduced and discussed the extensive properties of the polarization field of a general radio sky. In this chapter, we formally present the theory of polarization measurements. The objective of this section is straightforward: to equip students, especially those working with antenna far fields such as CST simulations, with the knowledge to represent general ground-based observations using the antenna far field.

6.1 Antenna voltage signal

The reception pattern of an antenna is usually characterized by the far field that results when the antenna is used as a transmitter. This is a consequence of the reciprocity principle. Both the antenna far-field pattern and the instantaneous incoming signal can be expressed as two-dimensional vector fields on a 2-sphere. Without introducing a coordinate system, we can generally express such a vector field as a physical entity:

$$\boldsymbol{V}(\hat{n}) = V^1(\hat{n})\,\vec{e}_1(\hat{n}) + V^2(\hat{n})\,\vec{e}_2(\hat{n}) \equiv V^\mu\,\vec{e}_\mu \tag{6.1}$$

where $\hat{n}$ denotes the direction[1] on the sphere, $\mu = 1, 2$, and $\hat{\epsilon}_\mu(\hat{n})$ are basis vectors of the tangent space at $\hat{n}$. We adopt the Einstein summation convention so that repeated Greek indices are summed.

The far-field pattern, denoted by $\boldsymbol{E}(\nu)$, is frequency dependent and usually assumed to be constant over observation time. However, the sky electric field, $\boldsymbol{\mathcal{E}}(\nu_i, t_j)$, is a time-varying vector field. We used discrete coordinates ν_i, t_j for two reasons: First, according to the uncertainty principle, the radiation electric field cannot be arbitrarily compact in both time and frequency, so one should bin the coordinates up to an uncertainty. Second, in frequency observations based on sampling the signal in the

[1]Note that $\hat{n}$ is a physical entity known as direction, and we have not introduced any coordinates to represent $\hat{n}$ until equation (6.7).

time domain[2], the component with frequency ν is extracted from the signal over a time period $(t-\tau/2, t+\tau/2)$, where $\tau \gg \nu^{-1}$. Thus, the sky electric field $\boldsymbol{\mathcal{E}}(\nu_i, t_j)$ can be understood as the ν_i component of the time segment of $\boldsymbol{\mathcal{E}}(t)$ centered at t_j.

The voltage signal of frequency ν_i at the receiver, induced by $\boldsymbol{\mathcal{E}}$ and observed at time t_i, can then be expressed as the spherical integral of the *inner product* of $\boldsymbol{E}(\nu_i)$ and $\boldsymbol{\mathcal{E}}(\nu_i, t_j)$ (Evans, 1990):

$$\mathcal{V}(\nu_i, t_j) \equiv \iint \mathrm{d}^2\hat{n}\, \boldsymbol{E}(\hat{n}, \nu_i) \cdot \boldsymbol{\mathcal{E}}(\hat{n}, \nu_i, t_j). \tag{6.2}$$

Strictly speaking, one should also introduce a phase term in equation (6.2) to account for the relative phase of the receiver's far field, since the far field is usually defined relative to a phase constant sphere around the antenna. However, this phase correction is actually trivial for data analysis. For autocorrelation of a receiver, the phase information is removed. For cross-correlation, there is no need for phase correction either, since the far fields of all receivers in the telescope or array usually follow the same phase convention.

Ideally, equation (6.2) could simply be rewritten as $\mathcal{V} = \iint \mathrm{d}^2\hat{n}\, E^1\mathcal{E}^1 + E^2\mathcal{E}^2$ if $\boldsymbol{E}$ and $\boldsymbol{\mathcal{E}}$ are locally expressed with the same orthonormal linear basis. In practical analysis, however, one tends to use different coordinate systems as well as linear bases for $\boldsymbol{E}$ and $\boldsymbol{\mathcal{E}}$ to maximize the ease of data analysis. For obvious reasons, the preferred coordinate system is the antenna static coordinate system (denoted by O) for the antenna far field $\boldsymbol{E}$ and the celestial coordinate system (denoted by O') for the sky signal $\boldsymbol{\mathcal{E}}$.

Using the coordinate basis as the linear basis of the vector space[3], a general vector field on the sphere can be expressed in either O or O':

$$\boldsymbol{V} = V^\alpha \vec{e}_\alpha = V^{\beta'} \vec{e}_{\beta'} = \Lambda^{\beta'}{}_{\alpha} V^\alpha \vec{e}_{\beta'} = \Lambda^{\alpha}{}_{\beta'} V^{\beta'} \vec{e}_\alpha \tag{6.3}$$

Here and below, the unprimed Greek indices correspond to the antenna coordinate system O, while the primed Greek indices correspond to the sky coordinate system O'. $\Lambda^{\beta'}{}_{\alpha}$ is the *Jacobian* matrix of the coordinate mapping

$$f : O \to O'. \tag{6.4}$$

Its inverse matrix, $\Lambda^{\alpha}{}_{\beta'}$, is the Jacobian of f^{-1}. More often than not, we want to use *normalized* coordinate bases $\{\hat{e}_\alpha\}$ and $\{\hat{e}_{\beta'}\}$, in terms of which a general vector $\boldsymbol{V}$ reads

[2]In standard Nyquist sampling, τ gives the inverse of the frequency resolution and the sampling frequency ν_{sample} satisfies $\nu_{\mathrm{sample}} > 2\nu_{\mathrm{max}}$. Before the sampling, a low-pass filter (LPF) is usually needed to avoid aliasing.

[3]While there are other widely used conventions introduced by Ludwig 1973, we have used the coordinate basis vectors as references for polarization and cross polarization.

as follows

$$\begin{aligned}\boldsymbol{V} = V^{\alpha}\hat{e}_{\alpha} &= \frac{|\vec{e}_{\alpha}|}{|\vec{e}_{\beta'}|}\Lambda^{\alpha}{}_{\beta'}V^{\beta'}\hat{e}_{\alpha} \equiv \tilde{\Lambda}^{\alpha}{}_{\beta'}V^{\beta'}\hat{e}_{\alpha} \\ &= V^{\beta'}\hat{e}_{\beta'} = \frac{|\vec{e}_{\beta'}|}{|\vec{e}_{\alpha}|}\Lambda^{\beta'}{}_{\alpha}V^{\alpha}\hat{e}_{\beta'} \equiv \tilde{\Lambda}^{\beta'}{}_{\alpha}V^{\alpha}\hat{e}_{\beta'}\end{aligned} \tag{6.5}$$

where we have defined the *orthogonalized* Jacobian matrix, $\tilde{\Lambda}$.

Expressing $\boldsymbol{E}$ and $\boldsymbol{\mathcal{E}}$ in their preferred coordinate systems, we can rewrite equation (6.2) as

$$\mathcal{V} = \iint \mathrm{d}^2\hat{n}\, E^{\alpha}\mathcal{E}^{\beta'}\hat{e}_{\alpha}\cdot\hat{e}_{\beta'} = \iint \mathrm{d}^2\hat{n}\,\tilde{\Lambda}^{\gamma}{}_{\beta'}E^{\alpha}\mathcal{E}^{\beta'}\delta_{\alpha\gamma} = \iint \mathrm{d}^2\hat{n}\,\tilde{\Lambda}^{\alpha}{}_{\beta'}E_{\alpha}\mathcal{E}^{\beta'} \tag{6.6}$$

where in the second equality we have applied the orthonormality of the coordinate basis. Otherwise, we have to use a non-trivial metric instead of $\delta_{\alpha\gamma}$.

Equation (6.6) only reads the field components. We also want to explicitly write the fields in their respective preferred coordinates. Similarly, we denote the spherical coordinates of a physical point $\hat{n}$ on the sphere in O and O' as unprimed n^{α} and primed $n^{\beta'}$, respectively. Then equation (6.6) can be rewritten as

$$\begin{aligned}\mathcal{V}(\nu_i, t_j) &= \iint \mathrm{d}^2\hat{n}\,\tilde{\Lambda}^{\alpha}{}_{\beta'}(\hat{n})E_{\alpha}(\hat{n},\nu_i)\mathcal{E}^{\beta'}(\hat{n},\nu_i,t_j) \\ &= \iint \mathrm{d}^2\hat{n}\,\frac{|\vec{e}_{\alpha}|}{|\vec{e}_{\beta'}|}\frac{\partial n^{\alpha}}{\partial n^{\beta'}}E_{\alpha}(n^{\alpha})\mathcal{E}^{\beta'}(n^{\beta'},\nu_i,t_j),\end{aligned} \tag{6.7}$$

where $n^{\beta'} = f(n^{\alpha})$. A useful equivalent to the above equations is to express the integrand as a linear algebra multiplication:

$$\mathcal{V} = \iint \mathrm{d}^2\hat{n}\,\mathbf{E}^T\tilde{\mathbf{\Lambda}}\mathcal{E}, \tag{6.8}$$

where $\mathbf{E}$ and $\mathcal{E}$ are two-dimensional column vectors, and $\tilde{\mathbf{\Lambda}}$ is a 2-by-2 orthogonal matrix. This expression has the merit of keeping us away from the Greek indices.

In the second identity of the equation (6.7), the integral is written in antenna coordinates, while the sky field is expressed in celestial coordinates, which is a friendly arrangement for data analysts. In the next section we will provide details on the coordinate mapping f and the field transformation $\tilde{\Lambda}^{\alpha}{}_{\beta'}$.

6.2 Mapping the sky on the ground: coordinate conversions

As elaborated in Section 6.1, the instantaneous antenna reception can be expressed as an integral over the celestial sphere. This reception process involves three coordi-

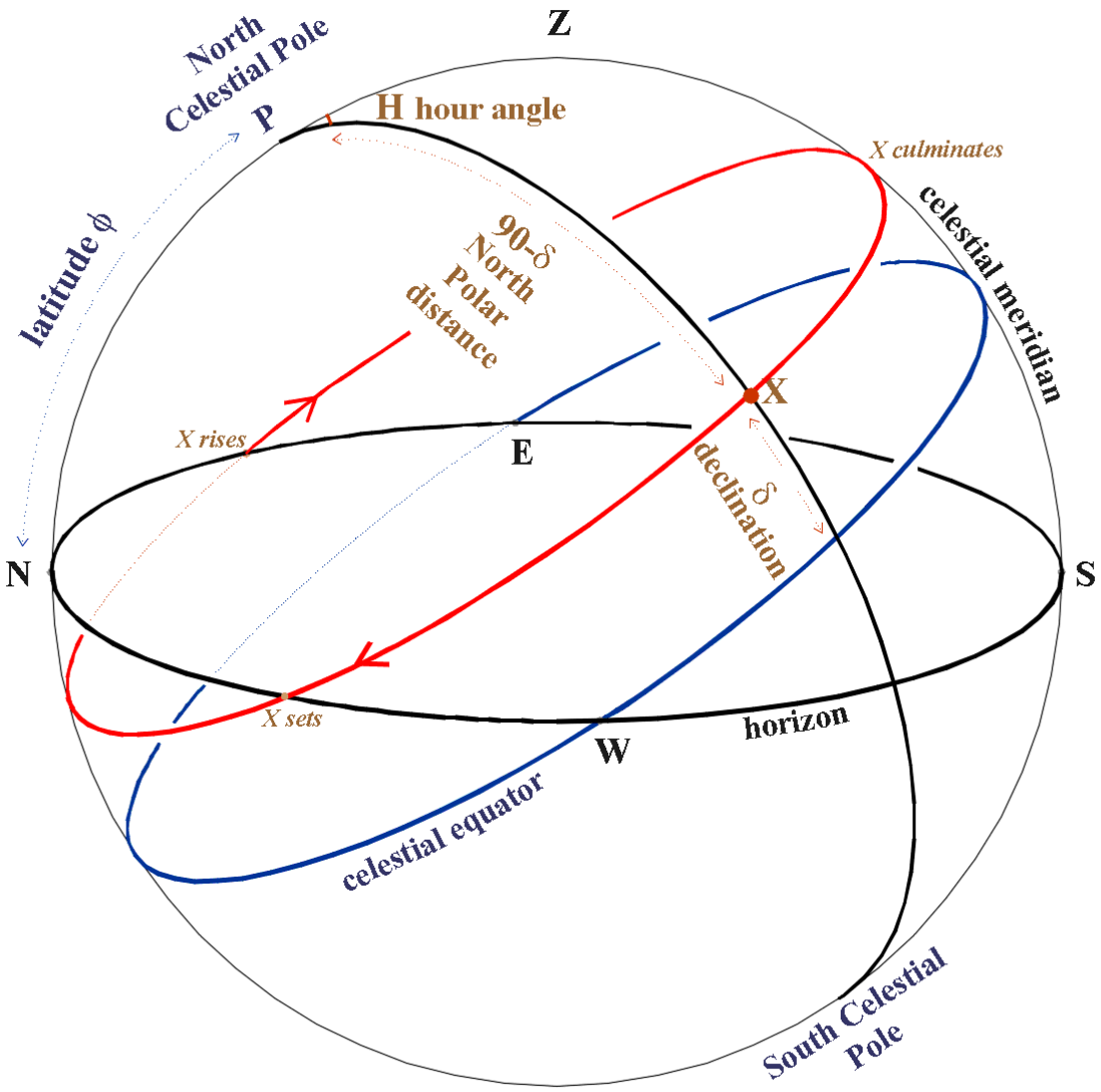

Figure 6.1: Diagram of the Equatorial and Horizontal coordinate systems.

nate systems, one fixed to the telescope, another fixed to the local site where antenna is deployed, and yet another fixed to the celestial sphere. In this section, we show the transformation matrices between the various coordinate systems and how the antenna signal is most conveniently represented in terms of equatorial coordinates. This is standard material, but here we indicate explicitly our conventions for describing polarization.

6.2.1 Convention

The antenna far-field pattern is usually expressed in the coordinate system where the antenna is static, which is terminated as the *antenna coordinate system*. However, the pointing direction of the telescope, or the beam direction, is static (and thus easily expressed) relative to the *horizontal coordinate system*. Furthermore, the sky catalogs

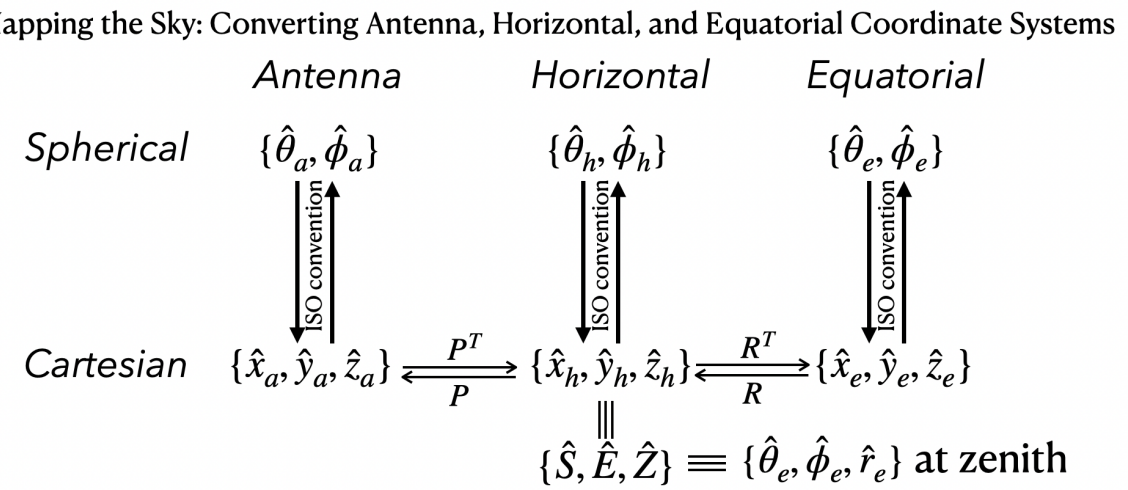

Figure 6.2: Conversions of coordinate basis vectors between antenna, horizontal and equatorial coordinate systems. The equatorial spherical coordinates, $\{\theta_e, \phi_e\}$, are basically the same thing as the $\{\mathrm{RA}, \mathrm{DEC}\}$ coordinates but with $\theta_e = \frac{\pi}{2} - \mathrm{DEC}$ and $\phi = \mathrm{RA}$, where all the angle objects are in radians. The ISO convention implies that the polar axis ($\theta = 0$) aligns with the $\hat{z}$ direction, $\hat{x}$ orients to the direction of $\theta = \pi/2$ and $\phi = 0$, and the y-axis can be given by the right hand convention. The antenna Cartesian coordinates, (x_a, y_a, z_a), and the equatorial Cartesian coordinates, (x_e, y_e, z_e), are naturally defined with the spherical coordinates under ISO convention. However, the horizontal (or local) Cartesian coordinates, (x_h, y_h, z_h), are defined using the Cardinal directions and the zenith of the antenna: it shares the same origin as the antenna Cartesian system; $\hat{z}_h \equiv \hat{\mathbf{Z}}$ is the direction from the antenna to its zenith; $\hat{x}_h \equiv \hat{\mathbf{S}}$ is the direction of the true South of the antenna; $\hat{y}_h \equiv \hat{\mathbf{E}}$ is the direction of the true East of the antenna. The rotation matrix R is defined in equation (6.11).

are constant in celestial coordinate systems fixed to stars, so one might want to use for example the *equatorial coordinate system* to map the sky. When it comes to a cosmological analysis, one might also prefer to formulate the observations using celestial coordinates. For each coordinate system, we define both the spherical and the Cartesian coordinates. By introducing Cartesian coordinates, one can the conveniently characterize the relationship between coordinate systems in terms of linear transformations. The conversions between these coordinate systems are illustrated in Figure 6.2.

Our goal is to formulate the observation with each component given in its favoured coordinates. To be exact, the problem can be defined within a rather practical observation scenario:

- The time and location of a measurement is given by the local sidereal time (LST, in seconds) and the latitude (lat, in radians) of the antenna.
- The antenna far-field pattern is simulated or measured in antenna spherical coordinates, θ_a and ϕ_a.
- The antenna alignment or orientation is specified by expressing the antenna axes,

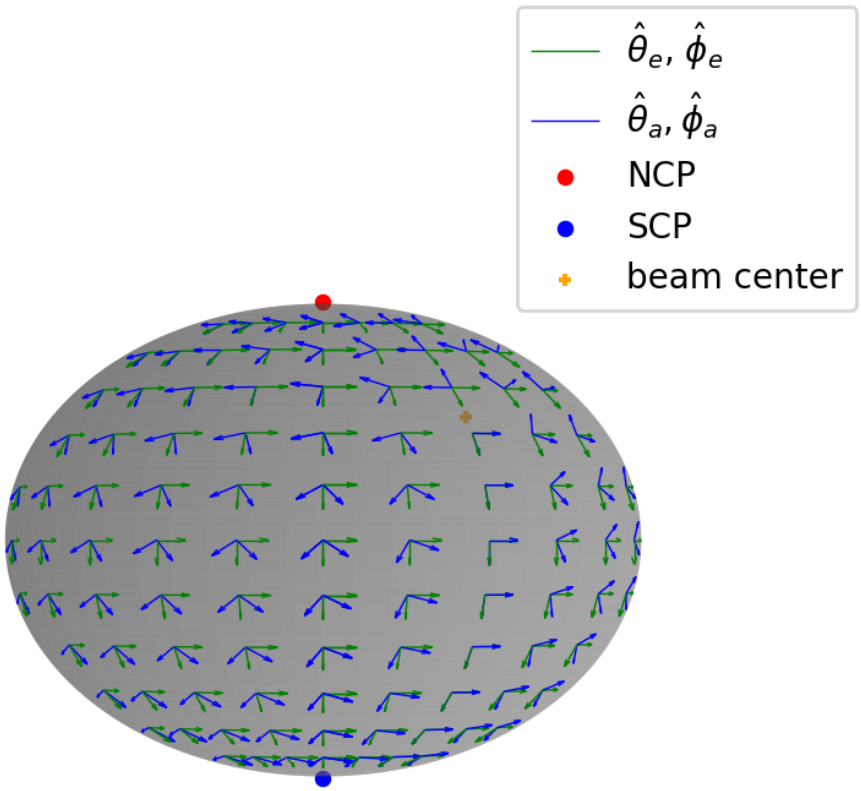

Figure 6.3: Diagram illustrating the spherical coordinate basis vectors, serving as polarization reference vectors, within both the antenna and equatorial coordinate systems.

$\{\hat{x}_a, \hat{y}_a, \hat{z}_a\}$, in terms of horizontal coordinate basis, $\{\hat{x}_h, \hat{y}_h, \hat{z}_h\}$. One can define the pointing matrix, P, such that

$$\begin{pmatrix} \hat{x}_a \\ \hat{y}_a \\ \hat{z}_a \end{pmatrix} = P \begin{pmatrix} \hat{\mathbf{S}} \\ \hat{\mathbf{E}} \\ \hat{\mathbf{Z}} \end{pmatrix} \tag{6.9}$$

where $\hat{\mathbf{S}} = \hat{x}_h$, $\hat{\mathbf{E}} = \hat{y}_h$ and $\hat{\mathbf{Z}} = \hat{y}_h$ align with the South, the East, and the Zenith of the antenna respectively, as defined in the caption of Figure 6.2.

- The task is to formulate the antenna reception in equatorial spherical coordinates, (θ_e, ϕ_e).

6.2.2 Coordinate mapping and field transformation

In this section, we shall formulate the transformations between the antenna, horizontal and equatorial systems. The antenna and horizontal systems are both centred at the antenna receiver, while the origin of the equatorial system is separated from the antenna by a distance around the radius of the Earth. Since what we observe are astrophysical sources, this difference in origin is considered negligible so that we shall assume the three systems have the same origin. Therefore, we can simply formulate the coordinate

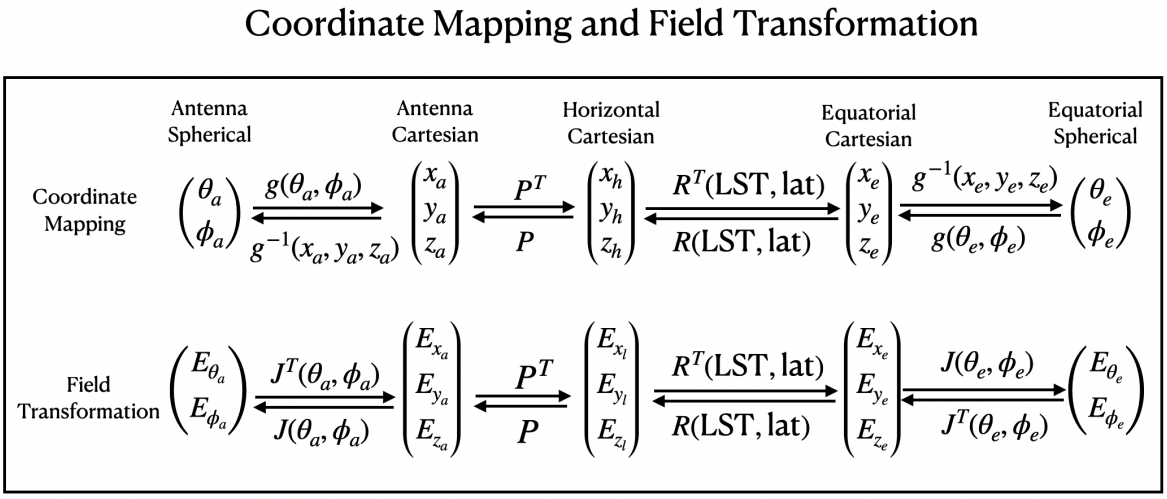

Figure 6.4: Illustrating diagram for coordinate mapping and field transformation between antenna, horizontal and equatorial coordinate systems.

transformations in terms of rotation matrices.

The connection between the antenna and horizontal systems has been given by the pointing matrix, P, as illustrated in equation (6.9). Similarly, one can get the transformation between horizontal and equatorial Cartesian coordinates by realizing that the basis of the horizontal spherical coordinate system just constitutes the spherical coordinate basis of the zenith in the equatorial system,

$$\begin{pmatrix}\hat{\mathbf{S}}\\ \hat{\mathbf{E}}\\ \hat{\mathbf{Z}}\end{pmatrix} = \begin{pmatrix}\hat{\theta}_e\\ \hat{\phi}_e\\ \hat{r}_e\end{pmatrix}_{\text{zenith}} = R\begin{pmatrix}\hat{x}_e\\ \hat{y}_e\\ \hat{z}_e\end{pmatrix} \tag{6.10}$$

where

$$R(\mathrm{lat},\mathrm{LST}) = \begin{pmatrix}\cos\theta_z\cos\phi_z & \cos\theta_z\sin\phi_z & -\sin\theta_z\\ -\sin\phi_z & \cos\phi_z & 0\\ \sin\theta_z\cos\phi_z & \sin\theta_z\sin\phi_z & \cos\theta_z\end{pmatrix} \tag{6.11}$$

and θ_z and ϕ_z are the equatorial spherical coordinates of the zenith (local to the antenna), so they are functions of the LST (in seconds) and the latitude (in radians) of the antenna:

$$\theta_z = \frac{\pi}{2} - \mathrm{lat}, \quad \phi_z = \frac{\mathrm{LST}}{3600}\times 15\times\frac{\pi}{180}. \tag{6.12}$$

With the coordinate basis transformations between the antenna, horizontal, and equatorial systems, we can easily map coordinates and transform field components between antenna and equatorial spherical coordinates, which is illustrated in Figure 6.4. The coordinate mapping at given lat and LST, f, is given by a composition:

$$f : (\theta_a, \phi_a; \text{lat}, \text{LST}) \to (\theta_e, \phi_e), \qquad \boxed{f \equiv g^{-1} \circ R^{-1} \circ P^{-1} \circ g}, \tag{6.13}$$

$$f^{-1} : (\theta_e, \phi_e; \text{lat}, \text{LST}) \to (\theta_a, \phi_a), \qquad \boxed{f^{-1} \equiv g^{-1} \circ P \circ R \circ g}, \tag{6.14}$$

where $g : (\theta, \phi) \xrightarrow{r=1} (x, y, z)^T$ maps spherical coordinates to Cartesian coordinates and g^{-1} denotes the inverse mapping. In matrix form, we have $R^{-1} = R^T$ and $P^{-1} = P^T$.

The transformation of field components is then specified by the orthogonalized Jocobian of the coordinate mapping,

$$\begin{pmatrix} E_{r_e} \\ E_{\theta_e} \\ E_{\phi_e} \end{pmatrix} = \mathcal{J}(\theta_e, \phi_e) R^T P^T \mathcal{J}^T(\theta_a, \phi_a) \begin{pmatrix} E_{r_a} \\ E_{\theta_a} \\ E_{\phi_a} \end{pmatrix} \tag{6.15}$$

where $(\theta_a, \phi_a) = f^{-1}(\theta_e, \phi_e)$ and $\mathcal{J}$ is defined as

$$\mathcal{J}(\theta, \phi) = \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{pmatrix}. \tag{6.16}$$

Since the radial field components are actually zero as we are considering tangent vector fields on the sphere, we don't really need the radial part of rotation matrix. Therefore, we can discard the first row of $\mathcal{J}$, and denote the residual 2×3 matrix as J:

$$J(\theta, \phi) = \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{pmatrix}. \tag{6.17}$$

Then the field transformation operators can be expressed as

$$\text{for } f : O \to O', \qquad \boxed{\tilde{\Lambda}^{\alpha'}{}_{\alpha} = J(\theta_e, \phi_e) R^T P^T J^T(\theta_a, \phi_a)}, \tag{6.18}$$

$$\text{for } f^{-1} : O' \to O, \qquad \boxed{\tilde{\Lambda}^{\alpha}{}_{\alpha'} = J(\theta_a, \phi_a) P R J^T(\theta_e, \phi_e)}, \tag{6.19}$$

where $(\theta_e, \phi_e) = f(\theta_a, \phi_a)$.

6.2.3 A closer look

In the previous section, we obtained a specific form of coordinate mapping and vector field transformation between the sky coordinate system and the ground-based antenna coordinate system. Theoretically, one could compute coordinate mappings for each position of the sphere to relate the far-field degrees of freedom of the antenna to the vector-field degrees of freedom of the sky. However, this is tedious and seems unnecessary considering the fact that as long as we specify the spatial orientation of the telescope, all degrees of freedom are implicitly mapped onto the celestial sphere. This means that we can formalize the vector field transformations in a form that

depends only on the antenna orientation, which is jointly determined by the pointing matrix (P) and the zenith sky coordinates ((θ_z,ϕ_z) or equivalently (lat, LST)), as well as the antenna coordinates (θ_a, ϕ_a). In this section we develop such a formalism by making the transformation operator $\tilde{\mathbf{\Lambda}}$ independent of the sky coordinates (θ_e, ϕ_e).

Recall the coordinate transformation from the antenna system to the sky system, $f : O \to O'$, where the mapping is represented as a function composition and the transformation operator is

$$\boxed{f \equiv g^{-1} \circ R^{-1} \circ P^{-1} \circ g} \qquad \boxed{\tilde{\mathbf{\Lambda}} = J(\theta_e, \phi_e) R^T P^T J^T(\theta_a, \phi_a)}. \tag{6.20}$$

As we will see, we can decompose these orthogonal matrices using two basic rotation matrices, which makes it easy to distinguish the roles of the different variables. We define

$$\mathcal{R}_2(\Delta) \equiv \begin{pmatrix} \cos\Delta & 0 & -\sin\Delta \\ 0 & 1 & 0 \\ \sin\Delta & 0 & \cos\Delta \end{pmatrix}, \qquad \mathcal{R}_3(\Delta) \equiv \begin{pmatrix} \cos\Delta & -\sin\Delta & 0 \\ \sin\Delta & \cos\Delta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{6.21}$$

then the transformations, equations (6.11) and (6.17), can be rewritten as

$$R = \mathcal{R}_2(\theta_z)\mathcal{R}_3(-\phi_z), \qquad J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \mathcal{R}_2(\theta)\mathcal{R}_3(-\phi). \tag{6.22}$$

We can also decompose the spherical-to-cartesian coordinate mapping, g, as

$$g(\theta, \phi) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} = \mathcal{R}_3(\phi)\mathcal{R}_2(-\theta) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{6.23}$$

The function composition relates the antenna coordinate (θ_a, ϕ_a) and the sky coordinate (θ_e, ϕ_e) by $g(\theta_e, \phi_e) = R^T P^T g(\theta_a, \phi_a)$, which can be expanded as

$$\mathcal{R}_3(\phi_e)\mathcal{R}_2(-\theta_e) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{6.24}$$

where we have defined

$$\begin{aligned} A(\phi_a, \theta_a, \phi_z, \theta_z) &\equiv R^T P^T \mathcal{R}_3(\phi_a)\mathcal{R}_2(-\theta_a) \\ &= \mathcal{R}_3(\phi_z)\mathcal{R}_2(-\theta_z) P^T \mathcal{R}_3(\phi_a)\mathcal{R}_2(-\theta_a). \end{aligned} \tag{6.25}$$

$J(\theta_e, \phi_e)$ is the only term in $\tilde{\mathbf{\Lambda}}$ that depends on the sky coordinates. Using equations (6.23) and (6.24), we can deduce that $J(\theta_e, \phi_e)$ can be fully expressed in terms of A,

as shown below:

$$J(\theta_e,\phi_e)=\frac{1}{\sqrt{1-\lambda^2}}\left\{\begin{pmatrix}0&0&\lambda\\0&0&0\end{pmatrix}A^T+\begin{pmatrix}0&0&0\\0&0&1\end{pmatrix}A^T\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}-\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}\right\} \tag{6.26}$$

where $\lambda = A_{33}$. Substituting equation (6.26) into $\tilde{\mathbf{\Lambda}}$, we have

$$\tilde{\mathbf{\Lambda}}=\lambda'\left\{\begin{pmatrix}0&0&0\\0&0&1\end{pmatrix}A^T\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}A-\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}A\right\}\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix} \tag{6.27}$$

where $\lambda' = 1/\sqrt{1-A_{33}^2}$.

A useful observation is that $\tilde{\mathbf{\Lambda}}$ is independent of ϕ_z or LST. The ϕ_z-dependence in A is encapsulated in $\mathcal{R}_3(\phi_z)$, which is canceled in the first term of equation (6.27) as

$$\mathcal{R}_3^T(\phi_z)\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}\mathcal{R}_3(\phi_z)=\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}. \tag{6.28}$$

While in the second term, we have

$$\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}\mathcal{R}_3(\phi_z)=\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix} \tag{6.29}$$

so that $\mathcal{R}_3(\phi_z)$ is also canceled. Therefore, defining

$$\boxed{\tilde{A}\equiv\mathcal{R}_3(-\phi_z)A=\mathcal{R}_2(-\theta_z)P^T\mathcal{R}_3(\phi_a)\mathcal{R}_2(-\theta_a)}, \tag{6.30}$$

equation (6.27) can be rewritten as

$$\boxed{\tilde{\mathbf{\Lambda}}(\theta_a,\phi_a;\text{lat},P)=\lambda''\left\{\begin{pmatrix}0&0&0\\0&0&1\end{pmatrix}\tilde{A}^T\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}-\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}\right\}\tilde{A}\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}} \tag{6.31}$$

where $\lambda'' = 1/\sqrt{1-(\tilde{A}_{33})^2}$.

In summary, in this section we have worked out a transformation formalism without sky coordinates. One can do similar work to make the transformations in sky coordinates only, without antenna coordinates. We found that this field transformation is independent of the LST. This makes sense because the spherical coordinate system is symmetric with respect to the azimuthal angle. In contrast, it is dependent on the colatitude angle, or *declination*.

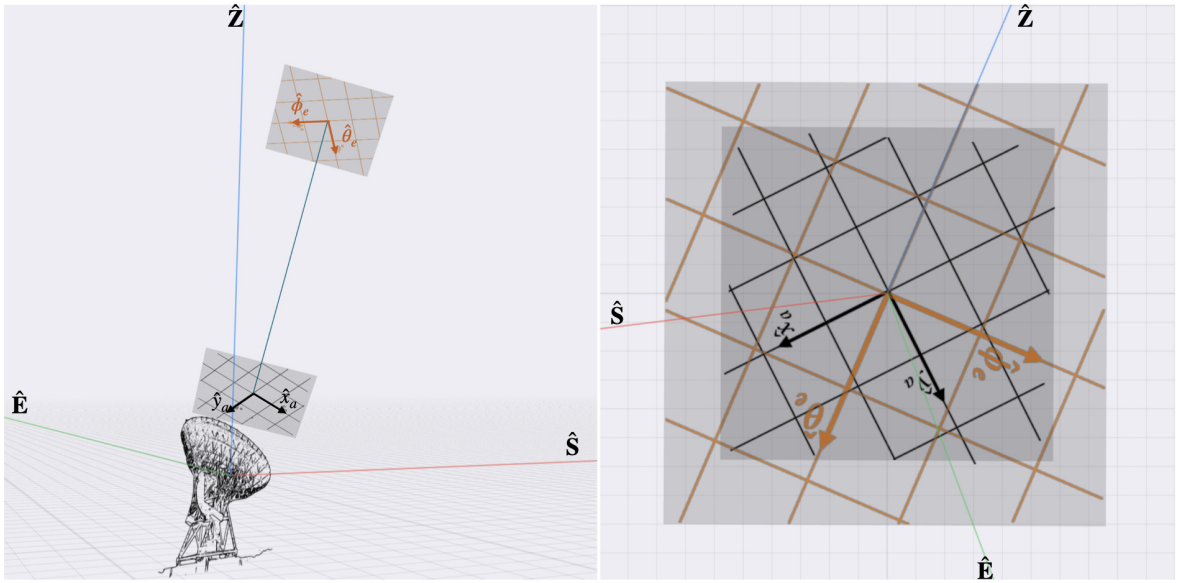

Figure 6.5: Illustration of the geometric orientation of an antenna. **Left panel**: The 3D orientation can be determined by the direction of a given axis and the tangential orientation on the plane perpendicular to the axis. A natural axis to be specified is the direction of antenna beam center, which is usually also the direction of the antenna Z-axis. $\hat{x}_a$ and $\hat{y}_a$ are the other two axes of the antenna, and $\hat{\theta}_e$ and $\hat{\phi}_e$ are spherical coordinate basis vectors of the celestial system at the pointing center. $\{\hat{x}_a, \hat{y}_a, \hat{z}_a\}$ and $\{\hat{\theta}_e, \hat{\phi}_e, \hat{z}_a\}$ can each form an orthonormal basis of the 3D Euclidean space. Obviously, $\hat{x}_a$ and $\hat{y}_a$ can be transformed from $\hat{\theta}_e$ and $\hat{\phi}_e$ by a rotation matrix, as implied in the **right panel** where the viewer's line of sight overlaps the antenna Z-axis but is looking from the sky towards the antenna. Two gray square planes both are perpendicular to the pointing direction, and the grids on them illustrate coordinate lines for the two coordinate systems respectively. The equatorial spherical coordinate lines (orange) are drawn in the flat-sky limit and thus look Euclidean.

6.2.4 Spatial orientation of the antenna

Section 6.2.2 provides a formalism that deals with all degrees of freedom of vector fields. This is often the application scenario required for simulation and data analysis. Sometimes, if we only want to understand the spatial relationship between the sky and the antenna during the measurement, we may only be interested in the 3D geometric orientation of the antennas relative to the celestial sphere. (For example, if you want to rotate the antenna during the observations.) In this case, we only need to specify the direction of the beam center (or *pointing center*) as well as the 2D orientation in the plane perpendicular to the pointing direction. They can be considered as the **axial** and **tangential** components of the geometric orientation of the antenna, respectively.

In the usual convention, the direction of the beam center is actually the same physical entity as the antenna Z-axis $\hat{z}_a$, while the tangential orientation of the antenna can be characterized by the connection between the other two antenna axes and the spherical coordinate basis vectors of the celestial system at the pointing center (see Figure 6.5).

Mathematically, the spatial orientation of the antenna can be specified by decomposing $\{\hat{x}_a, \hat{y}_a, \hat{z}_a\}$ with $\{\hat{\theta}_e, \hat{\phi}_e, \hat{r}_e\}$

$$\begin{pmatrix} \hat{x}_a \\ \hat{y}_a \\ \hat{z}_a \end{pmatrix} = P \begin{pmatrix} \hat{\mathbf{S}} \\ \hat{\mathbf{E}} \\ \hat{\mathbf{Z}} \end{pmatrix} = PR \begin{pmatrix} \hat{x}_e \\ \hat{y}_e \\ \hat{z}_e \end{pmatrix} = PR\mathcal{J}^T(\theta_e^c, \phi_e^c) \begin{pmatrix} \hat{r}_e \\ \hat{\theta}_e \\ \hat{\phi}_e \end{pmatrix} \tag{6.32}$$

where (θ_e^c, ϕ_e^c) is the equatorial spherical coordinates of the pointing center, which satisfies

$$\theta_e^c = \arccos(r_3), \quad \phi_e^c = \operatorname{sgn}(r_2) \arccos \frac{r_1}{\sqrt{r_1^2 + r_2^2}} \tag{6.33}$$

and $r_j = P_{3i}R_{ij}$. Finally, it's straightforward to find that the transformation in equation (6.32) has the following form

$$PR\mathcal{J}^T(\theta_e^c, \phi_e^c) \equiv \left(\begin{array}{c|cc} 0 & \mathcal{R}_{11} & \mathcal{R}_{12} \\ 0 & \mathcal{R}_{21} & \mathcal{R}_{22} \\ \hline 1 & 0 & 0 \end{array}\right) \tag{6.34}$$

where the upper right block transforming $\hat{\theta}_e$ and $\hat{\phi}_e$ to $\hat{x}_a$ and $\hat{y}_a$ has the form of some 2×2 rotation matrix $\mathcal{R}(\xi)$, and we shall use ξ to characterize the tangential orientation of the antenna.

In summary, the 3D geometric orientation of an antenna can be described by its axial direction $\hat{\mathbf{n}}^c \equiv (\theta_e^c, \phi_e^c)$ and the tangential orientation ξ, which can be derived from equation (6.34).

6.3 Power measurement

The time sequence of antenna voltage induced by the incident field is in the form of rapidly fluctuating incoherent noise. Typically, radio astronomers use narrowband bandpass filters to isolate a specific frequency range of interest from the broad range of radio frequencies received. Although after the bandpass filter the voltage is correlated in time from point to point [4], it is still a zero mean random variable in nature.

The simplest way to obtain continuous non-zero data is to correlate a pair of noise voltages and average them over a period of time. In other words, instead of measuring the voltage signal itself, we measure its power averaged over a time interval. As we will see below, such a power measurement, or let us say correlation or visibility, can be related to the power of the electric field of the sky, which is represented as a polarization tensor field (see section 5.2). This is due to the stochastic nature of the radio sky: The incoming sky electric field is generally incoherent in frequency and direction, and the

[4] This manifests as a low frequency modulation of the voltage envelope, estimated as the inverse of the frequency channel width, $1/\Delta\nu$.

two-point correlation of the field can be written as

$$\langle \mathcal{E}_i^*(\hat{n},\nu,t)\mathcal{E}_j(\hat{n}',\nu',t)\rangle_{\mathcal{T}} = \delta(\nu-\nu')\delta^2(\hat{n}-\hat{n}')\mathcal{P}_{ij}(\hat{n},\nu,t_0), \tag{6.35}$$

where the subscript $\mathcal{T}$ indicates that the average is taken over the time period $(t_0 - \mathcal{T}/2, t_0 + \mathcal{T}/2)$, and we have taken the continuous limit that the time and frequency coordinates are approximately continuous variables. Since the sky power $\mathcal{P}_{ij}$ is usually assumed to be constant over the observation time, the time coordinate t_0 is usually omitted.

6.3.1 Autocorrelation

The simplest way to measure power is to average the square of the voltage over a long period $\mathcal{T}$, the integration time. This type of measurement is called *autocorrelation*.

Since $\mathcal{T}$ is so short (say 2 seconds) that the Earth's rotation angle is negligible, the mapping between antenna coordinates and sky coordinates, f, is typically approximated as not changing with time. Correspondingly, $\tilde{\Lambda}$, the orthogonalized Jacobian of f is also constant over time. Thus, in linear algebra notations (see equation (6.8)), an autocorrelation data can be written as

$$\begin{aligned} d(\nu_i) &\equiv \langle \mathcal{V}\mathcal{V}^*\rangle_{\mathcal{T}} \\ &= \iint \mathrm{d}^2\hat{n} \iint \mathrm{d}^2\hat{n}'\ \mathbf{E}^T(\hat{n},\nu_i)\tilde{\mathbf{\Lambda}}(\hat{n})\big\langle \mathcal{E}(\hat{n},\nu_i,t)\mathcal{E}^\dagger(\hat{n}',\nu_i,t)\big\rangle_{\mathcal{T}}\tilde{\mathbf{\Lambda}}^T(\hat{n}')\mathbf{E}^*(\hat{n}',\nu_i) \\ &= \iint \mathrm{d}^2\hat{n}\ \mathbf{E}^T(\hat{n},\nu_i)\tilde{\mathbf{\Lambda}}(\hat{n})\mathcal{P}(\hat{n},\nu_i)\tilde{\mathbf{\Lambda}}^T(\hat{n})\mathbf{E}^*(\hat{n},\nu_i) \end{aligned} \tag{6.36}$$

where $\mathcal{P}$ is the sky polarization field, and we have assumed it to be constant over the observation time. Otherwise, both $\mathcal{P}$ and d will pick up the time coordinate.

Equation (5.3) allows us to decompose $\mathcal{P}$ into by Pauli matrices and Stokes parameters, so the autocorrelation data can be rewritten as

$$d(\nu_i) = \sum_{S=I,Q,U,V} \iint \mathrm{d}^2\hat{n}\ T_{\mathcal{S}}(\hat{n},\nu_i)\mathbf{E}^T(\hat{n},\nu_i)\tilde{\mathbf{\Lambda}}(\hat{n})\sigma^{\mathcal{S}}\tilde{\mathbf{\Lambda}}^T(\hat{n})\mathbf{E}^*(\hat{n},\nu_i). \tag{6.37}$$

where $T_{\mathcal{S}} = I, Q, U, V$ and $\sigma^{\mathcal{S}}$ are corresponding Pauli matrices. For convenience, we define the polarized beams

$$B^{\mathcal{S}}(\hat{n},\nu_i) \equiv \mathbf{E}^T(\hat{n},\nu_i)\tilde{\mathbf{\Lambda}}(\hat{n})\sigma^{\mathcal{S}}\tilde{\mathbf{\Lambda}}^T(\hat{n})\mathbf{E}^*(\hat{n},\nu_i), \tag{6.38}$$

Then the autocorrelation can be rewritten as

$$d(\nu_i) = \sum_{S=I,Q,U,V} \iint \mathrm{d}^2\hat{n}\; B^{\mathcal{S}}(\hat{n},\nu_i) T_{\mathcal{S}}(\hat{n},\nu_i). \tag{6.39}$$

Now let's take a closer look at the Stokes beams. Since $\tilde{\mathbf{\Lambda}}$ is an orthogonal matrix, we have $\tilde{\mathbf{\Lambda}}\sigma^{\mathcal{S}}\tilde{\mathbf{\Lambda}}^T = \sigma^{\mathcal{S}}$ for $\mathcal{S} = I, V$. Therefore, the Stokes I and V beams get rid of the transformation matrix and reduce to scalar beams

$$\boxed{B^{\mathcal{S}}(\hat{n},\nu_i) = \mathbf{E}^T(\hat{n},\nu_i)\sigma^{\mathcal{S}}\mathbf{E}^*(\hat{n},\nu_i), \qquad \text{for } \mathcal{S} = I, V.} \tag{6.40}$$

However, the Stokes Q and U beams depend on $\tilde{\mathbf{\Lambda}}$ and thus on the spatial orientation of the antenna, which is specified by the pointing matrix P, the local sidereal time, LST, and the latitude of the antenna location, alt. As discussed in section 6.2.3, $\tilde{\mathbf{\Lambda}}$ is independent of the galactic azimuths of the pointing center but dependent on the declination. Therefore, given 'lat' and P, the Q and U beams can be written in the antenna coordinates as

$$\boxed{B^{\mathcal{S}}(n^\alpha,\nu_i;P,\mathrm{lat}) = \mathbf{E}^T(n^\alpha,\nu_i)\tilde{\mathbf{\Lambda}}(n^\alpha,P,\mathrm{lat})\sigma^{\mathcal{S}}\tilde{\mathbf{\Lambda}}^T(n^\alpha,P,\mathrm{lat})\mathbf{E}^*(n^\alpha,\nu_i), \quad \text{for } \mathcal{S} = Q, U} \tag{6.41}$$

where $\tilde{\mathbf{\Lambda}}$ has been given by equation (6.31). For convenience, we define the transformed antenna far field $\tilde{\mathbf{E}} = \tilde{\mathbf{\Lambda}}^T\mathbf{E}$ and rewrite the above equation as

$$\boxed{B^{\mathcal{S}}(n^\alpha,\nu_i;P,\mathrm{lat}) = \tilde{\mathbf{E}}^T(n^\alpha,\nu_i;P,\mathrm{lat})\sigma^{\mathcal{S}}\,\tilde{\mathbf{E}}^*(n^\alpha,\nu_i;P,\mathrm{lat})}, \tag{6.42}$$

for $\mathcal{S} = I, Q, U, V$, although $\tilde{\mathbf{E}}$ is just $\mathbf{E}$ for Stokes I and V.

Since the integration time is finite, the averaged output power from the integrator is only an estimate of the true value, with an associated uncertainty that contributes as *noise*. In the Rayleigh-Jeans regime, both the average power and the fluctuations in the integrator output are expressed in temperature, The RMS fluctuation or "noise scale" is given by the radiometer equation:

$$\Delta T_{\mathrm{rms}} = \frac{T_{\mathrm{sys}}}{\sqrt{\Delta\nu\mathcal{T}}} \tag{6.43}$$

where T_{sys} is the system noise temperature and $\Delta\nu$ is the frequency channel width.

6.3.2 Cross-correlation and radio interferometry

Most telescopes have multiple receivers. Therefore, power measurement via correlation between voltage signals of two different receivers is possible. Such measurement is termed as the *cross-correlation* and the data acquired consists of the cross-correlations are known as *visibilities*.

For single-dish experiments, one can get a correlation measurement with pairs of receivers centered at the same position. While for interferometric observations, cross-correlation is more generally between two separate receivers and there is a constant phase difference between the signals from the same direction and frequency. As we do not consider the phased array, where time delay is introduced in integrator, the phase difference is determined by the projection of the baseline vector $\mathbf{b}$ onto the direction of the sky source $\hat{\mathbf{n}}$. Therefore, the visibility measured by cross-correlation of the receiver p on dish I and the receiver q on dish J reads

$$d_{Ip,Jq}(\nu_i) = \sum_{\mathcal{S}} \iint \mathrm{d}^2\hat{n}\ B^{\mathcal{S}}_{IJ,pq}(\hat{n},\nu_i) T_{\mathcal{S}}(\hat{n},\nu_i)\ \exp\left[-i(2\pi\nu_i/c)\hat{n}\cdot\vec{b}_{IJ}\right]. \tag{6.44}$$

where $\mathcal{S} = I, Q, U, V$ and the Stokes beams in the antenna coordinates n^α are given by

$$B^{\mathcal{S}}_{IJ,pq}(n^\alpha,\nu_i;P,\mathrm{lat}) = \tilde{\mathbf{E}}^T_{iA}(n^\alpha,\nu_i;P,\mathrm{lat})\sigma^{\mathcal{S}}\,\tilde{\mathbf{E}}^*_{jB}(n^\alpha,\nu_i;P,\mathrm{lat}) \tag{6.45}$$

The RMS fluctuations in the amplitude of the integrator output become

$$\Delta T_{\mathrm{rms,IJ}} = \sqrt{\frac{T_{\mathrm{sys,I}}T_{\mathrm{sys,J}}}{\Delta\nu\mathcal{T}}}. \tag{6.46}$$

6.4 Summary

This section describes the antenna measurements in detail. In some cases, the equations involve many parameters or indices, but space and readability did not allow a complete presentation of all dependencies. To prevent the reader from being overwhelmed by too many technical details and missing the forest for the trees, a short summary is given here.

We begin by reviewing a few physical entities:

- The arena is a celestial sphere. Think of it as a 2-sphere manifold. Each point on the surface of the sphere represents a physical direction, notated as $\hat{n}$.
- Fields on the celestial sphere:
  1. Sky electrical signal field, $\boldsymbol{\mathcal{E}}(\hat{n},\nu_i,t_j)$.
     - This is a 2-dimensional time-varying vector field.
     - Its power defines the sky polarization field, $\mathcal{P}(\hat{n},\nu_i) \equiv \langle \boldsymbol{E}\otimes\boldsymbol{E}^*\rangle$. This is an order 2 tensor field and is assumed to be constant over the observation.
  2. Antenna far field, $\boldsymbol{E}(\hat{n},\nu_i)$.
     - As a physical entity, it is per se time-invariant.

– However, as a vector field placed on the celestial sphere, it depends on the relative spatial relationship between the antenna and the celestial sphere. In this chapter we have used two ways of describing this spatial relation:

(a) $\{P, \text{lat}, \text{LST}\}$.

P is the pointing matrix of the antenna, which describes the local deployment of the antenna.

'lat' is the latitude of the antenna, and 'LST' is the local sidereal time of the observation.

(b) $\{\hat{n}_c, \xi\}$.

$\hat{n}_c$ is the point on the celestial sphere associated with the antenna's pointing center, which is the axial orientation of the antenna.

ξ is the tangent orientation. See section 6.2.4 for more details.

– Thus, if one considers observations pointed at multiple positions, it might be better to explicitly specify the antenna far field with its spatial orientations. We can write it as $\boldsymbol{E}(\hat{n}, \nu_i; P, \text{lat}, \text{LST})$ or $\boldsymbol{E}(\hat{n}, \nu_i; \hat{n}_c, \xi)$. But it would be even better to write it like this

$$\boxed{\boldsymbol{E} = \boldsymbol{E}(\hat{n} - \hat{n}_c, \nu_i, \xi)} \tag{6.47}$$

so that one can easily read the intrinsic invariance of the field. ξ is usually constant throughout the observations, so in most cases one can drop this coordinate and simply write the field as $\boldsymbol{E}(\hat{n} - \hat{n}_c, \nu_i)$.

Considering the practical needs of radio data analysis, we discussed in detail the representation of these physical entities in the antenna and celestial coordinate systems, as well as the relationships between them. They are summarized as follows:

- Two systems: the antenna coordinate system (O) and the celestial coordinate system (O'). We have

$$\begin{array}{lll} \hat{n} \xrightarrow{O} n^\alpha & \hat{n} \xrightarrow{O'} n^{\alpha'} & f(n^\alpha) = n^{\alpha'} \\ \boldsymbol{E} \xrightarrow{O} E^\alpha & \boldsymbol{E} \xrightarrow{O'} E^{\alpha'} & E^{\alpha'} = \Lambda^{\alpha'}{}_{\alpha} \mathcal{E}^\alpha \\ \boldsymbol{E} \xrightarrow{O} E^\alpha & \boldsymbol{\mathcal{E}} \xrightarrow{O'} \mathcal{E}^{\alpha'} & \mathcal{E}^{\alpha'} = \Lambda^{\alpha'}{}_{\alpha} \mathcal{E}^\alpha \end{array}$$

where $\Lambda^{\alpha'}$ is the Jocobian or orthogonalized Jacobian depending on the convention of the linear basis. (See section 6.1.) While in this work, we use normalized coordinate basis vectors so we should use the orthogonalized Jacobian, which is denoted as $\tilde{\Lambda}^{\alpha'}{}_{\alpha}$. The specific forms of f and $\tilde{\Lambda}$ are discussed in section 6.2.

- The voltage signal received by the antenna, in the form of matrix multiplication (equation (6.8)):

$$\begin{aligned}
\mathcal{V} &= \iint \mathrm{d}^2\hat{n}\, \mathbf{E}^T \tilde{\mathbf{\Lambda}} \mathcal{E} \\
&= \iint \mathrm{d}^2\hat{n}\, \mathbf{E}^T(\hat{n}-\hat{n}_c,\nu_i,\xi)\tilde{\mathbf{\Lambda}}(\hat{n},\hat{n}_c,\xi)\mathcal{E}(\hat{n},\nu_i,t_i) \\
\xrightarrow{O} &= \iint \mathrm{d}^2\hat{n}\, \mathbf{E}^T(n^\alpha,\nu_i)\tilde{\mathbf{\Lambda}}(n^\alpha;P,\mathrm{lat})\mathcal{E}(f(n^\alpha),\nu_i,t_i) \\
&= \iint \mathrm{d}^2\hat{n}\, \tilde{\mathbf{E}}^T(n^\alpha,\nu_i;P,\mathrm{lat})\mathcal{E}(f(n^\alpha),\nu_i,t_i) \\
&= \iint \mathrm{d}^2\hat{n}\, \tilde{\mathbf{E}}^T(\hat{n}-\hat{n}_c,\nu_i;P,\mathrm{lat})\mathcal{E}(\hat{n},\nu_i,t_i)
\end{aligned} \tag{6.48}$$

  where $\tilde{\mathbf{\Lambda}}$ is given by equation (6.31) and $\tilde{\mathbf{E}} = \mathbf{E}\tilde{\mathbf{\Lambda}}^T$.

- The power measurement of the antenna pair

$$d_{Ip,Jq}(\nu_i) = \sum_S \iint \mathrm{d}^2\hat{n}\; B^{\mathcal{S}}_{IJ,pq}(\hat{n}-\hat{n}_c,\nu_i;P,\mathrm{lat})T_{\mathcal{S}}(\hat{n},\nu_i)\; \exp\left[-i\,(2\pi\nu_i/c)\,\hat{n}\cdot\vec{b}_{IJ}\right], \tag{6.49}$$

  where the Stokes beams $B^{\mathcal{S}}_{IJ,pq}(\hat{n},\nu_i) \xrightarrow{O} B^{\mathcal{S}}_{IJ,pq}(n^\alpha,\nu_i;P,\mathrm{lat})$ is given by equation (6.45). Actually, the Stokes I and V beams are scalars that are independent of P and 'lat'. For a narrow sky-strip survey, the change in the P may be small so that we can ignore the change in the Stokes Q and U beams. Also, people usually do not change the antenna latitude during the observation. Under these assumptions, we can drop the variables P and 'lat' in $B^{\mathcal{S}}_{IJ,pq}$. For beams that are independent of P and 'lat', we can express them as

$$B^{\mathcal{S}}_{IJ,pq} = B^{\mathcal{S}}_{IJ,pq}(n^\alpha,\nu_i) = B^{\mathcal{S}}_{IJ,pq}(\hat{n}-\hat{n}_c,\nu_i). \tag{6.50}$$

  In this case, the measurement equation reduces to

$$\boxed{d_{Ip,Jq}(\nu_i) = \sum_S \iint \mathrm{d}^2\hat{n}\; B^{\mathcal{S}}_{IJ,pq}(\hat{n}-\hat{n}_c,\nu_i)T_{\mathcal{S}}(\hat{n},\nu_i)\; \exp\left[-i\,(2\pi\nu_i/c)\,\hat{n}\cdot\vec{b}_{IJ}\right]}, \tag{6.51}$$

  which is the usual formula we are familiar with.

Part III

Foregrounds

Chapter 7

Overview

Detecting the redshifted 21 cm signal is challenging. One of the biggest challenges is the extremely prominent foreground emission that obscures the cosmological signal. The foreground emission is typically two to five orders of magnitude brighter than the 21cm signal, depending on whether global or specific fluctuating modes are of interest. Although the foreground emission completely dominates the radio sky, we empirically expect it to have different spectral properties than the cosmological signal: For 21cm observations, we are only interested in the sky temperature, or Stokes I. While the spectral behavior of the Stokes I foreground is often thought to be smooth enough that most of its energy is concentrated in the low wavenumber modes in the Fourier space of frequency, the 21cm signal, which has a complex spectral structure, has the potential to dwarf the foreground in the higher wavenumber modes, making it deliverable to observers.

But this is not the whole story. On the one hand, foreground challenges could couple to instrumental systematics, which can lead to leakage of polarized emission into Stokes I measurements. Polarized foregrounds generally have a complex spectral structure due to Faraday rotation during propagation. This creates serious difficulties for foreground mitigation strategies based on spectral smoothness. On the other hand, 21 cm experiments are generally high dynamic range, i.e. we are interested in contrasts much smaller than the mean intensity. However, the known smoothness of the foreground spectrum, either as an extrapolation from the success of similar strategies in CMB experiments or as a result of theoretical toy modeling, does not provide direct evidence that the hypothesis would hold up to the dynamic range of 21 cm experiments. Therefore, a detailed study of the foreground emission is essential for 21 cm observations.

The diffuse low-frequency radio sky (see Figure 7.2) is dominated almost exclusively by relativistic synchrotron radiation, the radiation from accelerated electrons in the galactic magnetic field, and free-free emission, the radiation produced by the accelerated electrons in the Coulomb potential of ions. Our discussion of the diffuse

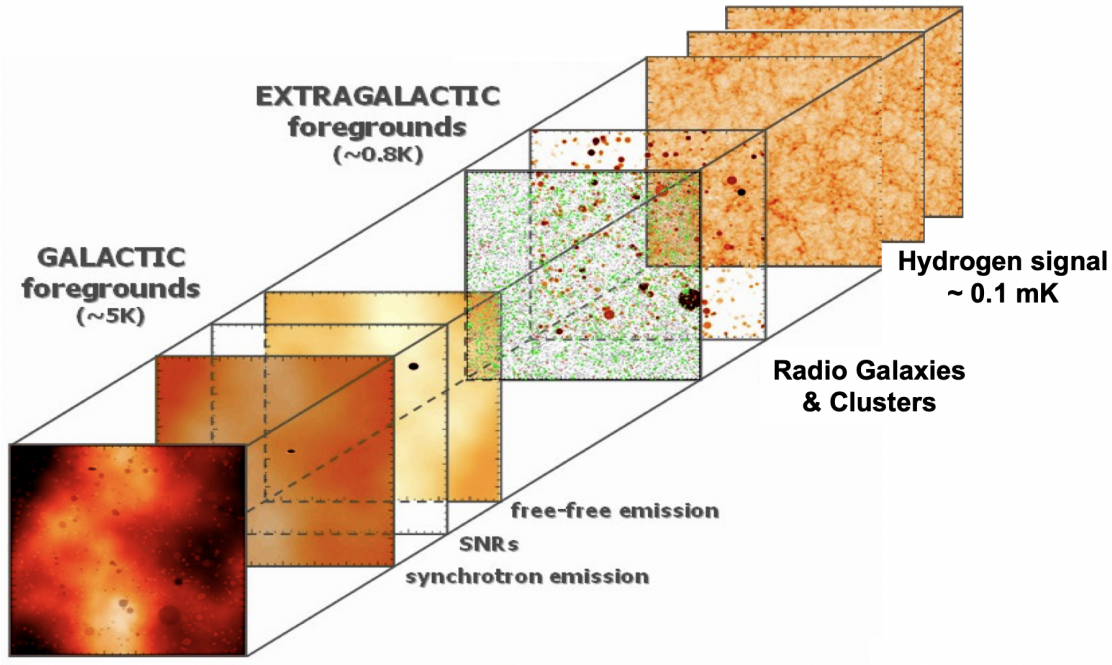

Figure 7.1: Decomposition of the total intensity of the radio sky. This figure is taken from Peter Timbie's slides.

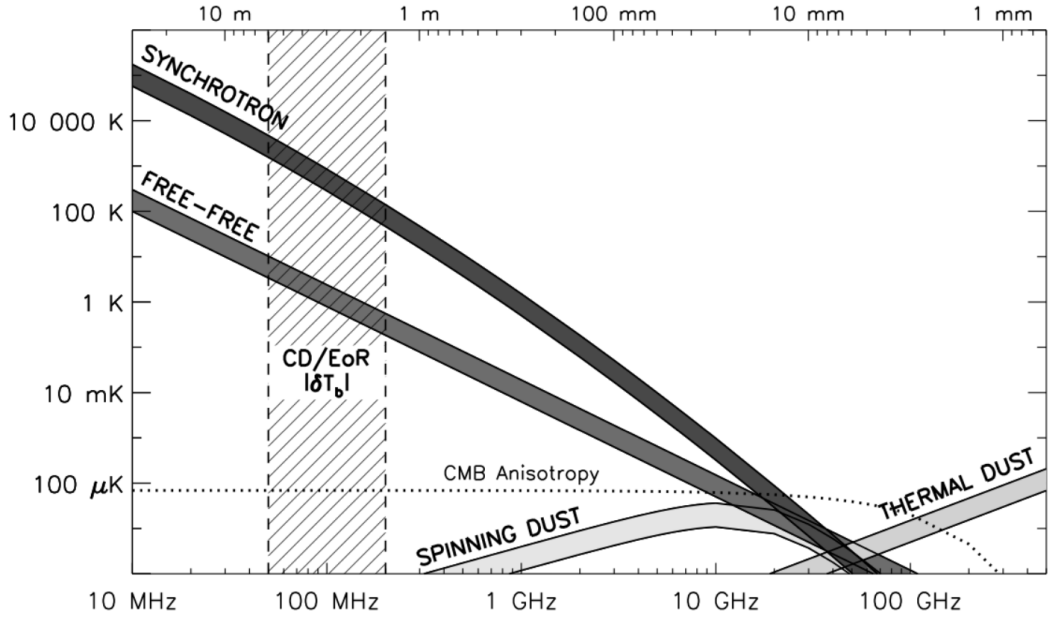

Figure 7.2: The total intensity of each of the major galactic diffuse foreground components. Synchrotron emission dominates at frequencies below $\sim$ 10 GHz. However, free-free emission also plays an important role in the 21cm experiments. This figure is taken from Chapman and Jelić (2019).

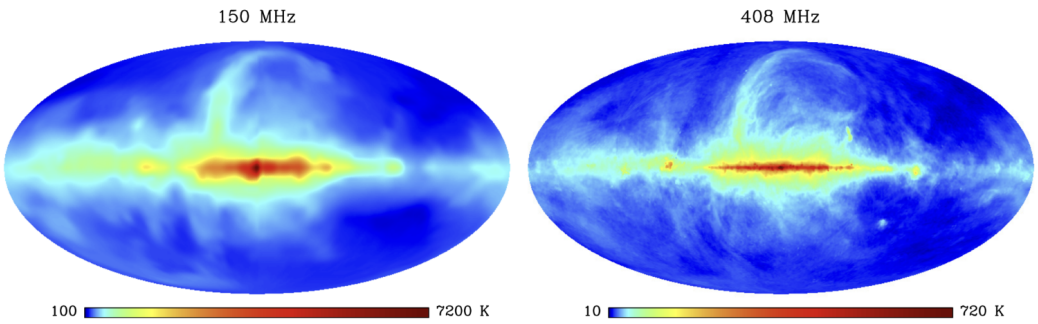

Figure 7.3: All sky maps of Galactic radio emission at 150 MHz (Landecker and Wielebinski, 1970) and 408 MHz (Haslam et al., 1982).

foreground is mainly concerned with the physical and observational properties of these two radiation processes.

In this overview chapter we briefly review the intensity (section 7.1) and polarization (section 7.2) of diffuse foregrounds at radio frequencies. In the remainder of this part III of the thesis we present physical analysis of the polarizatiobn of galactic synchron and free-free processes and the propagation of the radiation. The power spectrum description of the foreground intensity and polarization is dedicated in section 19.3 of part V, so we will not tedious on that this part.

7.1 Foreground instensity

As shown in the figure, 7.3 is the all-sky map at two radio frequencies. In the frequency range of interest for HI intensity mapping experiments, from 350 MHz to 1.4 GHz, the radio sky (the map as the sum of all emissions at the relevant frequency) is dominated by galactic synchrotron emission and galactic free-free emission. In this section we briefly review the smooth spectral behavior of the two emissions.

Intensity of galactic free-free emission

Galactic free-free emission, also known as thermal bremsstrahlung, arises when free electrons are scattered by ions without being captured. This process is devoid of quantum transitions between discrete energy levels, resulting in a continuous spectrum. At radio frequencies, free-free emission primarily operates in the optically thin regime (see e.g. (Condon and Ransom, 2016)). In this context, the majority of emitted photons escape without re-absorption. For this regime, the intensity of the radiation is given by:

$$I_\nu \propto \nu^{-2.1}$$

This is typical for HII regions. Due to this power-law dependence on frequency, we can observe a nearly consistent spectral index over radio wavelengths (Mezger and

Henderson, 1967; Churchwell et al., 2006). This consistency imparts the "smooth" spectral characteristic of free-free emission in the radio bands, distinguishing it from other spectral features associated with specific atomic or molecular transitions.

Intensity of galactic synchrotron emission

Relativistic electrons orbiting in magnetic fields are responsible for the radio emission of galactic synchrotron radiation. Although we will give a first-principles discussion of its power-law spectrum later in this part, here we qualitatively review several factors that cause its smooth spectral nature.

First, the cosmic ray electrons exhibit a continuous power-law energy distribution, which implies a non-preferential state for specific energy transitions as seen in atomic or molecular lines. Instead, electrons contribute to the emission over a wide energy range. This generally results in continuous synchrotron emission.

On the other hand, unlike the discrete spectral lines resulting from specific atomic or molecular transitions, synchrotron emission results from a broader mechanism, resulting in a broad and smooth spectrum. For a highly relativistic electron, the main lobe of the synchrotron radiation beam is very narrow. One rotation period of the electron leaves the observer with a smooth pulse, which in the time domain is much narrower than the rotation period of the electron. Therefore, the emission in the frequency domain has a broad, smooth structure with a cutoff frequency determined by the rotation period of the electron.

The emitted synchrotron radiation typically follows a power-law spectrum as a function of frequency, given by $I_\nu \propto \nu^\alpha$, where α is the spectral index (often negative). In general, for many galactic regions the spectral index revolves around $\alpha \approx -0.7$, but can show variability (Chapman and Jelić, 2019). The inherent power-law character results in a smooth spectral decline with increasing frequency.

7.2 Foreground polarization

The polarizations of both physical processes can be precisely computed, allowing for a theoretical discussion of their polarization. On a qualitative level, given that ions in a plasma can manifest in any direction relative to the incident electron beam, it's reasonable to consider unpolarized free-free emission as a suitable approximation when averaged over an ensemble.

Galactic synchrotron emission exhibits partial linear polarization. When we assume an isotropic power-law distribution of cosmic ray electrons, the linear polarization degree is described by equation (10.19). Typically, the Stokes V is approximated as 0, as detailed in section 10.3. In the vast regions of the Galactic magnetic field that encompass HII regions, the Faraday rotation effect impacts linearly polarized waves,

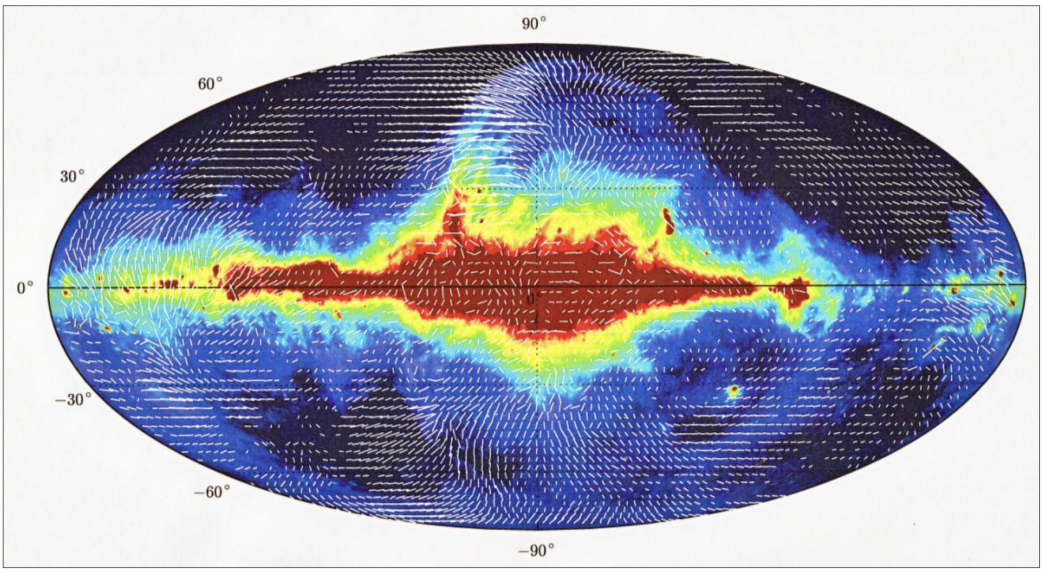

Figure 7.4: All sky polarizations (E-vectors) observed at 1400 MHz overlaid on a total intensity color map. This is taken from Wielebinski (2012).

leading to intricate spectral patterns in the linearly polarized foreground, as discussed in section 9.3.

Chapter 8

Radiation from moving charges

We are concisely review the necessary theoretical tools to analyze radiation emitted by moving charges. The velocities and positions of all charges can be described by $f(\boldsymbol{x}, \boldsymbol{p}, t)$, the phase space distribution of the charges in the observer's frame of reference. From $f(\boldsymbol{x}, \boldsymbol{p}, t)$, we can derive the 4-current density, j^α, and calculate the electromagnetic 4-potential A^α (Section 8.1). It is straightforward to obtain the electromagnetic fields using the differentiations of the potentials (Section 8.2).

8.1 Liénard-Wiechert potentials

For particles of identical charge q, a useful expression for the 4-current density is in terms of the distribution function in 6-dimensional phase space $f(\boldsymbol{x}, \boldsymbol{p}, t)$, which is described by the particles' trajectories in the phase space, $\boldsymbol{x}_P(t)$ and $\boldsymbol{p}_P(t)$:

$$f(\boldsymbol{x}, \boldsymbol{p}, t) = \sum_P \delta^3(\boldsymbol{x} - \boldsymbol{x}_P(t))\delta^3(\boldsymbol{p} - \boldsymbol{p}_P(t)) \tag{8.1}$$

Then the 4-current density is

$$j^\alpha = q \int d^3p \; \dot{x}^\alpha(\boldsymbol{p}, t) f(\boldsymbol{x}, \boldsymbol{p}, t) \tag{8.2}$$

Calculating out the integral over the momentum space we have

$$j^\alpha = q \sum_P \dot{x}^\alpha_P(t) \, \delta^3(\boldsymbol{x} - \boldsymbol{x_P}(t)) \tag{8.3}$$

The electromagnetic 4-potential then can be expressed as the 4-current convolved with the Green's function:

$$\begin{aligned} A^\alpha &= \frac{4\pi}{c} j^\alpha * G \\ &= \frac{4\pi q}{c} \sum_P \int d^3x' \int dt' \; \dot{x}_P^\alpha(t') \, \delta^3(\boldsymbol{x'} - \boldsymbol{x_P}(t')) G(\boldsymbol{x} - \boldsymbol{x'}, t - t') \\ &= \frac{4\pi q}{c} \sum_P \int dt' \; \dot{x}_P^\alpha(t') \, G(\boldsymbol{x} - \boldsymbol{x_P}(t'), t - t') \end{aligned} \tag{8.4}$$

where G is the Green's function of the d'Alembertian operator

$$G(\boldsymbol{x} - \boldsymbol{x'}, t - t') = \frac{\delta(t - t' - |\boldsymbol{x} - \boldsymbol{x'}|/c)}{4\pi|\boldsymbol{x} - \boldsymbol{x'}|}, \tag{8.5}$$

The delta function in the G is the Dirac delta function composed with a "retarded" function and the integral in this formula basically means that the particle P's contribution to the field at the point of observation at time t is determined by its state of motion at the earlier time t', for which the time of propagation from $\boldsymbol{x}_P(t')$ to the field point $\boldsymbol{x}$ just coincides with the difference $t - t'$. To distinguish, we shall call t the *observation time*, and t' the *emission time*, while $t - t'$ is named the *retarded time*. Let $\boldsymbol{R_P}(t') = \boldsymbol{x} - \boldsymbol{x_P}(t')$ be the radius vector from the particle to the field point and

$$R_P(t') \equiv |\boldsymbol{x} - \boldsymbol{x_P}(t')| \;\text{ and }\; \boldsymbol{n_P}(t') \equiv \frac{\boldsymbol{x} - \boldsymbol{x_P}(t')}{|\boldsymbol{x} - \boldsymbol{x_P}(t')|}, \tag{8.6}$$

then the signal emitted by particle P at $\boldsymbol{x_P}(t')$ at time t' can be observed at $\boldsymbol{x}$ at time

$$t = F_P(t') \equiv t' + R_P(t')/c. \tag{8.7}$$

To calculate the integral in equation (8.4), we shall regard the Green's function, equation (8.5), as a composition with a function of t'. Breaking down the composition, the Dirac delta function in the Green's function becomes

$$\delta(t - F_P(t')) = \frac{\delta(F_P^{-1}(t) - t')}{F_P'(t')}, \tag{8.8}$$

where F_P^{-1} is the inverse function of F_P and $F_P'(t')$ is the derivative of $F_P(t')$

$$F_P'(t') = 1 - \boldsymbol{n_P}(t') \cdot \boldsymbol{\beta_P}(t') \tag{8.9}$$

where $\boldsymbol{\beta_P} = \dot{\boldsymbol{x}}_{\boldsymbol{P}}/c$.

Substituting equation (8.8) into equation (8.4), we have

$$A^\alpha(t,\boldsymbol{x}) = q\sum_P \frac{\dot{x}_P^\alpha(t'_P)/c}{R_P(t'_P)(1-\boldsymbol{n_P}(t'_P)\cdot\boldsymbol{\beta_P}(t'_P))} \tag{8.10}$$

where $t'_P \equiv F_P^{-1}(t)$. Transforming to three-dimensional notation, we obtain the following expressions:

$$\phi = q\sum_P \frac{1}{R_P(1-\boldsymbol{n_P}\cdot\boldsymbol{\beta_P})}, \qquad \boldsymbol{A} = q\sum_P \frac{\boldsymbol{\beta_P}}{R_P(1-\boldsymbol{n_P}\cdot\boldsymbol{\beta_P})}, \tag{8.11}$$

where all the quantities on the right sides are evaluated at the emission time t'. These potentials are known as the *Liénard-Wiechert potentials*.

8.2 EM fields of moving charges

Next we shall find the electromagnetic field. There are several ways to do this. One is to find the fields in the instantaneous rest frame of the particle and to Lorentz-transform to the observer's frame. Another is to take derivatives of the integrals for the potentials, equation (8.4), a procedure followed in Jackson's *Classical Electrodynamics*. Yet another is simply taking derivatives of the Liénard-Wiechert potentials, equation (8.10) and (8.11), which is about to be employed here.

To calculate the field strengths of the electric and magnetic fields from Liénard-Wiechert potentials, we must differentiate the electromagnetic 4-potential with respect to the field coordinates $\boldsymbol{x}$ and the observation time t. However, the potentials are expressed as functions of t', while t' itself can be regarded as a function $t'(t,\boldsymbol{x})$. Therefore, to calculate the desired derivatives, we should first calculate the derivatives with respect to t'.

For a specific particle P, using the relation $R_P = c(t-t'_P)$ we can calculate the differentiations

$$\frac{\partial t'_P}{\partial t} = \frac{1}{1-\boldsymbol{n_P}\cdot\boldsymbol{\beta_P}} \tag{8.12}$$

and

$$\frac{\partial t'_P}{\partial x^i} = -\frac{1}{c}\frac{\partial R_P}{\partial x^i} = -\frac{1}{c}\left[-(\boldsymbol{n_P}\cdot\dot{\boldsymbol{x}}_{\boldsymbol{P}})\frac{\partial t'_P}{\partial x^i} + \boldsymbol{n_P}\right] \tag{8.13}$$

which implies

$$\frac{\partial t'_P}{\partial x^i} = -\frac{\boldsymbol{n_P}}{c(1-\boldsymbol{n_P}\cdot\boldsymbol{\beta_P})}. \tag{8.14}$$

Now the fields $\boldsymbol{E}$ and $\boldsymbol{B}$ can be calculated with the aid of these differentiations

$$\boldsymbol{E}(\boldsymbol{x},t) = \sum_P \boldsymbol{E_P} = \sum_P \left\{ q\left(\frac{\boldsymbol{n}-\boldsymbol{\beta}}{\gamma^2\kappa^3R^2}\right) + \frac{q}{c}\left[\frac{\boldsymbol{n}\times[(\boldsymbol{n}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}]}{\kappa^3 R}\right]\right\}_P \tag{8.15}$$

where $\kappa = 1 - \boldsymbol{n} \cdot \boldsymbol{\beta}$. For the sake of brevity, the lower index P that marks the particles is omitted. However, we should keep in mind that all physical quantities on the right side are evaluated at the t' of the corresponding charged particle. The magnetic field may be determined by

$$\boldsymbol{B}(\boldsymbol{x}, t) = \sum_P \{\boldsymbol{n_P} \times \boldsymbol{E_P}(\boldsymbol{x}, t)\} . \tag{8.16}$$

The first term in equation (8.15) is called the *velocity term* as it doesn't involve $\dot{\boldsymbol{\beta}}$; in contrast the second term is called the acceleration term. If there is no acceleration ($\dot{\boldsymbol{\beta}} = 0$), then the second term vanishes while the first term falls off with distance as $1/r^2$, which does not give rise to a net flux of electromagnetic fields to infinity. If $\dot{\boldsymbol{\beta}} \neq 0$, then the acceleration term is nontrivial and it falls off as $1/r$ giving rise to a net flux, or in jargon *radiation*. It means that the charges emit radiations only if they are accelerated.

8.3 Radiation from accelerated charge

Astrophysical synchrotron sources are distant enough so that we can take the limit of large R. In this case, only the radiation terms of the fields are kept. The electric field induced by a single particle becomes

$$\boldsymbol{E}(\boldsymbol{x}, t) = \frac{q}{c} \left[\frac{\boldsymbol{n} \times [(\boldsymbol{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{\kappa^3 R} \right] \tag{8.17}$$

To study the angular distribution of the radiation, it is useful to decompose all the vectors into parts parallel and perpendicular to $\boldsymbol{\beta}$ so that $\boldsymbol{n} = \boldsymbol{n}_\perp + \boldsymbol{n}_\parallel$ and $\dot{\boldsymbol{\beta}} = \dot{\boldsymbol{\beta}}_\perp + \dot{\boldsymbol{\beta}}_\parallel$. The electric field written in terms of these components is

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{x}, t) &= \frac{q}{c} \left[\frac{(\boldsymbol{n}_\perp + \boldsymbol{n}_\parallel) \times [(\boldsymbol{n}_\parallel - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}_\perp + \boldsymbol{n}_\perp \times (\dot{\boldsymbol{\beta}}_\parallel + \dot{\boldsymbol{\beta}}_\perp)]}{\kappa^3 R} \right] \\ &= \frac{q}{c\kappa^3 R} \left[-\dot{\boldsymbol{\beta}}_\perp (1 - \boldsymbol{n} \cdot \boldsymbol{\beta}) + (\boldsymbol{n} - \boldsymbol{\beta})(\boldsymbol{n}_\perp \cdot \dot{\boldsymbol{\beta}}_\perp) - \dot{\boldsymbol{\beta}}_\parallel (\boldsymbol{n}_\perp \cdot \boldsymbol{n}_\perp) + \boldsymbol{n}_\perp (\boldsymbol{n}_\parallel \cdot \dot{\boldsymbol{\beta}}_\parallel) \right] \end{aligned} \tag{8.18}$$

The right hand side of the above equation contains the four components of the electric field. They each represent the electric field components induced by the radial or tangential particle acceleration.

The power per solid angle radiated by the moving charge is denoted by $\mathcal{P}'(\boldsymbol{n}, t')$. Correspondingly, $\mathcal{P}(\boldsymbol{n}, t)$ is the observed power per solid angle. The radiated energy

can be equivalently expressed by both $\mathcal{P}'(\boldsymbol{n}, t')$ and $\mathcal{P}(\boldsymbol{n}, t)$ so that

$$dW = \mathcal{P}'(\boldsymbol{n}, t')\, d\Omega dt' = \mathcal{P}(\boldsymbol{n}, t)\, d\Omega dt = R^2 \boldsymbol{S} \cdot \boldsymbol{n}\, d\Omega dt. \tag{8.19}$$

By realizing that $\partial t'/\partial t = 1/\kappa$ which is implied by equation (8.12), the above equations give

$$\frac{\mathcal{P}'}{\kappa} = \mathcal{P} = R^2 \boldsymbol{S} \cdot \boldsymbol{n} \tag{8.20}$$

where $\boldsymbol{S}(\boldsymbol{x}, t)$ is the Poynting vector given by

$$\boldsymbol{S} = \frac{c}{4\pi} \boldsymbol{E} \times \boldsymbol{B}, \tag{8.21}$$

which represents the directional energy flux of the electromagnetic fields. Since only the radiation term is considered, we have $\boldsymbol{n} \cdot \boldsymbol{E} = \boldsymbol{n} \cdot \boldsymbol{B} = 0$ and $\boldsymbol{B} = \boldsymbol{n} \times \boldsymbol{E}$. Therefore, the Poynting vector can be rewritten as

$$\boldsymbol{S} = \frac{c}{4\pi} \boldsymbol{n} |\boldsymbol{E}|^2. \tag{8.22}$$

The observed intensity at time t depends upon the state of the particle at t', and the differential time elements are related by $dt = (1 - \boldsymbol{n} \cdot \boldsymbol{\beta}) dt'$. Thus, if one wants to calculate the radiated power observed at field point $\boldsymbol{x}$ at time t, then they will have to calculate the retardation. If a particle is impulsively accelerated for a time dt', a pulse of radiation will later appear at the observer of duration $(1 - \boldsymbol{n} \cdot \boldsymbol{\beta}) dt'$.

Here we calculate the total power instantaneously radiated by the moving particle, which is given by

$$\mathcal{P}'_{tot}(t') = \int d\Omega\, \mathcal{P}'(\boldsymbol{n}, t') = \frac{2q^2\gamma^6}{3c} [\dot{\beta}^2 - (\boldsymbol{\beta} \times \dot{\boldsymbol{\beta}})^2] \tag{8.23}$$

This is known as the relativistic generalization of the *Larmor formula*. Let the angle between β and $\dot{\beta}$ be θ_0. Then the total power radiated can be expressed as

$$\mathcal{P}'_{tot}(t') = \frac{2\dot{\beta}^2 q^2 \gamma^6}{3c} (1 - \beta^2 \sin^2 \theta_0) \tag{8.24}$$

$$= \frac{2q^2\gamma^4}{3c} \left(\dot{\beta}_\perp^2 + \gamma^2 \dot{\beta}_\parallel^2 \right) \tag{8.25}$$

In addition to the power of the radiation field, another result that often needs to be considered is the energy per unit frequency per unit solid angle. Using Parseval's theorem we have

$$\frac{\mathrm{d}W}{\mathrm{d}\Omega} \equiv \frac{c}{4\pi} \int |R\, \boldsymbol{E}(t)|^2\, \mathrm{d}t = \frac{c}{4\pi} \int \left| \int R\, \boldsymbol{E}(t) e^{-i\omega t}\, \mathrm{d}t \right|^2 \frac{\mathrm{d}\omega}{2\pi}, \tag{8.26}$$

then the spectrum is given by

$$\begin{aligned}\frac{\mathrm{d}W}{\mathrm{d}\omega\,\mathrm{d}\Omega} &= \frac{c}{8\pi^2}\left|\int R\,\boldsymbol{E}(t)e^{-i\omega t}\,\mathrm{d}t\right|^2 \\ &= \frac{q^2}{8\pi^2 c}\left|\int \frac{\boldsymbol{n}\times[(\boldsymbol{n}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}]}{\kappa^3}e^{-i\omega t}\,\mathrm{d}t\right|^2\end{aligned} \tag{8.27}$$

If the integration variable is replaced by retarded time, $\mathrm{d}t = \kappa\,\mathrm{d}t'$, and using the approximation

$$t' = t - \frac{R(t')}{c} \approx t - \frac{|\boldsymbol{x}| - \boldsymbol{n}\cdot\boldsymbol{x_P}(t')}{c}, \tag{8.28}$$

which is valid for the coordinate systems satisfying $|\boldsymbol{x_P}| \ll |\boldsymbol{x}|$, we can rewrite the above integral as

$$\frac{\mathrm{d}W}{\mathrm{d}\omega\,\mathrm{d}\Omega} = \frac{q^2}{8\pi^2 c}\left|\int \frac{\boldsymbol{n}\times[(\boldsymbol{n}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}]}{\kappa^2}\exp\left[-i\omega\left(t' - \frac{\boldsymbol{n}\cdot\boldsymbol{x_P}(t')}{c}\right)\right]\mathrm{d}t'\right|^2. \tag{8.29}$$

Integrating by parts, we have

$$\frac{\mathrm{d}W}{\mathrm{d}\omega\,\mathrm{d}\Omega} = \frac{q^2\omega^2}{8\pi^2 c}\left|\int \boldsymbol{n}\times(\boldsymbol{n}\times\boldsymbol{\beta})\exp\left[-i\omega\left(t' - \frac{\boldsymbol{n}\cdot\boldsymbol{x_P}(t')}{c}\right)\right]\mathrm{d}t'\right|^2. \tag{8.30}$$

Chapter 9

Propagation in the ISM

A gas is composed of neutral particles, and the interactions between neutral particles at large distances result from van der Waals forces, which are inversely proportional to the seventh power of the distance between interacting particles. When the average kinetic energy of the particles exceeds the ionization potential of the atoms, some neutral particles are ionized, and the gas becomes an overall electrically neutral system composed of positive ions, negative ions, and neutral particles. This state of matter is called *plasma*.

In addition to van der Waals forces between neutral particles in a plasma, there are also interactions between charged particles and interactions between neutral and charged particles. The interaction between two charged particles is described by Coulomb's law, and the potential is inversely proportional to the distance. The interaction between a neutral particle and a charged particle is also essentially electrostatic; the charged particle induces a dipole moment in the neutral particle that is proportional to the electric field of the charged particle, i.e., proportional to the inverse square of the distance. Since the potential of the interaction is proportional to the product of the dipole moment and the electric field, it is inversely proportional to the fourth power of the distance. If we compare the potentials of these three long-range interactions as a function of distance,

$$\text{neutral-neutral: } U_{nn} \propto r^{-6}$$
$$\text{neutral-charged: } U_{nc} \propto r^{-4}$$
$$\text{charged-charged: } U_{cc} \propto r^{-1}$$

we can clearly see that the Coulomb interaction between charged particles has the largest range. This explains why Coulomb interactions are more important than other interactions in most cases.

In radio astronomy, almost all of the radio emission in the Milky Way comes from HII regions. In the HII region, hydrogen atoms are highly ionized, and there are large

numbers of free electrons and ions that maintain the electrical neutrality of the plasma. However, because the mobility of the electrons far exceeds that of the ions, the electrons dominate the propagation of radiation in the plasma.

Since the radiation field propagating in the plasma frequently interacts with the charged particles, the radiation field can no longer be described as freely propagating waves in a vacuum over long distances. Fortunately, however, just as ordinary plane waves are fundamental solutions of homogeneous Maxwell's equations, if we consider the propagating radiation field as a perturbation of the charged particles, then we can solve for the linear responses of the plasma EM fields to the small perturbations. Our discussion in this chapter will illustrate this point. These linear responses are called plasma waves, and the simplest waves have the form of

$$\boldsymbol{E} = E_0 \exp i(\mathbf{k}\cdot\mathbf{r} - \omega t) \tag{9.1}$$

$$\boldsymbol{B} = B_0 \exp i(\mathbf{k}\cdot\mathbf{r} - \omega t) \tag{9.2}$$

We will consider the *cold plasma waves*; by "cold", we mean that the phase velocity is much greater than the thermal velocity of the particles in the plasma

$$\frac{\omega}{k} \gg v_t, \tag{9.3}$$

so that the particles with different velocities see approximately the same EM fields. For cold plasma, a non-relativistic description is a good approximation.

9.1 Dispersion in cold and isotropic plasma

We begin by discussing the propagation of radiation in a cold, isotropic plasma. The word 'cold' implies that the magnetic force on the electrons, which is of the order of v/c, can be neglected in the perturbation of the electromagnetic radiation. 'Isotropic' means that there is no external magnetic field. Therefore, the dynamical equation of the electrons is given by

$$m\dot{\boldsymbol{v}} = -e\boldsymbol{E} \quad \xrightarrow{\textbf{In Fourier space}} \quad \boldsymbol{v} = \frac{e\boldsymbol{E}}{im\omega}. \tag{9.4}$$

The current density, defined as $\boldsymbol{j} = -ne\boldsymbol{v}$, where n is the electron number density, is given by

$$\boldsymbol{j} = -\frac{ne^2}{im\omega}\boldsymbol{E}. \tag{9.5}$$

The charge density, ρ, is related to $\boldsymbol{j}$ by the charge conservation equation,

$$-i\omega\rho + i\boldsymbol{k}\cdot\boldsymbol{j} = 0, \tag{9.6}$$

which in turn gives the charge density

$$\rho = \omega^{-1}\boldsymbol{k}\cdot\boldsymbol{j} = -\frac{ne^2}{im\omega^2}\boldsymbol{k}\cdot\boldsymbol{E}. \tag{9.7}$$

The EM fields sourced by ρ and $\boldsymbol{j}$ are characterized by Maxewell's equations, which in Fourier space read as below

$$i\boldsymbol{k}\cdot\boldsymbol{E} = 4\pi\rho, \qquad \rightarrow \qquad i\boldsymbol{k}\cdot\epsilon\boldsymbol{E} = 0 \tag{9.8}$$

$$i\boldsymbol{k}\times\boldsymbol{E} = i\frac{\omega}{c}\boldsymbol{B}, \qquad \rightarrow \qquad i\boldsymbol{k}\times\boldsymbol{E} = i\frac{\omega}{c}\boldsymbol{B}, \tag{9.9}$$

$$i\boldsymbol{k}\cdot\boldsymbol{B} = 0, \qquad \rightarrow \qquad i\boldsymbol{k}\cdot\boldsymbol{B} = 0, \tag{9.10}$$

$$i\boldsymbol{k}\times\boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{j} - i\frac{\omega}{c}\boldsymbol{E}, \qquad \rightarrow \qquad i\boldsymbol{k}\times\boldsymbol{B} = -i\frac{\omega}{c}\epsilon\boldsymbol{E} \tag{9.11}$$

where in the right column we have used the expressions of $\boldsymbol{j}$ and ρ in terms of EM fields. And we have defined the

$$\epsilon \equiv 1 - \frac{4\pi ne^2}{m\omega^2}. \tag{9.12}$$

Since in the right column, Maxwell's equations have been rendered in a "source free" way, from which we get the homogeneous wave equations

$$\left(\epsilon\frac{\omega^2}{c^2} - k^2\right)\boldsymbol{E} = 0, \qquad \left(\epsilon\frac{\omega^2}{c^2} - k^2\right)\boldsymbol{B} = 0. \tag{9.13}$$

Thus, we have the dispersion relation for both $\boldsymbol{B}$ and $\boldsymbol{E}$ waves:

$$c^2k^2 = \epsilon\omega^2 = \omega^2 - \omega_p^2 \tag{9.14}$$

where

$$\omega_p = \sqrt{4\pi ne^2/m} \tag{9.15}$$

is the so-called *plasma frequency*. We see that for $\omega < \omega_p$, the wave $\exp i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)$ decreases exponentially. For $\omega > \omega_p$, the electromagnetic radiation propagates with the phase velocity

$$v_p \equiv \frac{\omega}{k} = \frac{c}{\sqrt{1 - \omega_p^2/\omega^2}}, \tag{9.16}$$

which is greater than c, and the group velocity is

$$v_g \equiv \frac{\partial\omega}{\partial k} = c\sqrt{1 - \omega_p^2/\omega^2}, \tag{9.17}$$

which is always less than c. The factor $\sqrt{1 - \omega_p^2/\omega^2} = \sqrt{\epsilon}$ is defined as the *index of refraction*.

9.2 Dispersion in anisotropic plasmas

In the previous section, we examined the behavior of electromagnetic radiation in isotropic plasma. However, in the interstellar medium (ISM) of the Milky Way galaxy, there exist pervasive magnetic fields known as the galactic magnetic fields, which serve as an exerted background magnetic field for the plasma.

The magnetic field in the ISM has both large-scale and small-scale components. The large-scale magnetic field, often referred to as the mean or regular field, is characterized by a coherent structure that can be described by a simple model Beck (2016). In the Milky Way, the regular field has a spiral pattern, following the spiral arms of the galaxy, with both azimuthal and radial components.

Small-scale irregular magnetic fields, also known as turbulent fields, are superimposed on the regular field. These turbulent fields are the result of various processes, such as supernova explosions, stellar winds, and magneto-hydrodynamic (MHD) turbulence, which inject energy into the ISM and cause fluctuations in the magnetic field strength and direction Ferriere (2001). For more details, we refer readers to a dedicated overview on the galactic magnetic fields, e.g. Beck (2004). In this section, we assume a background magnetic field $\boldsymbol{B^0}$ constant in time and uniform in space. This approximation works well for the background large-scale magnetic field.

When the background magnetic field, $\boldsymbol{B^0}$, is much stronger than the field strengths of the propagating wave and is assumed to be constant in time, the equation of motion for an electron in the plasma can be approximated as:

$$m\frac{d\boldsymbol{v}}{dt} = -e\boldsymbol{E} - \frac{e}{c}\boldsymbol{v} \times \boldsymbol{B^0} \tag{9.18}$$

where m is electron mass. In Fourier coordinates, this equation can be written as:

$$im\omega\boldsymbol{v} + \frac{e}{c}\boldsymbol{B^0} \times \boldsymbol{v} = e\boldsymbol{E} \xrightarrow{\boldsymbol{\omega_B} \equiv e\boldsymbol{B^0}/mc} i\omega\boldsymbol{v} + \boldsymbol{\omega_B} \times \boldsymbol{v} = \frac{e}{m}\boldsymbol{E} \tag{9.19}$$

where we have defined the *cyclotron frequency*, ω_B:

$$\boldsymbol{\omega_B} = \frac{e\boldsymbol{B^0}}{mc} = \begin{pmatrix} eB_1^0/mc \\ eB_2^0/mc \\ eB_3^0/mc \end{pmatrix}. \tag{9.20}$$

Equation (9.19) can be further expressed in the matrix form as following

$$\Omega \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \frac{e}{m} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}, \tag{9.21}$$

where the matrix Ω is given by

$$\Omega \equiv \begin{pmatrix} i\omega & -\omega_{B_3} & \omega_{B_1} \\ \omega_{B_3} & i\omega & -\omega_{B_2} \\ -\omega_{B_1} & \omega_{B_2} & i\omega \end{pmatrix}, \tag{9.22}$$

whose inverse matrix Ω^{-1} is

$$\Omega^{-1} = \frac{1}{i\omega(\omega^2 - \omega_B^2)} \begin{pmatrix} \omega^2 - \omega_{B_2}^2 & -\omega_{B_1}\omega_{B_2} + i\omega\omega_{B_3} & -\omega_{B_2}\omega_{B_3} - i\omega\omega_{B_1} \\ -\omega_{B_1}\omega_{B_2} - i\omega\omega_{B_3} & \omega^2 - \omega_{B_1}^2 & -\omega_{B_1}\omega_{B_3} + i\omega\omega_{B_2} \\ -\omega_{B_2}\omega_{B_3} + i\omega\omega_{B_1} & -\omega_{B_1}\omega_{B_3} - i\omega\omega_{B_2} & \omega^2 - \omega_{B_3}^2 \end{pmatrix} \tag{9.23}$$

where $\omega_B = |\boldsymbol{\omega_B}| = \sqrt{\omega_{B_1}^2 + \omega_{B_2}^2 + \omega_{B_3}^2}$. The velocity is solved for as follows

$$\boldsymbol{v} = \frac{e}{m}\Omega^{-1}\boldsymbol{E}. \tag{9.24}$$

The current density defined as $\boldsymbol{j} = -ne\boldsymbol{v}$ is obtained by substituting in the expression for $\boldsymbol{v}$:

$$\boldsymbol{j} = -(ne^2/m)\Omega^{-1}\boldsymbol{E} \tag{9.25}$$

The charge density, ρ, can be derived using the charge conservation equation:

$$\rho = \omega^{-1}\boldsymbol{k}\cdot\boldsymbol{j} = -\frac{ne^2}{m\omega}\,\boldsymbol{k}\cdot\,\Omega^{-1}\boldsymbol{E} \tag{9.26}$$

Using these expressions for $\boldsymbol{j}$ and ρ, Maxwell's equations can be equivalently written in a 'source-free' way:

$$i\boldsymbol{k}\cdot\boldsymbol{\epsilon}\boldsymbol{E} = 0, \qquad i\boldsymbol{k}\cdot\boldsymbol{B} = 0, \tag{9.27}$$

$$i\boldsymbol{k}\times\boldsymbol{E} = i\frac{\omega}{c}\boldsymbol{B}, \qquad i\boldsymbol{k}\times\boldsymbol{B} = -i\frac{\omega}{c}\boldsymbol{\epsilon}\boldsymbol{E}, \tag{9.28}$$

where $\boldsymbol{\epsilon}$ is the di-electric tensor given by

$$\boldsymbol{\epsilon} \equiv I - i\frac{4\pi ne^2}{m\omega}\Omega^{-1} = I - i\frac{\omega_p^2}{\omega}\Omega^{-1}, \tag{9.29}$$

where ω_p is the plasma frequency given by equation (9.15). Define $\omega_B = |\boldsymbol{\omega_B}| = \sqrt{\omega_{B_1}^2 + \omega_{B_2}^2 + \omega_{B_3}^2}$, then $\boldsymbol{\epsilon}$ can be written as

$$\boldsymbol{\epsilon} = I - \frac{\omega_p^2}{\omega^2(\omega^2 - \omega_B^2)} \begin{pmatrix} \omega^2 - \omega_{B_2}^2 & -\omega_{B_1}\omega_{B_2} + i\omega\omega_{B_3} & -\omega_{B_2}\omega_{B_3} - i\omega\omega_{B_1} \\ -\omega_{B_1}\omega_{B_2} - i\omega\omega_{B_3} & \omega^2 - \omega_{B_1}^2 & -\omega_{B_1}\omega_{B_3} + i\omega\omega_{B_2} \\ -\omega_{B_2}\omega_{B_3} + i\omega\omega_{B_1} & -\omega_{B_1}\omega_{B_3} - i\omega\omega_{B_2} & \omega^2 - \omega_{B_3}^2 \end{pmatrix}. \tag{9.30}$$

Combining equations (9.28), we obtain:

$$\boldsymbol{k} \times (\boldsymbol{k} \times \boldsymbol{E}) + \frac{\omega^2}{c^2}\boldsymbol{\epsilon}\boldsymbol{E} = 0, \tag{9.31}$$

which can be rewritten in terms of matrices

$$\begin{pmatrix} -k_2^2 - k_3^2 & k_1 k_2 & k_1 k_3 \\ k_1 k_2 & -k_1^2 - k_3^2 & k_2 k_3 \\ k_1 k_3 & k_2 k_3 & -k_1^2 - k_2^2 \end{pmatrix} \boldsymbol{E} = -\frac{\omega^2}{c^2}\boldsymbol{\epsilon}\boldsymbol{E} \tag{9.32}$$

or more conveniently

$$\left(\varepsilon_{iab}\varepsilon_{bcj}k^a k^c + \frac{\omega^2}{c^2}\boldsymbol{\epsilon}_{ij}\right) E^j = 0 \tag{9.33}$$

where ε_{ijk} is the Levi-Civita symbol.

9.3 Faraday rotation

In the previous section we obtained the wave equation for an anisotropic cold plasma, equation (9.33), which can be seen as a linear system of equations. We can see that, for a given propagating wave vector $\boldsymbol{k}$, these equations allow us to study the polarization of the $\boldsymbol{E}$ field; or, for a given oscillating mode $\boldsymbol{E}$, one can solve for the corresponding dispersion relations. Note that not all $\boldsymbol{k}$ ($\boldsymbol{E}$) have nontrivial solutions $\boldsymbol{E}$ ($\boldsymbol{k}$). The observation is that the linear equation system has nonzero solutions if and only if

$$\left|\boldsymbol{K} + \frac{\omega^2}{c^2}\boldsymbol{\epsilon}\right| = 0, \tag{9.34}$$

where

$$\boldsymbol{K} \equiv \begin{pmatrix} -k_2^2 - k_3^2 & k_1 k_2 & k_1 k_3 \\ k_1 k_2 & -k_1^2 - k_3^2 & k_2 k_3 \\ k_1 k_3 & k_2 k_3 & -k_1^2 - k_2^2 \end{pmatrix}. \tag{9.35}$$

Here we will not brute-force solve for the general solutions of E and the corresponding dispersion relations. We will only discuss a special case where $B_1 = B_2 = 0, B_3 = B$. The di-electric tensor then reduces to

$$\boldsymbol{\epsilon} = I - \frac{\omega_p^2}{\omega^2(\omega^2 - \omega_B^2)} \begin{pmatrix} \omega^2 & i\omega\omega_B & 0 \\ -i\omega\omega_B & \omega^2 & 0 \\ 0 & 0 & \omega^2 - \omega_B^2 \end{pmatrix}. \tag{9.36}$$

From Rybicki and Lightman (2008) we found that the circularly polarized waves,

$$\boldsymbol{E}_{L(R)}(\omega, \boldsymbol{k}) = E\hat{e}_1 \pm iE\hat{e}_2 \tag{9.37}$$

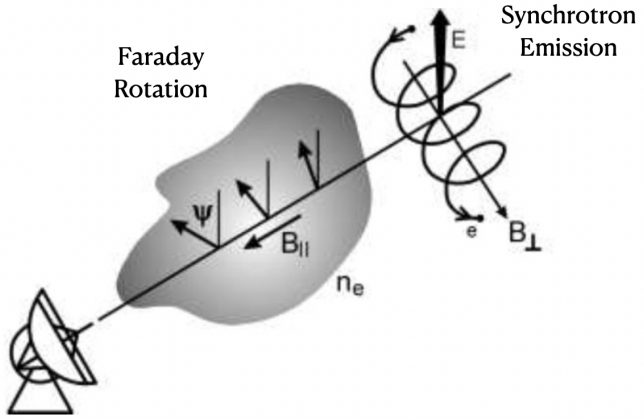

Figure 9.1: Synchrotron emission and Faraday rotation. This is taken from Beck (2016). However, a special clarification is needed here: the linear polarization is a "arrowless vector" and we should not be misled by the arrows in the figure.

where '+' corresponds to left circular polarization and '−' corresponds to right circular polarization, are stable solutions of this special case.

Substituting $\boldsymbol{E}_{L(R)}(\boldsymbol{k}, \omega)$ into equation (9.33), we obtain the dispersion relations for the circularly polarized waves:

$$c^2 k^2 = \omega^2 - \frac{\omega_p^2}{\omega^2 - \omega_B^2} \left(\omega^2 \pm i\omega\omega_B\right). \tag{9.38}$$

In the limit of $\omega \gg \omega_B$ and $\omega \gg \omega_p$, we have

$$\frac{c}{\omega} k_{L,R} \approx 1 - \frac{1}{2}\frac{\omega_p^2}{\omega^2}\left(1 \pm \frac{\omega_B}{\omega}\right). \tag{9.39}$$

This equation implies that left and right circularly polarized waves propagate at different speeds. Thus, as a linear superposition of left and right circularly polarized waves, a linearly polarized electromagnetic wave no longer maintains a fixed plane of polarization; it will rotate as it propagates. This is known as *Faraday rotation* (see Figure 9.1). The phase ϕ of a wave can be calculated as the integral of the spatial frequency k:

$$\phi_{L,R} = \int_0^d k_{L,R}\, \mathrm{d}s, \tag{9.40}$$

where $s = 0$ is the zero phase reference point. The phase difference between the left and right circularly polarized waves is

$$\Delta\phi = \phi_R - \phi_L \approx \int_0^d \frac{\omega_B}{c}\frac{\omega_p^2}{\omega^2}\, \mathrm{d}s. \tag{9.41}$$

The rotation angle of the plane of polarization is given by

$$\Delta\theta = \frac{1}{2}\Delta\phi = \frac{2\pi e^3}{m^2 c^2 \omega^2} \int_0^d nB\, \mathrm{d}s. \tag{9.42}$$

where we have replaced ω_B and ω_p in the second equality.

Chapter 10

Galactic synchrotron emission

10.1 Synchrotron radiation field

We assume a uniform magnetic field and an unprimed reference frame $\mathcal{O}$ that is static to it, and the x^1-axis is aligned with the direction of the magnetic field. Then, the Faraday tensor of this environment magnetic field is constant:

$$F_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -B \\ 0 & 0 & B & 0 \end{pmatrix} \xrightarrow{\dot{\gamma}=0} \begin{pmatrix} \dot{v}_1 \\ \dot{v}_2 \\ \dot{v}_3 \end{pmatrix} = \frac{-eB}{\gamma mc} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \tag{10.1}$$

where the assumption of $\dot{\gamma} = 0$ holds only if the radiation reaction effect is negligible, and we have applied the Lorentz force law and the work equation in terms of the Faraday tensor:

$$\frac{dp_\nu}{dt} = q\frac{\dot{x}^\mu}{c} F_{\mu\nu}. \tag{10.2}$$

Equation (10.1) is a simple linear system of ordinary differential equations, whose solutions are easily found as

$$v_1(t) = v_1^0 \qquad x_1(t) = v_1^0 t \tag{10.3}$$

$$v_2(t) = \cos(\omega_B t)\, v_2^0 - \sin(\omega_B t)\, v_3^0 \qquad x_2(t) = \frac{\sin(\omega_B t)}{\omega_B} v_2^0 - \frac{1-\cos(\omega_B t)}{\omega_B} v_3^0 \tag{10.4}$$

$$v_3(t) = \sin(\omega_B t)\, v_2^0 + \cos(\omega_B t)\, v_3^0 \qquad x_3(t) = \frac{1-\cos(\omega_B t)}{\omega_B} v_2^0 + \frac{\sin(\omega_B t)}{\omega_B} v_3^0 \tag{10.5}$$

where $\boldsymbol{v}(t=0) = (v_1^0, v_2^0, v_3^0)$ is the initial velocity, and the right column gives the world line of the particle. ω_B is the relativistic cyclotron frequency defined by

$$\omega_B \equiv \frac{eB}{\gamma mc}. \tag{10.6}$$

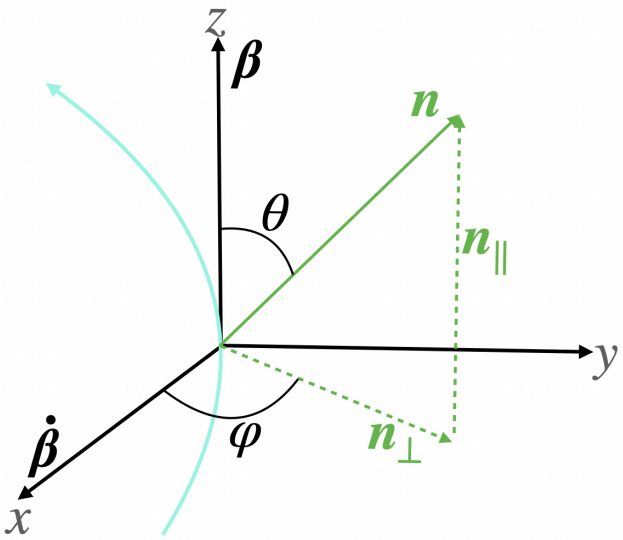

Figure 10.1: Diagram of an electron in synchrotron motion. This is used to help the reader understand the variables in equation (10.7): The origin is the electron, $\boldsymbol{\beta}$ is the velocity vector of the particle, and $\dot{\boldsymbol{\beta}}$ is the acceleration vector of the particle. $\boldsymbol{n}$ is the direction vector of the field point with respect to the particle.

Using equation (8.18), we get the radiation field as follows

$$\boldsymbol{E}(\boldsymbol{x},t) = \frac{e\dot{\beta}}{cR}\left[\frac{\cos\phi(\beta-\cos\theta)\hat{\boldsymbol{\theta}} + \sin\phi(1-\beta\cos\theta)\hat{\boldsymbol{\varphi}}}{(1-\beta\cos\theta)^3}\right] \tag{10.7}$$

where the angular variables are illustrated in Figure 10.1.

A closer look at the square of the magnitude of the directional radiation field,

$$|\boldsymbol{E}|^2 = \frac{e^2\dot{\beta}^2}{c^2R^2}\cdot\left[\frac{\cos^2\phi\,(\beta^2+\cos^2\theta-\beta^2\cos^2\theta-1)+(1-\beta\cos\theta)^2}{(1-\beta\cos\theta)^6}\right], \tag{10.8}$$

suggests that the radiation is strongly peaked in the forward direction ($\theta = 0$, i.e., parallel to the velocity and perpendicular to the acceleration), with an opening angle of about $1/\gamma$.

The total radiated power is given by the relativistic Lamor formula, given by equation (8.25). Since the acceleration has only a transverse component, we have

$$\mathcal{P}'_{tot}(t') = \frac{2\dot{\beta}^2e^2\gamma^4}{3c}. \tag{10.9}$$

From equation (10.1), $\dot{\beta}^2$ can be obtained:

$$\dot{\beta}^2 = \omega_B^2\left(v_2^2+v_3^2\right) = \omega_B^2\beta_\perp^2. \tag{10.10}$$

Substituting the above equation into equation (10.9), we get the total emitted radiation as follows

$$\mathcal{P}'_{tot}(t') = \frac{2e^4B^2\beta_\perp^2\gamma^2}{3m^2c^3}. \tag{10.11}$$

Rybicki and Lightman (2008) provides a detailed discussion on the spectrum of synchrotron radiation, where for the highly relativistic electron, the power per unit

frequency emitted by the electron is decomposed into two polarization states

$$\begin{aligned} P_{\perp}(\omega) &= \frac{\sqrt{3}}{4\pi}\frac{e^3 B\sin\alpha}{mc^2}\left[F\left(\frac{\omega}{\omega_c}\right)+G\left(\frac{\omega}{\omega_c}\right)\right] \\ P_{\parallel}(\omega) &= \frac{\sqrt{3}}{4\pi}\frac{e^3 B\sin\alpha}{mc^2}\left[F\left(\frac{\omega}{\omega_c}\right)-G\left(\frac{\omega}{\omega_c}\right)\right] \\ P(\omega) &= \frac{\sqrt{3}}{2\pi}\frac{e^3 B\sin\alpha}{mc^2}F\left(\frac{\omega}{\omega_c}\right) \end{aligned} \tag{10.12}$$

where $\perp$ and $\parallel$ denote the directions parallel and perpendicular to the projection of the magnetic field on the sky plane, respectively. α is the pitch angle between the field and velocity, and ω_c is the critical frequency that characterizes the cutoff of the broad feature in the spectrum:

$$\omega_c \equiv \frac{3}{2}\gamma^3\omega_B\sin\alpha. \tag{10.13}$$

The function $F(x)$ and $G(x)$ are given in an asymptotic form as follows (Rybicki and Lightman, 2008)

$$G(x)\sim\Gamma\left(\frac{2}{3}\right)\left(\frac{x}{2}\right)^{1/3}, \qquad F(x)\sim\frac{4\pi}{\sqrt{3}\Gamma\left(\frac{1}{3}\right)}\left(\frac{x}{2}\right)^{1/3}, \qquad x\ll 1. \tag{10.14}$$

10.2 Synchrotron radiation from cosmic ray electrons

In the previous section we discussed the synchrotron radiation from a single electron. However, the diffuse galactic synchrotron radiation we observe in astrophysical surveys is the sum of a large number of cosmic ray electrons passing through the interstellar medium (ISM) of the Milky Way Galaxy.

These electrons are one component of the larger population of cosmic rays, which also includes protons, helium nuclei, and other charged particles. A variety of sources, including supernova explosions and other astrophysical phenomena, are thought to produce cosmic ray electrons. These sources accelerate charged particles to very high energies, creating a population of cosmic ray electrons that are injected into the ISM.

One of the most striking features of cosmic ray electrons is their energy spectrum, which follows a power-law distribution:

$$n_{\rm CR}(E)dE\propto E^{-p}dE, \tag{10.15}$$

where p is a constant that characterizes the slope of the distribution. The power law distribution of cosmic ray electrons has been observed to extend over a wide energy range, from MeV to TeV. A typical value of p is 2.4 (Strong et al., 2007).

For a cosmic ray electron ensemble described by the power law distribution, equation

(10.15), we can integrate the equation (10.12) and obtain the ensemble average, which is in a power law over frequency (Rybicki and Lightman, 2008)

$$P_{\text{tot}}(\omega) \propto \left(\frac{m\omega c}{3eB\sin\alpha}\right)^{-(p-1)/2}. \tag{10.16}$$

10.3 Polarization of synchrotron radiation

In this section we simply review the polarization of synchrotron radiation. There are a few points of emphasis in this regard:

1. The radiation from a single charge is elliptically polarized.

   As discussed in section 5.1.1, monochromatic waves are by definition fully polarized. Since equation (10.12) tells us that single-electron synchrotron radiation in general has a non-zero power spectral density in both polarization directions, the monochromatic wave of single-electron synchrotron radiation is elliptically polarized.

2. Single-electron synchrotron radiation field is symmetric with respect to the $\boldsymbol{\beta}\dot{\boldsymbol{\beta}}$ plane, or the xz plane in Figure 10.1.

   Observing the radiation field equation, equation (10.7), we compare the electric fields at two directions:

$$\begin{array}{c} \boldsymbol{n} = (\theta_0, \phi_0) \\ \boldsymbol{E}(\boldsymbol{n}) = E_\theta\hat{\boldsymbol{\theta}}(\boldsymbol{n}) + E_\phi\hat{\boldsymbol{\phi}}(\boldsymbol{n}) \end{array} \longrightarrow \begin{array}{c} \boldsymbol{n}' = (\theta_0, -\phi_0) \\ \boldsymbol{E}(\boldsymbol{n}') = E_\theta\hat{\boldsymbol{\theta}}(\boldsymbol{n}') - E_\phi\hat{\boldsymbol{\phi}}(\boldsymbol{n}') \end{array} \tag{10.17}$$

   These spherical coordinate basis vectors satisfy that

   - $\hat{\boldsymbol{\theta}}(\theta_0, \phi_0)$ and $\hat{\boldsymbol{\theta}}(\theta_0, -\phi_0)$ are symmetric about the xz plane.
   - $\hat{\boldsymbol{\phi}}(\theta_0, \phi_0)$ and $-\hat{\boldsymbol{\phi}}(\theta_0, -\phi_0)$ are also symmetric with respect to the xz plane.

   Thus, $\boldsymbol{E}(\boldsymbol{n})$ and $\boldsymbol{E}(\boldsymbol{n}')$ are symmetric with respect to the xz plane. In other words, using the xz plane as a mirror, the electric fields on either side of the xz plane are mirror images of each other.

3. Ideally, circular polarizations could be canceled out.

   The field symmetry illustrated in the previous point suggests that two observers on opposite sides of the xz plane can observe opposite circular polarizations.

   A similar scenario is when an observer receives radiation beams from many pitch angles at the same time. For some of the radiation, the observer is on the side containing only the left circular polarization, while for the rest, the

observer is on the side containing only the right circular polarization. An ideal situation is one in which all of the left and right circular polarizations cancel each other out. In this case, the observer receives only partially linearly polarized radiation.

4. If the circular polarization is trivial, then the degree of linear polarization can be obtained from equation (10.12):

$$\Pi(\omega) = \frac{P_\perp(\omega) - P_\parallel(\omega)}{P_\perp(\omega) + P_\parallel(\omega)}. \tag{10.18}$$

5. For a power law cosmic ray electron ensemble, equation (10.15), we have

$$\Pi = \frac{p+1}{p+\frac{7}{3}}. \tag{10.19}$$

Chapter 11

Galactic free-free emission

Free-free emission, also known as thermal bremsstrahlung radiation, is another diffuse Galactic emission in the frequency range of significant interest for 21cm statistical and global measurements. In the context of astrophysics, free-free emission occurs when free and thermally hot electrons are accelerated by the Coulomb force of nearby ions, causing them to emit radiation in a wide range of frequencies. Free-free is so named because the initial and final states of the electron are both unbound (Smoot, 1998).

We consider collisions between electrons and ions with charge Ze. In general, in a coordinate system where the plasma is static as a whole, the electrons move much faster than the ions. However, in the electron rest frame, the collision will be viewed as an ion moving at high speed running into an electron. As we will see in the following discussion, since the Coulomb field of a highly relativistic ion can be approximated as a radiation impulse, the free-free process appears to be the scattering of the radiation field off the electron, which is also known as Compton scattering. Therefore, the well studied Compton scattering mechanism can be used to frame the free-free emission of extremely relativistic electrons.

11.1 'Impulse radiation' of uniformly moving ions

The velocity term in equation (8.15) gives us the electric field of a uniformly moving charge:

$$\boldsymbol{E}(\boldsymbol{x},t) = q\left(\frac{\boldsymbol{n}-\boldsymbol{\beta}}{\gamma^2\kappa^3 R^2}\right). \tag{11.1}$$

Without loss of generality, we assume that the charge propagates in the positive direction along the x axis. Then the electromagnetic fields at a distance b from the ion trajectory are are given by equation (11.1) evaluated at

$$t_{ret} = \frac{t - \sqrt{\beta^2 t^2 + (1-\beta^2)(b^2/c^2)}}{1-\beta^2} \tag{11.2}$$

Then the EM fields are

$$E_x = -\frac{qv\gamma t}{(\gamma^2 v^2 t^2 + b^2)^{3/2}} \qquad B_x = 0 \tag{11.3}$$

$$E_y = \frac{q\gamma b}{(\gamma^2 v^2 t^2 + b^2)^{3/2}} \qquad B_y = 0 \tag{11.4}$$

$$E_z = 0 \qquad B_z = \beta E_y \tag{11.5}$$

We see that for a highly relativistic charge, $\beta \approx 1$, we have $\mathbf{E} \approx E_y \hat{\boldsymbol{e}}_y$ and $\mathbf{B} \approx E_y \hat{\boldsymbol{e}}_z$. Thus, the electromagnetic field of the moving charge can be treated as if it were a radiation pulse, whose direction of propagation is parallel to the motion of the charge, and the field is confined in the transverse plane of the trajectory.

The electric field in the frequency domain is given by

$$E(\omega) = \int E_y(t) e^{-i\omega t}\,\mathrm{d}t = \frac{2q}{bv}\frac{b\omega}{\gamma v} K_1\left(\frac{b\omega}{\gamma v}\right) \tag{11.6}$$

The spectrum of the pulse electromagnetic radiation is (Rybicki and Lightman, 2008)

$$\frac{\mathrm{d}W}{\mathrm{d}A\,\mathrm{d}\omega} = \frac{c}{4\pi}|E(\omega)|^2 = \frac{q^2 c}{\pi b^2 v^2}\left(\frac{b\omega}{\gamma v}\right)^2 K_1^2\left(\frac{b\omega}{\gamma v}\right) \tag{11.7}$$

11.2 Free-free processes as Compton scattering

In this section, we discuss the free-free processes in the framework of Compton scattering, where the conservation of momentum and energy is expressed as

$$\lambda_1 - \lambda = \lambda_c(1 - \cos\theta) \tag{11.8}$$

where λ_c is the Compton wavelength given by

$$\lambda_c \equiv \frac{h}{mc}, \tag{11.9}$$

which is 0.02426Å for electrons.

The differential cross section for unpolarized radiation is given by the Klein–Nishina formula

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{r_0^2}{2}\frac{\lambda^2}{\lambda_1^2}\left(\frac{\lambda_1}{\lambda} + \frac{\lambda}{\lambda_1} - \sin^2\theta\right) \tag{11.10}$$

where r_0 is the classical electron radius. Integrating this equation in all directions gives the total cross section (Rybicki et al., 1986):

$$\sigma = \sigma_T \cdot \frac{3}{4}\left[\frac{1+x}{x^3}\left\{\frac{2x(1+x)}{1+2x} - \ln(1+2x)\right\} + \frac{1}{2x}\ln(1+2x) - \frac{1+3x}{(1+2x)^2}\right] \tag{11.11}$$

where $x \equiv h\nu/mc^2$ and $\sigma_T = (8\pi/3)r_0^2$ is the Thomson cross section.

The spectrum of the scattered radiation is given by

$$\frac{\mathrm{d}W_1}{\mathrm{d}\omega} = \sigma \frac{\mathrm{d}W}{\mathrm{d}A\,\mathrm{d}\omega} \tag{11.12}$$

In the lab frame, $\omega' = \gamma\omega(1 + \beta\cos\theta)$, which reduces to $\omega' = \gamma\omega$ when averaged over directions. Since the energy-to-frequency ratio is Lorentz invariant, the power spectral density observed in the lab frame is

$$\frac{\mathrm{d}W_1'}{\mathrm{d}\omega'} = \left.\frac{\mathrm{d}W_1}{\mathrm{d}\omega}\right|_{\omega=\omega'/\gamma} \tag{11.13}$$

The total spectrum for a medium with ion density n_i and electron density n_e at a fixed electron velocity v

$$\frac{\mathrm{d}W_{\mathrm{medium}}}{\mathrm{d}\omega'\,\mathrm{d}V\,\mathrm{d}t} = \int_{b_{min}}^{\infty} (n_e v)\left(n_i \frac{\mathrm{d}W_1'}{\mathrm{d}\omega'}\right) 2\pi b\,\mathrm{d}b \tag{11.14}$$

where $b_{min} = h/mv$ is evaluated using the uncertainty principle $\Delta x \Delta p \geq \hbar$ and taking $\Delta p \sim mv$.

Part IV

Optimal Stokes-I Extraction for Intensity Mapping

Chapter 12

Introduction

For 21cm intensity mapping the Stokes-I signal is what is of interest for doing cosmology (i.e., BAO science (Chang et al., 2008; Villaescusa-Navarro et al., 2017; Wyithe et al., 2008)), whereas the Stokes linearly polarized Q and U components arise from galactic synchrotron emission, which is highly polarized and unfortunately has a complex structure owing to Faraday rotation. So the cosmologists would like to detect solely the 21cm intensity without bothering themselves with looking into messy details of the polarized foregrounds. However, on the technical side, it is hard to construct a feed sensitive to total intensity. Thus, current intensity mapping experiments (Bandura et al. (2014); Newburgh et al. (2016); Chen et al. (2015), etc) are generally using pairs of polarimeters; usually each telescope element uses a pair of feeds sensitive to orthogonal circular or linear polarizations. In other words, polarization measurements have to be made even though only the intensity measurement is desired.

This technical dilemma brings about the challenges in the complex analysis of the galactic polarized emissions. A lot of effort (e.g., Cunnington et al. (2021)) has been devoted to understanding the impact of polarization leakage, while a parallel effort of great interest is to remove polarization from the data to the maximum extent possible. This methodology coincides with possible practical demand in data reduction: for most experiments, especially the interferometers, there are a large number of unnecessary degrees of freedom in the data array which makes the analysis hard to be performed. So an optimal technical route then becomes obvious: one wants to extract the cosmological information in Stokes-I with high fidelity and little contamination from the polarized components.

To achieve such a goal, one might perform data projections as discussed in section 14.2 based on SVD methods, where a linear mapping is decomposed into the product of an orthogonal matrix, a diagonal rectangular matrix, and another orthogonal matrix. Such a technique has been employed in (Shaw et al., 2014), although they only defined the inner product structure in the data space while a discussion of the appropriate

choice of metric on the linear space of sky signals is lacking.[1]

In this chapter, we suggest that what one wants to do is more complicated than a simple projection. One must rather balance two conflicting objectives. On the one hand, one wants to completely remove any admixture of the polarization by excluding linear combinations of data that include polarization. On the other hand, one wants to combine many measurements in order to minimize the total noise in the final temperature sky map—or here temperature dominated sky modes, in terms of the formalism used in the Shaw papers (Shaw et al., 2014, 2015).

In order to analyze how to achieve this balance optimally, additional information is needed. More specifically, one needs to know what is the expectation, or prior, on the Stokes I sky and the Stokes Q/U sky. These are expressed in terms of *a priori* power spectra. In this part, we sketch how to find the appropriate temperature, or Stokes-I, dominated modes. We demonstrate that this is a matter of integrating out the polarization.

[1] Or say, the implied prior is a white noise prior of the same amplitude for the temperature and polarization sky signals.

Chapter 13

Linear measurements

13.1 Linear map between linear spaces

In the previous chapter, we have formalised the measurement equation as an integral over a sphere, or as the sum of infinite-dimensional sequences over a set of basis functions. While in practical data analysis, at the resolution limit of the instrument, we can coarse-grain the integral over the continuous directional variable to obtain a discrete measurement equation over a finite number of pixels. As to a measurement equation expressed with a set of basis functions, the instrumental resolution works as a cutoff, which also makes the measurement equation a summation of finite terms. Therefore, in the absence of noise or measurement error, a single measurement can be formulated as the following *discrete measurement equation*:

$$d(\nu) = \sum_{i=1}^{N_s} B_i(\nu) s_i(\nu) \tag{13.1}$$

where d is the single measurement result at frequency ν. B and s are respectively the discretized antenna beam and the discretized sky, with i indexing all the "degrees of freedom" (DoF). As an example, if one uses the Healpix scheme for discretization, then a sky DoF represents the sky intensity for a specific polarization and pixel; the total number of Dof of the radio sky $s(\nu)$ is therefore $N_s = N_{pix} \times 4$; Correspondingly, the antenna beam for this measurement also has N_s DoF. A single measurement can therefore be understood as the sum over all sky degrees of freedom, each weighted by the corresponding beam degrees of freedom. For convenience, we call a tuple of all sky degrees of freedom a *sky vector*. Similarly, a *beam vector* is naturally defined in a measurement.

If we assume that the measured polarized radio sky is the same throughout the data set, i.e. the sky vector remains constant in the observation, then we can obtain a system of linear equations by listing the measurement equations for all measurements

under the same discretization regime:

$$\boldsymbol{d}_{(N_d)} = \boldsymbol{B}_{(N_d \times N_s)} \boldsymbol{s}_{(N_s)}. \tag{13.2}$$

Here, the data set of N_d measurements is represented as an N_d-dimensional (column) vector; because each measurement has its own beam vector, the antenna responses of all measurements are grouped as a *beam matrix* of N_d rows, where each row vector is the N_s-dimensional beam vector for the corresponding measurement. Equation(13.2) suggests the data vector can be written as the multiplication between the beam matrix and the sky vector.

In abstract terms, the discretized measurement equations we established above can be viewed as a linear mapping from one linear space to another. Specifically, we may define the N_s-dimensional linear space where the sky vector resides as the *sky space*, which is denoted by $\mathcal{S}$. Similarly, we define $\mathcal{D}$ as the *data space* (or sometimes called the *telescope space* or *visibility space*), from which the data vector is drawn. Then, the linear mapping formalism of the measurements reads

$$\mathcal{B} : \mathcal{S} \to \mathcal{D}, \tag{13.3}$$

which provides an equivalent description of equation (13.2).

In intensity mapping, it is customary to partition the sky degrees of freedom into its Stokes I and polarization components. Consequently, the measurement can be conceptualized as the aggregate of two linear mappings. Differentiating between the temperature of the sky and its polarization, the data can be represented by the relationship:

$$\boldsymbol{d} = \mathbf{A}\boldsymbol{i} + \mathbf{B}\boldsymbol{p} + \boldsymbol{n} \tag{13.4}$$

Here, $\boldsymbol{i}$ and $\boldsymbol{p}$ denote sky vectors of unpolarized and polarized degrees of freedom, respectively. The matrices $\mathbf{A}$ and $\mathbf{B}$ correspond to the respective beam matrices.

13.2 SVD

13.2.1 Singular Value Decomposition

A vector in a linear space can be decomposed on any complete orthonormal basis of the linear space. Assuming that we have found such a set of basis in the sky space $\mathcal{S}$, which consists of N_s orthonormal vectors, denoted as $\{\hat{\boldsymbol{v}}_1, \ldots, \hat{\boldsymbol{v}}_{N_s}\}$, then a sky vector $\boldsymbol{s}$ can be expressed as

$$\boldsymbol{s} = s_1 \hat{\boldsymbol{v}}_1 + \cdots + s_{N_s} \hat{\boldsymbol{v}}_{N_s} \tag{13.5}$$

where s_i is the projection of $\boldsymbol{s}$ onto the basis vector, i.e., $s_i = \hat{\boldsymbol{v}}_i^T \boldsymbol{s}$. Such a decomposition can also be expressed in a compact form,

$$\boldsymbol{s} = \begin{pmatrix} \hat{\boldsymbol{v}}_1 & \dots & \hat{\boldsymbol{v}}_{N_s} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{v}}_1^T \boldsymbol{s} \\ \vdots \\ \hat{\boldsymbol{v}}_{N_s}^T \boldsymbol{s} \end{pmatrix} = \mathbf{V}\mathbf{V}^{\mathbf{T}} \boldsymbol{s} \tag{13.6}$$

where the whole basis set is represented by an N_s by N_s square orthogonal matrix, $\mathbf{V}$, whose i-th column vector is defined to be the basis vector $\hat{\boldsymbol{v}}_i$.

Similarly, the data vector $\boldsymbol{d}$ in the data space $\mathcal{D}$ can also be linearly decomposed into

$$\boldsymbol{d} = \mathbf{U}\mathbf{U}^{\mathbf{T}} \boldsymbol{d} \tag{13.7}$$

where $\mathbf{U}$ is an N_d by N_d orthogonal matrix, and its column vectors $\{\hat{\boldsymbol{u}}_1, \dots, \hat{\boldsymbol{u}}_{N_d}\}$ can form a complete orthonormal basis for the data space $\mathcal{D}$.

The selections of the bases of the linear spaces $\mathcal{S}$ and $\mathcal{D}$ are arbitrary, or in other words, one could have infinite choices to construct the orthogonal matrices, $\mathbf{V}$ and $\mathbf{U}$. However, for the linear map $\mathcal{B}(s)$ characterized by the matrix $\boldsymbol{B}$, there is one special pair of $\mathbf{V}$ and $\mathbf{U}$, which satisfies that

$$\boxed{\boldsymbol{B}\hat{\boldsymbol{v}}_1 = \sigma_1 \hat{\boldsymbol{u}}_1 \quad \cdots \quad \boldsymbol{B}\hat{\boldsymbol{v}}_r = \sigma_r \hat{\boldsymbol{u}}_r} \qquad \boxed{\boldsymbol{B}\hat{\boldsymbol{v}}_{r+1} = 0 \quad \cdots \quad \boldsymbol{B}\hat{\boldsymbol{v}}_{N_s} = 0} \tag{13.8}$$

where the $\hat{\boldsymbol{v}}$'s are named as the *right singular vectors* and the $\hat{\boldsymbol{u}}$'s are the *left singular vectors*. The number r is the *rank* of $\boldsymbol{B}$, and the r numbers, σ's, are the so called *singular values*, which are usually arranged as positive numbers in descending order $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. These singular vectors and singular values are defined from the *Singular Value Decomposition* (SVD) of the beam matrix $\boldsymbol{B}$,

$$\boldsymbol{B} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathbf{T}} = \sigma_1 \hat{\boldsymbol{u}}_1 \hat{\boldsymbol{v}}_1^T + \cdots + \sigma_r \hat{\boldsymbol{u}}_r \hat{\boldsymbol{v}}_r^T . \tag{13.9}$$

where $\boldsymbol{\Sigma}$ is an N_d by N_s matrix

$$\boldsymbol{\Sigma} \equiv \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right) . \tag{13.10}$$

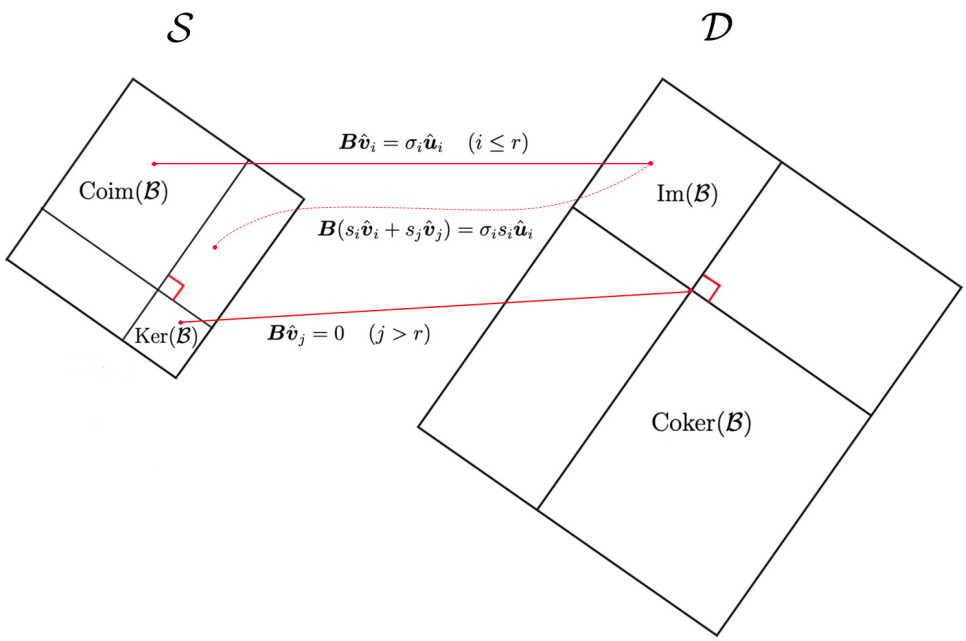

Figure 13.1: Four fundamental subspaces of the linear map.

13.2.2 SVD viewpoint of the linear map

Now we can use the SVD of $\boldsymbol{B}$ to understand the linear map in equation (13.3). It provides with us an intuitive understanding of the measurement equation:

$$\begin{aligned} \boldsymbol{B}\boldsymbol{s} &= \boldsymbol{B}(s_1\hat{\boldsymbol{v}}_1 + \cdots + s_r\hat{\boldsymbol{v}}_r + s_{r+1}\hat{\boldsymbol{v}}_{r+1} + \cdots s_{N_s}\hat{\boldsymbol{v}}_{N_s}) \\ &= \sigma_1 s_1 \hat{\boldsymbol{u}}_1 + \cdots + \sigma_r s_r \hat{\boldsymbol{u}}_r \end{aligned} \tag{13.11}$$

where you can see that the sky vector $\boldsymbol{s}$ is projected in N_s linearly independent directions. For $i \le r$, $\mathcal{B}$ maps the $\mathcal{S}$ space direction, $\hat{\boldsymbol{v}}_i$, to the $\mathcal{D}$ space direction, $\hat{\boldsymbol{u}}_i$, with the component value being scaled by the factor σ_i. The sky components in the remaining $N_s - r$ directions contribute nothing to the measurement, or linear map.

In summary, the mapping of the sky vector is effectively the sum of the mappings of all the sky components. The sensitivity or "gain" of the measurement process to the sky component at $\hat{\boldsymbol{v}}_i$ is characterized by σ_i. For $i > r$, the gains are simply zero.

13.2.3 Four fundamental subspaces

For a better grasp of the properties of the linear map, we define the four fundamental subspaces (see Figure 13.1), namely, *image*, *kernel*, *cockernel* and *coimage*, where kernel and coimage are subspaces of $\mathcal{S}$ and image and cokernel are subspaces of $\mathcal{D}$. Specifically, the *image* of the map is defined as the set $\{\mathcal{B}(\boldsymbol{s})|\boldsymbol{s} \in \mathcal{S}\}$, which is denoted as Im($\mathcal{B}$). We define the *kernel* (or *null space*) of $\mathcal{B}$, Ker($\mathcal{B}$), as the set $\{\boldsymbol{s} \in \mathcal{S}|\mathcal{B}(\boldsymbol{s}) = 0\}$.

In the data space, it would also be useful to look at the linear subspace orthogonal

to the image, which is called the *cokernel* denoted as Coker($\mathcal{B}$). A direct understanding of the cokernel is that, without considering any system noise or random error of the instrument, the sky does not have any mapping into the cokernel. In the sky space, we define the linear subspace orthogonal to the kernel as the *coimgage*, denoted as Coim($\mathcal{B}$). It can be seen that the sky vector can be divided into two parts, each residing in two linear subspaces of S, kernel and coim. Among them, the contribution of the sky components in the kernel to the mapping (or measurement results) is 0, while the sky components in the coim have a non-trivial contribution to the measurement.

More intuitively, the four fundamental subspaces can be spanned by the singular vectors of $\boldsymbol{B}$:

$$\begin{aligned}
\text{Ker}(\mathcal{B}) &= \text{span}\{\hat{\boldsymbol{v}}_{r+1}, \cdots, \hat{\boldsymbol{v}}_{N_s}\} \\
\text{Im}(\mathcal{B}) &= \text{span}\{\hat{\boldsymbol{u}}_1, \cdots, \hat{\boldsymbol{u}}_r\} \\
\text{Coker}(\mathcal{B}) &= \text{span}\{\hat{\boldsymbol{u}}_{r+1}, \cdots, \hat{\boldsymbol{u}}_{N_d}\} \\
\text{Coim}(\mathcal{B}) &= \text{span}\{\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_r\}
\end{aligned}$$

In addition to the above representations, it is easy to see that Im($\mathcal{B}$) can also be expressed as the space spanned by the columns of $\boldsymbol{B}$, i.e., Im($\mathcal{B}$) = C($\boldsymbol{B}$). Also, the Coim($\mathcal{B}$) can be spanned by the row vectors of B, which means Coim($\mathcal{B}$) = C($\boldsymbol{B^T}$).

13.3 Linear map between inner product spaces

In the context of statistical measurements, the radio sky can be viewed as an instance drawn from a random field. Different components of the sky might be characterized by different statistical models: the foreground components are usually given by empirical models whose parameters are to be calibrated by observation, while the cosmological component may come from theoretical models.

Taking the statistical models of the radio sky as the presumptions, we can predict the 2pt correlation between any pair of sky degrees of freedom. A complete prediction about the whole discrete radio sky is established by calculating the correlations of all pairs of sky degrees of freedom, which can be organized as a covariance matrix, $\boldsymbol{C} = \langle \boldsymbol{s}\boldsymbol{s}^T \rangle$, so that C_{ij} represents the correlation between the i-th and j-th sky degrees of freedom. A bit of jargon in statistics: we shall also call $\boldsymbol{C}$ the *prior* of the radio sky.

With the prior of the sky in hand, what can we do in data analysis? First of all, continuing our linear map theory, we can try to characterize the "signal strength" of the sky vector in different directions for this prior. We can characterize the expectation value of the projection of the sky vector into a certain direction in $\mathcal{S}$, say $\hat{\boldsymbol{v}}$, then the

square of the magnitude of $\boldsymbol{s}$ in this direction is [1]:

$$\langle(\hat{\boldsymbol{v}}^T\boldsymbol{s})^2\rangle = \langle\hat{\boldsymbol{v}}^T\boldsymbol{s}\boldsymbol{s}^T\hat{\boldsymbol{v}}\rangle = \hat{\boldsymbol{v}}^T\langle\boldsymbol{s}\boldsymbol{s}^T\rangle\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}^T\boldsymbol{C}\hat{\boldsymbol{v}} = C_{ij}v^iv^j \tag{13.12}$$

The above equations can be effectively understood as the definition of the *inner product structure* for the sky space $\mathcal{S}$, with the metric being C_{ij}. For example, the squared vector modulus

$$C(\boldsymbol{v},\boldsymbol{v}) \equiv C_{ij}v^iv^j \tag{13.13}$$

Then the squared vector modulus defined with C_{ij}, denoted as $C(\boldsymbol{v},\boldsymbol{v})$, characterizes the *prior signal strength* in the direction along $\boldsymbol{v}$.

On the other hand, the instrumental noise is unavoidable for any practical measurement. Taking noise terms into account, the discrete measurement equation, i.e., equation (13.2), can be rewritten as

$$\boldsymbol{d}_{(N_d)} = \boldsymbol{B}_{(N_d\times N_s)}\boldsymbol{s}_{(N_s)} + \boldsymbol{n}_{(N_d)}, \tag{13.14}$$

where $\boldsymbol{n}$ is the noise vector grouping the noise terms for all measurements. $\mathbf{N} \equiv \langle\boldsymbol{n}\boldsymbol{n}^T\rangle$ is defined as the covariance matrix of $\boldsymbol{n}$. Just like the role of $\mathbf{C}$ in $\mathcal{S}$, $\mathbf{N}$ also defines an inner product structure for the data space $\mathcal{D}$ so that the *total noise scale* of the data set in the direction, $\hat{\boldsymbol{u}}$, can be defined as

$$\langle(\hat{\boldsymbol{u}}^T\boldsymbol{n})^2\rangle = \langle\hat{\boldsymbol{u}}^T\boldsymbol{n}\boldsymbol{n}^T\hat{\boldsymbol{u}}\rangle = \hat{\boldsymbol{u}}^T\langle\boldsymbol{n}\boldsymbol{n}^T\rangle\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}^T\mathbf{N}\hat{\boldsymbol{u}} = N_{ij}u^iu^j \tag{13.15}$$

With the definitions of metrics C_{ij} and N_{ij}, now both linear spaces $\mathcal{S}$ and $\mathcal{D}$ are endowed with the inner product structure. In this scenario, the discrete measurement equation can be expressed as a map of two inner product spaces:

$$\mathcal{B} : (\mathcal{S}, C_{ij}) \to (\mathcal{D}, N_{ij}) \tag{13.16}$$

Taking the SVD viewpoint discussed in Section 13.2.2, we can build a fuller understanding of the linear map. Having decomposed the sky vector $\boldsymbol{s}$ using the right singular vectors, for each component we have:

1. The sky component along $\hat{\boldsymbol{v}}_i$ has a priori signal strength of $C(\hat{\boldsymbol{v}}_i, \hat{\boldsymbol{v}}_i)$.

2. The measurement process amplifies the signal by σ_i and it becomes the $\hat{\boldsymbol{u}}_i$ data component in the data space. (For directions with $i > r$, σ_i is 0.)

3. However, the random noise also contributes to the data in the $\hat{\boldsymbol{u}}_i$ direction; the total noise power can be described as $N(\hat{\boldsymbol{u}}_i, \hat{\boldsymbol{u}}_i)$

[1] Here we use the vector multiplication of general linear algebra to calculate the projection, that is, the ordinary metric is used by default.

Hence, We can roughly establish a definition of signal-to-noise ratio for the mapping at each single direction

$$\mathrm{SNR}[i] = \frac{\sigma_i \, C(\hat{\boldsymbol{v}}_i, \hat{\boldsymbol{v}}_i)}{N(\hat{\boldsymbol{u}}_i, \hat{\boldsymbol{u}}_i)} \tag{13.17}$$

It should be noted that, in order to establish the theoretical framework, we simply regarded all components of the radio sky as "signals" and only the instrumental noise as "noise". However, the definition of "signal" and "noise" depends on the specific scientific goals. For example, you may want to extract cosmological signals and the foreground should be part of the noise, for which the signal-to-noise radio might be defined as

$$\mathrm{SNR}[i] = \frac{\sigma_i \, C_{\cos}(\hat{\boldsymbol{v}}_i, \hat{\boldsymbol{v}}_i)}{N(\hat{\boldsymbol{u}}_i, \hat{\boldsymbol{u}}_i) + \sigma_i \, C_{\mathrm{fg}}(\hat{\boldsymbol{v}}_i, \hat{\boldsymbol{v}}_i)}. \tag{13.18}$$

Chapter 14

Linear data projections

In Chapter 13, we established a formalism for measurement. We discussed how to decompose the linear map within the SVD. Furthermore, we defined the four fundamental subspaces, each spanned by a set of singular vectors. We also proposed that, after introducing the prior of the sky and the noise power spectrum, one can characterize the signal strength and the noise scale for each "submap."

Given the linear formalism of the measurement, we naturally want to apply it to data projection, that is, to obtain the part of the data vector in some linear subspace of $\mathcal{D}$. From the perspective of expected effects, our goals can be divided into the following two types:

- *Data reduction* Ordinarily, there are more data points than sky degrees of freedom. From the perspective of discrete linear measurement process, the number of measurements must be greater than the number of sky DoF after discretization, and the number of sky DoF must also be greater than the needs of scientific extraction. As a result, raw data sets are often large, contain many redundant data points, and are dominated by instrument noise. Occasionally, managing the entire dataset is straightforward. However, there are instances when the volume of data becomes overwhelming. In such cases, an ideal solution would be to compress the raw data without losing any information—this is what we refer to as "data reduction." In Section 14.1, we look at achieving this by using linear data projectors.

- *Selection or avoidance of data components*

  Sometimes we want to eliminate some data components because their complex properties pose challenges for science extraction. For example, there might be some beam sidelobes which are intractable, or certain sky DoF (pixels, multipoles, or polarizations) which are incomprehensible. When we judge that removing them will do more good than harm, we might choose to remove (or at least down-weight) these components. This strategy is essentially different from the

lossless data reduction discussed above, as it really discards some sky information. So the signal-to-noise ratio in the overall data set is not guaranteed to rise or fall. Data projection for this purpose is discussed in Section 14.2, where we also give a toy model to the increased noise if one rejects all polarizations.

14.1 Data reduction

14.1.1 Noisy data reduction

In the previous discussion, we decomposed the sky vector in N_s directions with the help of the SVD of the beam matrix. Among all the components of the sky, only the components at r directions, which are $\hat{\boldsymbol{v}}_1, \ldots, \hat{\boldsymbol{v}}_r$, can have non-trivial mappings, which reside in the subspace of $\mathcal{D}$, $\mathrm{Im}(\mathcal{B})$, and can be completely described by the components at directions $\hat{\boldsymbol{u}}_1, \ldots, \hat{\boldsymbol{u}}_r$ in $\mathcal{D}$.

However, in addition to the mappings of the sky vector, the real data also includes the contribution of instrumental noise. Because of the random nature of the instumental noise, the noise vector may have nontrivial components in all N_d directions in $\mathcal{D}$. In other words, the data vector $\boldsymbol{d}$ with noise may appear in all the directions; $\boldsymbol{d}$-components in $\mathrm{Im}(\mathcal{B})$ may contain both sky and noise, while the components at the remaining $(N_d - r)$ directions are completely from the noise.

Therefore, a lossless data compression scheme is obviously at hand. By projecting the data vector onto $\mathrm{Im}(\mathcal{B})$ (Shaw et al., 2015), all the sky information is preserved, while what is lost is purely noise. This can be done mathematically by a projection operation

$$\boldsymbol{d}' = \mathbf{U}^T \boldsymbol{d} \tag{14.1}$$

where $\mathbf{U}$ is the data projection matrix, whose columns are the left singular vectors, $\hat{\boldsymbol{u}}_1, \ldots, \hat{\boldsymbol{u}}_r$.

This is a rather direct way of reducing a larger data set to a smaller, more manageable size. It can be understood as a number of linear recombinations of the raw data, and one can find the projector simply by writing out the column vectors of the beam matrix and computing all the vectors perpendicular to the column vectors. This method has nothing to do with the sky or noise modelling, but its losslessness depends on the accuracy of the beam matrix.

14.1.2 Noisy data reduction with priors

The data compression method discussed in the previous section only involves the concept of fundamental linear subspace. However, our linear mapping toolbox allows us to reduce the data set even more, if one needs so.

First, we can make use of the singular values, σ_i's. Because σ_i describes the measurement gain of the sky component in the $\hat{\boldsymbol{v}}_i$ direction, we can set a threshold for the singular values to filter out those components whose measurement gains are numerically zero.

Beyond the use of singular values as measurement gains, an example of a more accurate signal-to-noise description is given in equation (13.17). By introducing the covariance matrices of sky and noise, we can calculate the SNR in all r directions in $\mathrm{Im}(\mathcal{B})$ and abandon the components with way too low SNRs.

A similar idea has been used in Shaw et al. (2014, 2015). Specifically, they first whiten the data space with the noise power spectrum

$$\boldsymbol{d'} = \mathbf{N}^{-\frac{1}{2}}\boldsymbol{d}, \tag{14.2}$$

which also gives a new beam matrix, $\boldsymbol{B'} = \mathbf{N}^{-\frac{1}{2}}\boldsymbol{B}$. Then the data is compressed by projecting $\boldsymbol{d'}$ onto $\mathrm{Im}(\mathcal{B}')$. Within the linear map framework, this data projection scheme could be regarded as a special case of the data projection, with an inner product structure allocated to the data space while assuming a trivial metric for the sky space.

14.2 Selection or avoidance of data components

Sometimes, we may only wish to use only a part of the sky or instrument DOF, or try to avoid some incomprehensible DOF. Depending on the specific needs, we might want to divide all the column vectors of the beam matrix, $\boldsymbol{B}$, into two parts, the part that is needed, denoted as $\boldsymbol{B_g}$, and the part that needs to be avoided, denoted as $\boldsymbol{B_b}$. Corresponding to the column vectors, the sky degrees of freedom are also grouped into two vectors, $\boldsymbol{s_g}$ and $\boldsymbol{s_b}$. Then the measurement process reads

$$\boldsymbol{d} = \boldsymbol{B_g}\boldsymbol{s_g} + \boldsymbol{B_b}\boldsymbol{s_b}. \tag{14.3}$$

In abstract terms, the linear map is decomposed such that $\mathcal{B} = \mathcal{B}_g + \mathcal{B}_b$, where

$$\mathcal{B}_g : \mathcal{S}_g \to \mathcal{D} \tag{14.4}$$

$$\mathcal{B}_b : \mathcal{S}_b \to \mathcal{D} \tag{14.5}$$

Below we discuss two hypothetical situations to explain possible usage scenarios:

- **Case 1: Polarization Rejection**

  Some beam or sky DoF may have too complex behaviors that might introduce difficulties in specific data analysis. For example, one may want to avoid components from the complex polarization foreground in the data, so these unwanted sky DoF are grouped as $\boldsymbol{s_b}$. Then $\boldsymbol{s_g}$, $\boldsymbol{B_b}$, and $\boldsymbol{B_g}$ are also specified

at the same time. The data operation that can reject the polarization is to project data onto the subspace, $\mathrm{Im}(\mathcal{B}_g) \bigcap \mathrm{Coker}(\mathcal{B}_b)$.

- **Case 2: De-risking Problematic Beam Modes**

One may have a high uncertainty in the simulation or measurement of specific beam components, which can be understood as that some column vectors of the beam matrix are simulated or measured with a large uncertainty, which defines the $\mathcal{B}_b$. In this situation, it is no longer suitable for us to project data into the cokernel of $\mathcal{B}_b$, since it could deviate from the real beam a lot. But instead, one can project the data onto the image of $\mathcal{B}_g$, which keeps the whole mappings of $\mathcal{B}_g$ but somewhat down-weights the mappings of $\mathcal{B}_b$ too.

We should note the essential difference between the two situations. In case one, we rely on the beam matrix to remove some DoF to the greatest extent. However, in case 2, some modes of the beam matrix themselves might have large uncertainties, and we hope to down-weight the data components from these uncertain beam modes.

14.3 Estimator Noise: A Worked Example

In the previous discussion we explained how to use data projection to circumvent or down weight some data modes. But there are trade-offs to doing so. Whether it is necessary to completely remove certain degrees of freedom requires a specific pros and cons analysis, and there is no unified answer. In this section, we discuss a simple toy model to illustrate the effect that polarization rejection may have on the noise level of the sky map estimator.

Rejecting polarization does free the cosmologists from analyzing complex structures of the galactic polarized emissions. However it also comes at the cost of throwing away relevant information that would allow for a better reconstruction of the temperature sky.

In this section, we give an example, for which we calculate the variance of the temperature determination using the data that remains after projection onto $\mathrm{Coker}(\mathcal{B}_\mathcal{P})$. In this example we see that the resulting temperature determination has increased variance. The relative increase is determined by the expected ratio of the polarization signal to the expected instrumental noise. This example is contrived to be simple and illustrate certain points and do not address how the comparison would work for the actual data.

Suppose that

$$\mathbf{B}_{4\times 4} = \left(\begin{array}{c|ccc} \mathbf{B}_T & \mathbf{B}_P \end{array} \right) = \left(\begin{array}{c|ccc} t_0 & 0 & 0 & 0 \\ t_1 & \epsilon_{p1} & 0 & 0 \\ t_2 & 0 & \epsilon_{p2} & 0 \\ t_3 & 0 & 0 & \epsilon_{p3} \end{array} \right) \tag{14.6}$$

and $\mathbf{s} = \left(\begin{array}{c|ccc} \mathbf{s}_T & \mathbf{s}_P \end{array} \right)^{\mathbf{T}} = \left(\begin{array}{c|ccc} s_t & s_{p1}, & s_{p2}, & s_{p3} \end{array} \right)^{\mathbf{T}}$. The measurement then reads

$$\mathbf{v} = \left(\begin{array}{c|ccc} t_0 & 0 & 0 & 0 \\ t_1 & \epsilon_{p1} & 0 & 0 \\ t_2 & 0 & \epsilon_{p2} & 0 \\ t_3 & 0 & 0 & \epsilon_{p3} \end{array} \right) \left(\begin{array}{c} s_t \\ s_{p1} \\ s_{p2} \\ s_{p3} \end{array} \right) + \left(\begin{array}{c} n_0 \\ n_1 \\ n_2 \\ n_3 \end{array} \right)$$

We assume that the noise has been whitened. The three columns of $\mathbf{B}_P$ are orthogonal and thus form a natural orthonormal basis for range($\mathbf{B}_P$), i.e. the column space of $\mathbf{B}_P$. As the space orthogonal to range($\mathbf{B}_P$), cokernel($\mathbf{B}_P$) is spanned by the only left natural basis vector for $\mathbb{R}^4$, i.e., span$\{(1,0,0,0)^{\mathbf{T}}\}$. The projection operator (on the linear space of visibilities) is $\mathbf{P}_{P-free} = \text{diag}(1,0,0,0)$.

After the cokernel($\mathbf{B}_P$) projection, we are left with only the one equation

$$v_1 = t_0 s_t + n_0. \tag{14.7}$$

The minimum variance estimation $\hat{s}_t$ and the corresponding variance for this equation is

$$\hat{s}_t^{P-free} = \frac{v_1}{t_0}, \quad \sigma^{2,P-free} = \frac{1}{t_0^2}. \tag{14.8}$$

If we do not project onto the cokernel, we have the following equations

$$\left(\begin{array}{c} t_0 s_t + n_0 \\ t_1 s_t + \epsilon_{p1} s_{p1} + n_1 \\ t_2 s_t + \epsilon_{p2} s_{p2} + n_2 \\ t_3 s_t + \epsilon_{p3} s_{p3} + n_3 \end{array} \right) = \left(\begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \right). \tag{14.9}$$

For this simple example, we obtain

$$\frac{1}{\sigma^2(\hat{s}_{ti})} = \frac{t_i^2}{\epsilon_{pi}^2 \left\langle s_p^2 \right\rangle + 1}.$$

Since (by construction) the polarization and other noise of the estimators $\hat{s}_{ti}$ are statistically independent, the minimum variance unbiased estimator constructed from a

linear combination of these has a variance obeying

$$\frac{1}{\sigma^2(\hat{s}_p)} = \sum_i \frac{t_i^2}{\epsilon_{pi}^2 \left\langle s_p^2 \right\rangle + 1}$$

showing that in certain situations a much better estimator can be obtained by allowing for some polarization leakage.

Chapter 15

Optimal Stokes-I extraction for 21 cm intensity mapping

In this chapter, we present an optimal method for Stokes-I data extraction in the presence of polarization. Separating the temperature sky and the polarized sky, the data can be expressed as the process

$$\boldsymbol{d} = \mathbf{A}\boldsymbol{i} + \mathbf{B}\boldsymbol{p} + \boldsymbol{n} \tag{15.1}$$

where $\boldsymbol{i}$ and $\boldsymbol{p}$ are respectively the sky vectors composed of unpolarized and polarized degrees of freedom, and $\mathbf{A}$ and $\mathbf{B}$ are corresponding beam matrices. $\boldsymbol{n}$ denotes the instrumental noise, whose power spectrum is given as $\mathbf{N}$. We shall assume negligible correlations between polarized and non-polarized degrees of freedom so that a block diagonal covariance matrix for the sky is implied as

$$\mathbf{C}_S = \mathrm{diag}(\mathbf{C}_I, \mathbf{C}_P). \tag{15.2}$$

We further postulate a Gaussian prior of the form

$$p_{\mathrm{prior}} \sim \exp\left[-\frac{1}{2}\boldsymbol{p}^T\mathbf{C}_P^{-1}\boldsymbol{p} - \frac{1}{2}\boldsymbol{i}^T\mathbf{C}_I^{-1}\boldsymbol{i}\right]. \tag{15.3}$$

The *posterior*, with both the temperature and polarization prior included, is of the form

$$p_{\mathrm{posterior}} \sim \exp\left[-\frac{1}{2}\chi^2(\boldsymbol{i}, \boldsymbol{p}; \boldsymbol{d})\right] \tag{15.4}$$

where

$$\chi^2 = (\boldsymbol{d} - \mathbf{A}\boldsymbol{i} - \mathbf{B}\boldsymbol{p})^T\mathbf{N}^{-1}(\boldsymbol{d} - \mathbf{A}\boldsymbol{i} - \mathbf{B}\boldsymbol{p}) + \boldsymbol{p}^T\mathbf{C}_P^{-1}\boldsymbol{p} + \boldsymbol{i}^T\mathbf{C}_I^{-1}\boldsymbol{i}. \tag{15.5}$$

To integrate out the polarization $\boldsymbol{p}$, which is not of interest for our purpose, we

collect all the terms linear or quadratic in $\boldsymbol{p}$ in the above, obtaining

$$\boldsymbol{p}^T(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})\boldsymbol{p}-(\boldsymbol{d}-\mathbf{A}\boldsymbol{i})^T\mathbf{N}^{-1}\mathbf{B}\boldsymbol{p}-\boldsymbol{p}^T\mathbf{B}^T\mathbf{N}^{-1}(\boldsymbol{d}-\mathbf{A}\boldsymbol{i}), \tag{15.6}$$

which has the form

$$\boldsymbol{p}^T\mathcal{A}\boldsymbol{p}+\boldsymbol{q}^T\boldsymbol{p}+\boldsymbol{p}^T\boldsymbol{q}=(\boldsymbol{p}+\mathcal{A}^{-1}\boldsymbol{q})^T\mathcal{A}(\boldsymbol{p}+\mathcal{A}^{-1}\boldsymbol{q})-\boldsymbol{q}^T\mathcal{A}^{-1}\boldsymbol{q}. \tag{15.7}$$

Upon integrating over $\boldsymbol{p}$, the first term becomes a constant factor with no dependence on $\boldsymbol{i}$, leaving a factor proportional to

$$-\boldsymbol{q}^T\mathcal{A}^{-1}\boldsymbol{q} \tag{15.8}$$

which we may expand into

$$-(\boldsymbol{d}-\mathbf{A}\boldsymbol{i})^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}(\boldsymbol{d}-\mathbf{A}\boldsymbol{i}) \tag{15.9}$$

and then into terms quadratic and linear in $\boldsymbol{i}$, obtaining

$$\begin{aligned}&\boldsymbol{i}^T\left[-\mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\mathbf{A}\right]\boldsymbol{i}\\&+\boldsymbol{d}^T\left[\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\mathbf{A}\right]\boldsymbol{i}\\&+\boldsymbol{i}^T\left[\mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\right]\boldsymbol{d}\end{aligned} \tag{15.10}$$

Combining the above term the remaining terms in equation (15.5), we obtain

$$\begin{aligned}&\boldsymbol{i}^T\left[\mathbf{C}_I^{-1}+\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A}-\mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\mathbf{A}\right]\boldsymbol{i}\\&+\boldsymbol{d}^T\left[\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\mathbf{A}-\mathbf{N}^{-1}\mathbf{A}\right]\boldsymbol{i}\\&+\boldsymbol{i}^T\left[\mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}-\mathbf{A}^T\mathbf{N}^{-1}\right]\boldsymbol{d}\end{aligned} \tag{15.11}$$

The maximum likelihood intensity map is obtained by solving

$$\begin{aligned}&\left[\mathbf{C}_I^{-1}+\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A}-\mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\mathbf{A}\right]\boldsymbol{i}_{\rm ML}\\&\qquad=\left[\mathbf{A}^T\mathbf{N}^{-1}-\mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\right]\boldsymbol{d}\end{aligned} \tag{15.12}$$

Consequently, the posterior reads

$$p_{\rm posterior}\sim\exp\left[-\frac{1}{2}(\boldsymbol{i}-\boldsymbol{i}_{\rm ML})^T\mathcal{C}_{ii}^{-1}(\boldsymbol{i}-\boldsymbol{i}_{\rm ML})\right] \tag{15.13}$$

where

$$\mathcal{C}_{ii}^{-1}\equiv\mathbf{C}_I^{-1}+\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A}-\mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B}+\mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\mathbf{A}. \tag{15.14}$$

The right-hand side of equation (15.12) defines a data projection operator

$$\mathcal{U} \equiv \mathbf{A}^T\mathbf{N}^{-1} - \mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B} + \mathbf{C}_P^{-1})^{-1}\mathbf{B}^T\mathbf{N}^{-1} \tag{15.15}$$

so that

$$\boldsymbol{i}_{\rm ML} = \mathcal{C}_{ii}\mathcal{U}\boldsymbol{d} \tag{15.16}$$

optimally extracts Stokes I from the data with the polarization prior included.

Chapter 16

Discussion

In our earlier discussions, we presented an optimal Stokes-I extraction strategy by incorporating a prior for the sky. In this section, we expand on the topic, exploring Stokes-I extraction from various polarization viewpoints. Additionally, through a simple numerical example, we will compare the variance among these Stokes-I estimators.

16.1 Stokes I Strategies Across Polarization Views

We can treat polarization using four different approaches.

1. View polarization and Stokes I together as sky degrees of freedom to be solved.

2. Consider only Stokes I as the sky degree of freedom to be solved and remove all polarization components from the measurement.

3. Consider only Stokes I as the sky degree of freedom to be solved, but retain the polarization components in the measurement treating them as noise with a power spectrum C_p.

4. Solve for Stokes I with prior information on both polarization and Stokes I. (As in section 15.)

More specifically, starting with initial linear measurements again, we separate the sky Stokes I degree of freedom, denoted as $\boldsymbol{i}$, from the polarization degree of freedom, $\boldsymbol{p}$. Thus the measurement can be expressed as

$$\boldsymbol{d} = \mathbf{M}\,\boldsymbol{s} + \boldsymbol{n} = \mathbf{A}\boldsymbol{i} + \mathbf{B}\boldsymbol{p} + \boldsymbol{n} \tag{16.1}$$

where $\mathbf{A}$ represents the Stokes I beam and $\mathbf{B}$ denotes the polarized beam components.

If we use the least-squares estimation method, under the first of the approaches previously mentioned, we can derive the Maximum Likelihood Estimate (MLE) for the

entire sky $\boldsymbol{s}$

$$\hat{\boldsymbol{s}} = \begin{pmatrix} \boldsymbol{i} \\ \boldsymbol{p} \end{pmatrix} = \mathbf{K_1}\,\boldsymbol{d}, \tag{16.2}$$

where

$$\mathbf{K_1} = (\mathbf{M}^\dagger \mathbf{N}^{-1} \mathbf{M})^{-1} \mathbf{M}^\dagger \mathbf{N}^{-1}, \tag{16.3}$$

and the Stokes I sky $\hat{\boldsymbol{i}}$ is part of $\hat{\boldsymbol{s}}$.

To implement the second approach, we can use linear data projection with SVD, as discussed earlier. However, a faster approach might be to integrate out 'p' in the likelihood, similar to how in the previous section we marginalized over $\boldsymbol{p}$ in the posterior. Without going into repetitive computations, we find that $\hat{\boldsymbol{i}} = \mathbf{K_2}\boldsymbol{d}$ where

$$\mathbf{K_2} = \left[\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A} - \mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\mathbf{A}\right]^{-1} \left[\mathbf{A}^T\mathbf{N}^{-1} - \mathbf{A}^T\mathbf{N}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{N}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{N}^{-1}\right]. \tag{16.4}$$

The third approach is straightforward to implement. The estimator is given by $\hat{\boldsymbol{i}} = \mathbf{K_3}\boldsymbol{d}$ where

$$\mathbf{K_3} = \left[\mathbf{A}^\dagger\left(\mathbf{B}C_P\mathbf{B}^\dagger + \mathbf{N}\right)^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^\dagger\left(\mathbf{B}C_P\mathbf{B}^\dagger + \mathbf{N}\right)^{-1}. \tag{16.5}$$

The fourth approach mentioned corresponds to the 'optimal' estimator defined in section 15. The associated operator, represented as $\mathbf{K_4}$, is given by

$$\mathbf{K_4} = \mathcal{C}_{ii}\mathbf{U}, \tag{16.6}$$

which is defined in equation (15.16).

16.2 Numerical variance analysis

In this section, we introduce a basic numerical example to compare the variance among different Stokes I estimators.

Simulation Scenario

We consider a single HIRAX prime focus reflector equipped with a single feed, situated on the Earth's equator and operated in drift-scan mode. We focus on the $m = 140$ mode of the measurement.

The antenna is directed towards galactic latitudes, incrementing by 0.5° from 75°S to 75°N. Figure 16.1 displays the Stokes beam of the antenna at three distinct pointings, with galactic latitudes of 65°N, 32.5°N, and 0°.

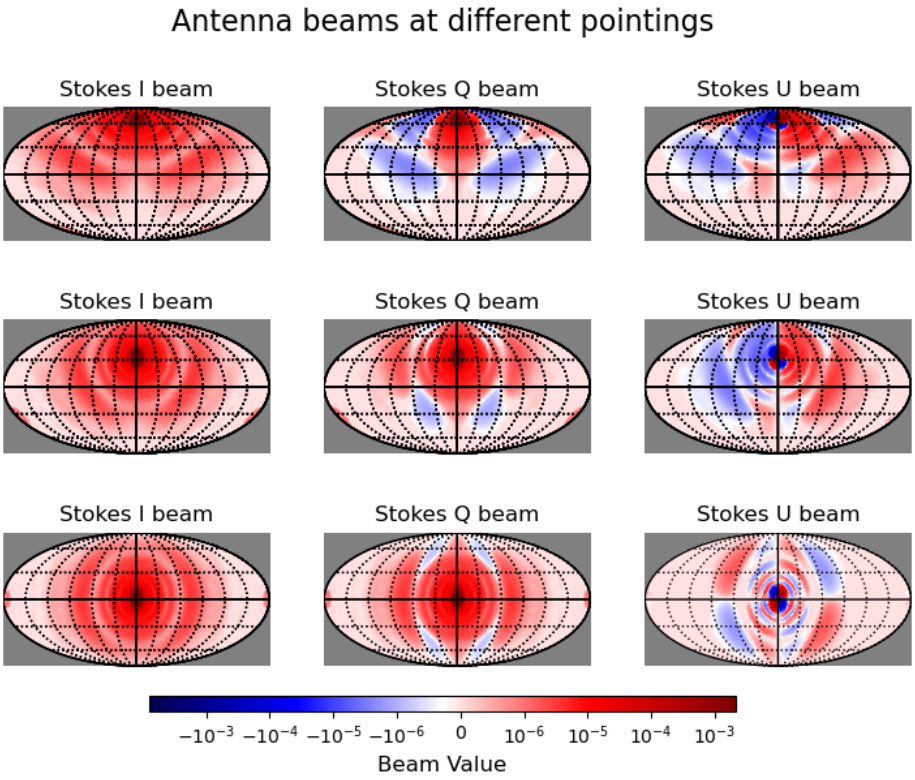

Figure 16.1: Beam Stokes parameters for three different pointing directions.The HIRAX beam data (Crichton et al., 2022) used here are from CST simulations by Carla Pieterse. Here we have assumed that the Stokes V sky signal is negligible.

Settings and Variance Analysis

For our m-mode settings, we have $m = 140$ and $l_{\max} = 191$. This configuration results in an overdetermined system, aiming to solve for sky modes where $m \leq l \leq l_{\max}$ and $m = 140$.

The variance comparisons for the four distinct estimators are illustrated in Figure 16.2. Under the assumption that both input noise and the sky power spectra closely approximate the real power spectra, our illustrative example indicates that the optimal Stokes I extraction exhibits the lowest variance among the considered estimators.

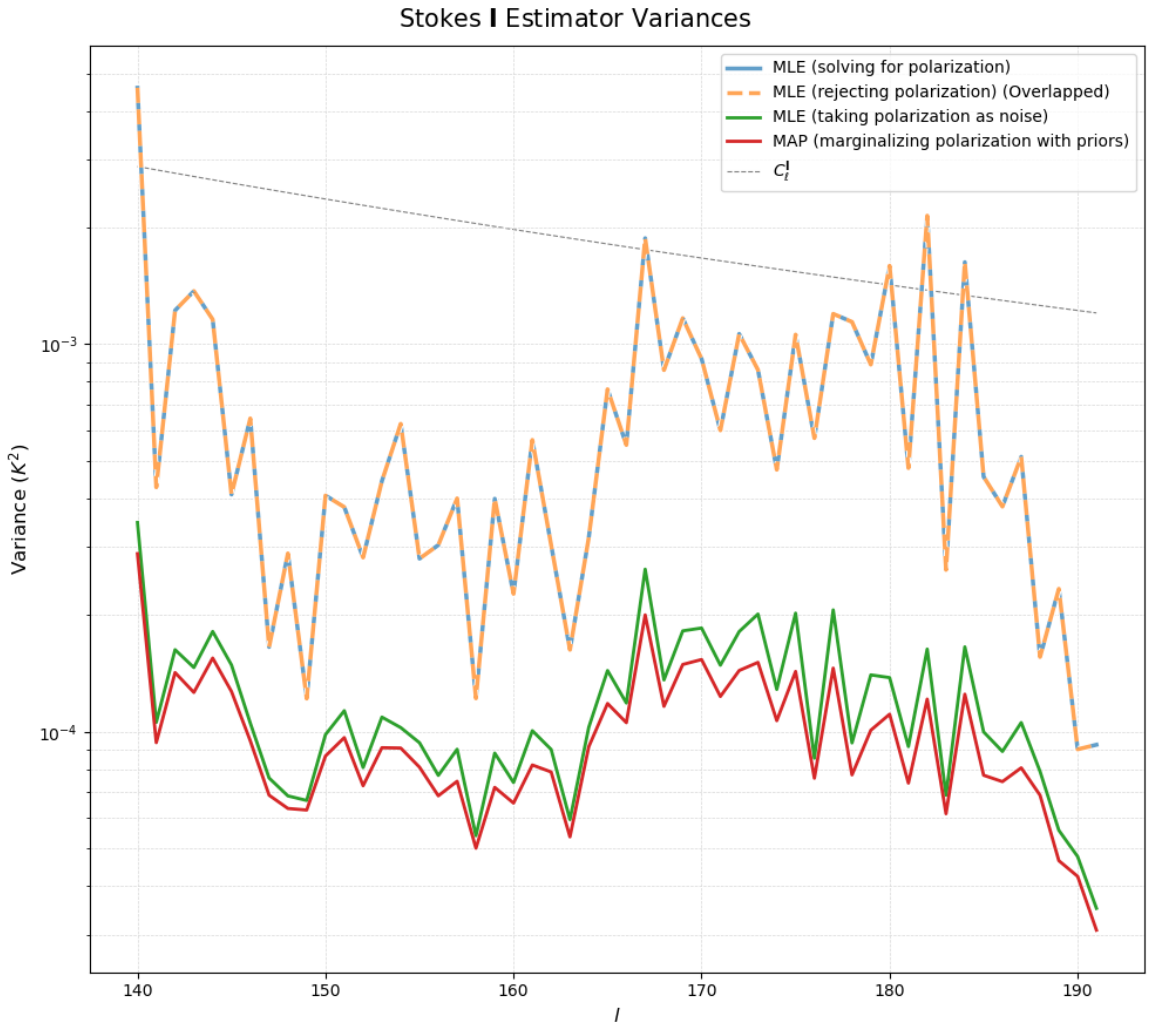

Figure 16.2: Variance comparison among various Stokes-I estimators. The blue line represents the MLE estimator using $\mathbf{K_1}$. The yellow-dashed line corresponds to the MLE estimator with $\mathbf{K_2}$, which projects out polarization. The green line depicts the MLE estimator with $\mathbf{K_3}$, accounting for polarization as noise. Lastly, the red line showcases the maximum posterior estimator using $\mathbf{K_4}$, identified as the optimal Stokes I estimator.

Part V

Figure-of-Merit Beam Characterization for HI Intensity Mapping

Chapter 17

Introduction

The last decade has seen a surge of interest in using the 21cm intensity mapping (IM) as a probe of large scale structure. This idea takes advantage of the optically thin 21cm spectral line arising from the hyperfine transition of neutral atomic hydrogen (HI), which has pervaded the universe after recombination. Although in the late universe reionization has destroyed most of the neutral hydrogen, in galaxies the overdense HI has a recombination rate, proportional to n^2, which is able to overcome the ionization rate, proportional to n, due to the background of ionizing radiation. Moreover, for larger, dense clouds, shielding needs to be taken into account. When the spin temperature is higher than the cosmic microwave background (CMB) temperature, overdense stimulated emission dominates over absorption, which leads to net 21cm emission and can be observed as the difference between the observed temperature and the CMB temperature. Observing the redshifted 21cm emission line traces the three-dimensional galaxy distribution, we can study the clustering of collapsed halos over a wide range of redshifts and spatial scales. Such data provides a powerful way to constrain cosmological models.

The detection of 21cm signal is challenging because this signal is contaminated by astrophysical foregrounds and their coupling with the instrument. Since the foregrounds are orders of magnitude stronger, any inaccuracy of our knowledge about the instrument, for example an imprecise beam characterization, can lead to non-trivial contamination to the cosmological signal.

In work with Bucher, Moodley, Pieterse, Zhang and de Villiers (Bucher et al., in prep.), we developed a figure-of-merit (FoM) that estimates the impact of imprecise beam characterization and beam complexity on intensity mapping science extraction.[1] Compared to complete end-to-end simulations, this technique provides a simple and

[1]In this paper, the contribution of Martin Bucher was the idea and the initial formalism of the isolated FoM beam evaluation. Kavilan Moodley contributed the input power spectra both for the signal and the foregrounds as well as the apodization in frequency scheme. Carla Pieterse computed the Hirax beam data using CST and beam error simulations. Prof Dirk de Villiers provided guidance on the antenna modelling aspects of the project. My contribution was the improve of the formalism and the numerical realisation of the analysis.

quick way to isolate the effect of beam systematics from other aspects of the measurement process. Specifically, we formulated the contribution of the beam error or uncertainty to the reconstructed sky. This beam uncertainty power spectrum, together with the instrumental noise and galactic foreground power spectra, are combined to yield the total noise power spectrum. Correspondingly, we define the signal-to-noise ratio (SNR) in Fourier space. The SNRs of all modes combine to give a figure of merit.

In a more concrete context, consider the example of a single dish experiment as discussed in Bucher et al. (in prep.). In this work, the transfer matrix, also referred to as the beam matrix and denoted as $\mathbf{M}$, is subject to perturbations represented by the stochastic variable $\delta\mathbf{M}$. This uncertainty is integral when assessing the error in reconstructing the cosmological signal, leading to the modified equation:

$$\mathbf{d} = (\mathbf{M} + \delta\mathbf{M})\,\mathbf{s} + \mathbf{n}. \tag{17.1}$$

Assumptions about $\delta\mathbf{M}$ posit it as stemming from a Gaussian stochastic process with a mean of zero. Its probability distribution is solely defined by its covariance matrix $\mathrm{Cov}(\delta\mathbf{M})$. A representation for the contribution of $\delta\mathbf{M}$ to the error in sky reconstruction is given by:

$$\langle \delta\mathbf{s}_{\delta\mathbf{M}}\ \delta\mathbf{s}_{\delta\mathbf{M}} \rangle = \mathbf{M}^{-1}\delta\mathbf{M}\,\mathbf{s}\,\mathbf{s}^T\delta\mathbf{M}^T{\mathbf{M}^{-1}}^T. \tag{17.2}$$

Of particular importance are the off-diagonal components of $\delta\mathbf{M}$. These elements can introduce significant noise by merging various galactic emission modes. The covariance associated with the sky map is described by:

$$\mathrm{Cov}(\delta\mathbf{s}_{\delta\mathbf{M}}) = \mathbf{M}^{-1}\delta\mathbf{M}\,\mathbf{P}_{gal}\,\delta\mathbf{M}^T{\mathbf{M}^{-1}}^T. \tag{17.3}$$

Zooming into the specifics, the operator $\mathbf{M}$ under the flat-sky approximation is diagonal in $\mathbf{k}_\perp$. This leads to the equation:

$$\mathbf{d}(\mathbf{k}_\perp, \nu) = \mathbf{M}(\mathbf{k}_\perp, \nu)\ \mathbf{s}(\mathbf{k}_\perp, \nu). \tag{17.4}$$

Yet, transforming from this space yields a more complex equation:

$$\mathbf{d}(\mathbf{k}_\perp, k_\parallel) = \int dk'_\parallel\ \mathbf{M}(\mathbf{k}_\perp, k'_\parallel)\ \mathbf{s}(\mathbf{k}_\perp, k_\parallel - k'_\parallel). \tag{17.5}$$

Such convolution potentially causes a mixture of modes, especially when the beam pattern's frequency dependency is taken into account.

However, the techniques presented there ignored polarization, which ignores the details of the Stokes-I extraction and the uncertainty of the polarized Stokes beam: it was assumed that the sky is unpolarized, or that the reception system is sensitive only to intensity. Another case getting rid of polarization is to fully reject the polarized

components using the "full" knowledge, which is an ideal assumption, of the polarized response of the primary beam as discussed in Section 14.2. These ideal assumptions could be broken down or at least come under close scrutiny when considering the complexities of real observations. On the one hand, the Galactic synchrotron emission is highly polarized and suffers from Faraday rotation effects as it propagates through the galactic magnetic field. As a result, the observed polarized foregrounds have a complex spectral structure, which makes it difficult for a foreground removal strategy based on the spectral smoothness of the galactic foregrounds (though this strategy works very well for the galactic Stokes I foreground). On the other hand, almost all 21cm intensity mapping experiments are constructed with orthogonal feed polarizations, generally in the form of co-located orthogonally oriented dipoles. Any imperfectness of the practical polarimetry implementation can lead to an entanglement of different polarized components in the data, which is termed "polarization leakage" (see Liu and Shaw 2020 for a further introduction). Although we are only interested in the 21cm intensity, the contamination from galactic polarization emission in the instrumental Stokes I data hinders the correct estimation of the cosmological signal. Thus, we inevitably need to address the problem of polarization for beam evaluations.

In this Part we further develop the figure-of-merit formalism for beam evaluation, with particular emphasis on the role of beam uncertainty in the 21cm power spectrum extracted from polarization measurements. One should note that the actual beam error propagation depends on the actual realizations of the Stokes-I extraction and power spectrum estimation. For our purpose, it suffices to discuss in a certain scenario using the maximum-likelihood estimator and the common unpolarized measurement strategy (XX+YY). The discussion here can be generalized to other linear strategies like the optimal Stokes-I extraction and the polarization rejection strategy introduced in Part IV of this thesis.

The group of chapters is organized as follows. After formulating the general polarization measurement in a scalar way and establishing the figure-of-merit formalism in Chapter 18, we model the input power spectra of the cosmological signal and various contaminants in Chapter 19. In Chapter 20, we summarize the work and present our conclusions.

Chapter 18

Analysis Formalism

18.1 Scalar formalism of polarization measurements

Polarized signals are typically measured using a pair of feeds that are sensitive to orthogonal linear polarizations at long wavelengths, or a single horn feed that separates the orthogonal circular or linear polarizations at shorter wavelengths. In this work we consider antennas with linearly polarized feeds, since that is what is most commonly used for most 21cm intensity mapping experiments. But the discussion should be easily generalized to circular polarization analysis.

The correlation between a pair of feeds has been given by equation (6.51), which can be rewritten as

$$V_{Ip,Jq} = \sum_{\mathcal{S}=\mathbf{I},\mathbf{Q},\mathbf{U},\mathbf{V}} \int \mathrm{d}^2\hat{\mathbf{n}}\, E^{\alpha *}_{Ip}(\hat{\mathbf{n}}-\hat{\mathbf{n}}_c) E^{\beta}_{Jq}(\hat{\mathbf{n}}-\hat{\mathbf{n}}_c)\, \sigma^{\mathcal{S}}_{\alpha\beta} \exp\left(-2\pi i \frac{\hat{\mathbf{n}}\cdot\mathbf{b}_{IJ}}{\lambda}\right) T_{\mathcal{S}}(\hat{\mathbf{n}}) + n \tag{18.1}$$

where I and J are indices of antennas, and $p, q = X, Y$ refer to the linear polarized feeds on the antennas. The repeated Greek indices, α and β, representing the polarization basis on the sky sphere[1], are summed over by convention. E^{α}_{Ip} represents the α-component of $\mathbf{E}_{Ip}$, the far field of the p-polarized feed of antenna I. The baseline vector $\mathbf{b}_{IJ}$ describes the relative position of the feed pair. The degenerate case of equation (18.1) with $I = J$ and $p = q$ is called the auto-correlation of the feed. In general, one should also consider the pointing dependence of the Stokes $Q(U)$ beams, as discussed in Section 6.3, but here we assume a narrow sky strip survey so that the Stokes $Q(U)$ beams are assumed to be invariant over the survey area.

Although equation (18.1) directly relates the far fields of the antennas to the correlation measurements, the Stokes Q and U sky maps and antenna beams are not scalar fields we can instead use the E and B maps. For more details on a general scalar description of a polarization field, see section 5.2.2).

[1] We shall just use the celestial coordinate basis vectors, $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\phi$.

We now proceed to define the E and B beams and introduce the flat-sky limit together. Neglecting the noise and the constant phase shift term caused by the part of the baseline along the antenna pointing direction, the measurement in the flat-sky approximation is given by

$$V_{Ip,Jq} = \sum_{\mathcal{S}=\mathbf{I},\mathbf{Q},\mathbf{U},\mathbf{V}} \iint \mathrm{d}^2\mathbf{x}\, B^{\mathcal{S}}_{pq}(\mathbf{x})\, T_{\mathcal{S}}(\mathbf{x_0}-\mathbf{x}) \exp\left(-2\pi i \frac{\mathbf{x}\cdot\mathbf{b}_{IJ,\perp}}{\lambda}\right) \tag{18.2}$$

where $\mathbf{x}$ is the Cartesian beam coordinates, and $\mathbf{x_0}$ is the beam coordinates of the origin of the sky coordinate system. $\mathbf{b}_\perp$ is the part of $\mathbf{b}$ perpendicular to the pointing direction. $B^{\mathcal{S}}_{pq} = E^{\alpha *}_{Ip} E^{\beta}_{Jq} \sigma^{\mathcal{S}}_{\alpha\beta}$ is the primary beam of the cross-correlation, where the antenna indices have been dropped as we assume identical feeds and dishes for all baselines. Equation (18.2) can be rewritten in Fourier conjugate coordinates of $\mathbf{x}$, which is denoted by $\mathbf{q}$:

$$V_{Ip,Jq} = \sum_{\mathcal{S}=\mathbf{I},\mathbf{Q},\mathbf{U},\mathbf{V}} \iint \frac{\mathrm{d}^2\mathbf{q}}{(2\pi)^2}\, B^{\mathcal{S}}_{pq}\left(\mathbf{q}+\frac{2\pi\mathbf{b}_{IJ,\perp}}{\lambda}\right) T_{\mathcal{S}}(\mathbf{q})\, e^{i\mathbf{q}\cdot\mathbf{x_0}} \tag{18.3}$$

where $T_{\mathcal{S}}(\mathbf{q})$ is the Fourier $\mathbf{q}$ mode in the sky coordinate system, while the extra phase $e^{i\mathbf{q}\cdot\mathbf{x_0}}$ shifts it to the beam coordinate system.

In Fourier space, Stokes Q and U fields are related to the E and B modes by (Dodelson and Schmidt, 2020)

$$\begin{aligned} T_{\mathbf{Q}}(\mathbf{q}) &= \cos\left(2\phi(\mathbf{q})\right) T_{\mathbf{E}}(\mathbf{q}) - \sin\left(2\phi(\mathbf{q})\right) T_{\mathbf{B}}(\mathbf{q}) \\ T_{\mathbf{U}}(\mathbf{q}) &= \sin\left(2\phi(\mathbf{q})\right) T_{\mathbf{E}}(\mathbf{q}) + \cos\left(2\phi(\mathbf{q})\right) T_{\mathbf{B}}(\mathbf{q}) \end{aligned} \tag{18.4}$$

where $\phi(\mathbf{q})$ is the azimuthal angle of the 2D wavevector, which satisfies $\mathbf{q} = (q_x, q_y) = (\cos\phi(\mathbf{q}), \sin\phi(\mathbf{q}))|\mathbf{q}|$. Substituting $T_{\mathbf{Q}(\mathbf{U})}$ by $T_{\mathbf{E}(\mathbf{B})}$, equation (18.3) can be rewritten in terms of E and B modes

$$V_{Ip,Jq} = \sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} \iint \frac{\mathrm{d}^2\mathbf{q}}{(2\pi)^2}\, B^{\mathcal{S}'}_{pq,\mathbf{b}_{IJ}}\left(\mathbf{q}+\frac{2\pi\mathbf{b}_{IJ,\perp}}{\lambda}\right) T_{\mathcal{S}'}(\mathbf{q})\, e^{i\mathbf{q}\cdot\mathbf{x_0}} \tag{18.5}$$

where $B^{\mathbf{E}(\mathbf{B})}_{pq,\mathbf{b}_{IJ}}$ are the effective beams defined by

$$\begin{pmatrix} B^{\mathbf{E}}_{pq,\mathbf{b}_{IJ}}\left(\mathbf{q}+\frac{2\pi\mathbf{b}_{IJ,\perp}}{\lambda}\right) \\ \\ B^{\mathbf{B}}_{pq,\mathbf{b}_{IJ}}\left(\mathbf{q}+\frac{2\pi\mathbf{b}_{IJ,\perp}}{\lambda}\right) \end{pmatrix} = \begin{pmatrix} \cos\left(2\phi(\mathbf{q})\right) & \sin\left(2\phi(\mathbf{q})\right) \\ \\ -\sin\left(2\phi(\mathbf{q})\right) & \cos\left(2\phi(\mathbf{q})\right) \end{pmatrix} \begin{pmatrix} B^{\mathbf{Q}}_{pq}\left(\mathbf{q}+\frac{2\pi\mathbf{b}_{IJ,\perp}}{\lambda}\right) \\ \\ B^{\mathbf{U}}_{pq}\left(\mathbf{q}+\frac{2\pi\mathbf{b}_{IJ,\perp}}{\lambda}\right) \end{pmatrix} \tag{18.6}$$

Rewritten in the real space, equation (18.5) is in the form of the sum of four convolu-

tions of scalar fields:

$$V_{Ip,Jq}(\mathbf{x_0}) = \sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} \int \mathrm{d}^2\mathbf{x}\, B^{\mathcal{S}'}_{pq,\mathbf{b}_{IJ}}(\mathbf{x})\, T_{\mathcal{S}'}(\mathbf{x_0}-\mathbf{x}) \exp\left(-2\pi i \frac{\mathbf{x}\cdot\mathbf{b}_{IJ,\perp}}{\lambda}\right). \tag{18.7}$$

18.2 Unpolarized measurement and beam uncertainty

In single-dish experiments, the two feeds are both auto-correlated and cross-correlated if polarization measurements are desired. However, for total intensity measurements, it is most commonly just taking the sum of two auto-correlations, which takes advantage of the opposite sign responses of XX and YY to the linear polarizations and no responses to Stokes V. Similarly for interferometers, two cross-correlations of parallel feeds on a baseline are combined as a total intensity measurement. Therefore, a usual unpolarized measurement can be generally represented as $d_{IJ} = V_{Ix,Jx} + V_{Iy,Jy}$.

Rendering all the unpolarized measurements into a data vector, we get the linear measurement equation

$$\mathbf{d} = \mathbf{M_I}\,\mathbf{T_I} + \mathbf{n}. \tag{18.8}$$

Generally there are more equations than unknowns and because of noise and imperfections in determining $\mathbf{M}$, these equation as is are inconsistent and formally have no solution. The standard solution is to use least squares or generalized least squares to find a best compromise solution, informed by the noise covariance matrix $\mathbf{N}$. Usually we take the Gaussian approximation to the noise statistics. Then an estimator of the intensity sky is chosen to minimize the negative log-likelihood

$$\chi^2 = \frac{1}{2}\left(\mathbf{d} - \mathbf{M_I}\,\mathbf{T_I}\right)^T \mathbf{N^{-1}} \left(\mathbf{d} - \mathbf{M_I}\,\mathbf{T_I}\right), \tag{18.9}$$

where $\mathbf{N}$ is the noise covariance matrix $\left\langle \mathbf{n}\mathbf{n}^\dagger \right\rangle$. The maximum likelihood (ML) estimator of sky map is therefore given by

$$\mathbf{T}_{est} = \mathbf{K}\,\mathbf{d}, \tag{18.10}$$

where

$$\mathbf{K} = (\mathbf{M_I}^T\mathbf{N}^{-1}\mathbf{M_I})^{-1}\mathbf{M_I}^T\mathbf{N}^{-1}. \tag{18.11}$$

For single-dish measurements, $\mathbf{M_I}$ might be square and non-degenerate, in which case $\mathbf{K}$ reduces to $\mathbf{M_I}^{-1}$.

However, in practice the beam matrix obtained from measurements or simulations includes an error. On the one hand, defects in the manufacturing process or acquired beam errors caused by deformations are problematic. On the other hand, due to the

change of environmental conditions with the observation time, factors like temperature and winds, or some other perturbations, may introduce uncertainties in the actual beam response. Furthermore, even if the beam is perfect and immune to all defects and deformations, the intrinsic errors in the simulation and measurement of the an antenna beam can also play a role.

In the absence of a perfect absolute description of beams, one can seek for adequate statistical descriptions of beam errors in addition to measurements or simulations. According to the time-varying nature of beam errors, we can roughly divide them into two categories: *stochastic beam errors* and *time varying beam errors*. Stochastic beam errors refer to errors that do not change over observation, and thus can be viewed as an instance drawn from some statistical model. In data analysis based on the convolution of the beam and the sky, the stochastic beam error will be block diagonal in coordinates of tangential wave number after the deconvolution, just as what a primary beam does. While time varying beam errors are defined as beam errors changing during the scanning of the sky, which might lead to the mode mixing or "aliasing" of tangential Fourier modes.

18.3 General figure-of-merit formalism

In this section, we propose a figure-of-merit beam characterization in the context of 21cm intensity mapping with full treatment of polarization. This scheme is designed to characterize the bias and variance of the error power spectrum (PS) of the reconstructed sky caused by beam uncertainties. In particular, we emphasize the mode mixing effect, especially the leakage from Galactic dominated modes into high SNR modes for the cosmological signal as well as the leakage from polarized radio sky.

Less abstractly, we assume that the true feed response $\mathbf{E}_{Jp}$ differs from the evaluated far field pattern $\mathbf{E}_p$ by $\delta\mathbf{E}_{Jp}$:[2]

$$\mathbf{E}_{Jp} \to \mathbf{E}_p + \delta\mathbf{E}_{Jp}. \tag{18.12}$$

As a result, the beam matrix of an unpolarized measurement to the Stokes I sky is correspondingly perturbed, which can be expressed as

$$\mathbf{M_I} \to \mathbf{M_I} + \delta\mathbf{M_I}. \tag{18.13}$$

Also, differences in the linear polarization responses of two identical but crossed feeds lead to nonzero polarized beams, which are written as

$$\mathbf{M}_{\mathcal{S}'} \to 0 + \delta\mathbf{M}_{\mathcal{S}'} \tag{18.14}$$

[2]Here we used the antenna index J. This is because uncertainties of different feeds can be understood as different instances of statistical beam error models.

for $\mathcal{S}' \in \{\mathbf{E}, \mathbf{B}, \mathbf{V}\}$. We leave the specific form of the beam matrices to later sections. In this section, we focus on defining a general form for beam characterization.

An unpolarized measurement can be decomposed into the estimated beams and the beam errors

$$\mathbf{d} = \mathbf{M_I}\,\mathbf{T_I} + \sum_{\mathcal{S}'} \delta\mathbf{M}_{\mathcal{S}'}\,\mathbf{T}_{\mathcal{S}'} + \mathbf{n}. \tag{18.15}$$

The estimator of the sky intensity is then given by

$$\mathbf{T}_{est} = \mathbf{T_I} + \delta\mathbf{T} + \mathbf{K}\,\mathbf{n}, \tag{18.16}$$

where

$$\delta\mathbf{T} = \sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} \mathbf{K}\,\delta\mathbf{M}_{\mathcal{S}'}\,\mathbf{T}_{\mathcal{S}'} \tag{18.17}$$

is the error in the reconstructed sky due to the beam error. The power spectrum of the reconstructed Stokes I sky can be divided into several components, as shown below:

PS of the reconstructed sky

$\mathbf{P}_{\delta\mathbf{M}}$: Beam Uncertainty PS

$$\left\langle \mathbf{T}_{est}\mathbf{T}_{est}^{\dagger} \right\rangle = \left\langle \mathbf{T_I}\mathbf{T_I}^{\dagger} \right\rangle + \left\langle \delta\mathbf{T}\mathbf{T_I}^{\dagger} \right\rangle + \left\langle \mathbf{T_I}\delta\mathbf{T}^{\dagger} \right\rangle + \left\langle \delta\mathbf{T}\delta\mathbf{T}^{\dagger} \right\rangle + \mathbf{K}\left\langle \mathbf{n}\mathbf{n}^{\dagger} \right\rangle \mathbf{K}^{\dagger} \tag{18.18}$$

$\mathbf{P}_{sky}$: Foreground + 21cm Signal

$\mathbf{P}_{\mathbf{n}}$: Noise PS

More explicitly, assuming that the polarized emission is uncorrelated, the error power spectrum caused by the beam uncertainty is given by

$$\mathbf{P}_{\delta\mathbf{M}} = \sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} \mathbf{K}\,\delta\mathbf{M}_{\mathcal{S}'}\,\mathbf{P}^{\mathcal{S}'}_{sky}\,\delta\mathbf{M}^{\dagger}_{\mathcal{S}'}\,\mathbf{K}^{\dagger} + \mathbf{K}\,\delta\mathbf{M}_{\mathbf{I}}\,\mathbf{P}^{\mathbf{I}}_{sky} + \mathbf{P}^{\mathbf{I}}_{sky}\,\delta\mathbf{M_I}^{\dagger}\,\mathbf{K}^{\dagger} \tag{18.19}$$

Note that the beam matrices $\delta\mathbf{M}_{\mathcal{S}'}$ in equation (18.19) are actual beam errors. A statistical description of $\mathbf{P}_{\delta\mathbf{M}}$ is possible, as long as we consider an ensemble of $\delta\mathbf{M}_{\mathcal{S}'}$, whose elements are described by a set of independent statistical variables. Usually, one can take the ensemble average, $E(\mathbf{P}_{\delta\mathbf{M}}) = \langle\mathbf{P}_{\delta\mathbf{M}}\rangle$, as the bias, or take its standard deviation $\sigma(\mathbf{P}_{\delta\mathbf{M}}) = \langle|\mathbf{P}_{\delta\mathbf{M}} - E(\mathbf{P}_{\delta\mathbf{M}})|\rangle$ as a characterization of the 'noise', or take the root of mean squared $\mathbf{P}_{\delta\mathbf{M}}$ as its statistical characterization, which is given by

$$\sqrt{E\left(\mathbf{P}^2_{\delta\mathbf{M}}\right)} \equiv \sqrt{E^2(\mathbf{P}_{\delta\mathbf{M}}) + \sigma^2(\mathbf{P}_{\delta\mathbf{M}})}. \tag{18.20}$$

In this work, we will take the expectation value of the quadratic term as well as the

standard deviation of the first order term so that

$$\begin{aligned}\overline{\mathbf{P}_{\delta\mathbf{M}}} &\equiv E\left(\sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} \mathbf{K}\,\delta\mathbf{M}_{\mathcal{S}'}\,\mathbf{P}^{\mathcal{S}'}_{sky}\,\delta\mathbf{M}^{\dagger}_{\mathcal{S}'}\,\mathbf{K}^{\dagger}\right) + \sigma\left(\mathbf{K}\,\delta\mathbf{M}_{\mathbf{I}}\,\mathbf{P}^{\mathbf{I}}_{sky} + \mathbf{P}^{\mathbf{I}}_{sky}\,\delta\mathbf{M}_{\mathbf{I}}{}^{\dagger}\,\mathbf{K}^{\dagger}\right) \\ &= \sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} \mathbf{K}\,\overline{\delta\mathbf{M}}_{\mathcal{S}'}\,\mathbf{P}^{\mathcal{S}'}_{sky}\,\overline{\delta\mathbf{M}}^{\dagger}_{\mathcal{S}'}\,\mathbf{K}^{\dagger} + \mathbf{K}\,\overline{\delta\mathbf{M}}_{\mathbf{I}}\,\mathbf{P}^{\mathbf{I}}_{sky} + \mathbf{P}^{\mathbf{I}}_{sky}\,\overline{\delta\mathbf{M}}^{\dagger}_{\mathbf{I}}\,\mathbf{K}^{\dagger}\end{aligned} \tag{18.21}$$

where $\overline{\delta\mathbf{M}}_{\mathcal{S}'} = \sigma(\delta\mathbf{M}_{\mathcal{S}'})$ is the beam error matrix whose entries are the standard deviation of the stochastic beam uncertainty of the corresponding beam element. The expectation of the quadratic term is conveniently given by the standard deviation of the random variable. However, if we take the expectation of the entire error power spectrum, $E(\mathbf{P}_{\delta\mathbf{M}})$, we will lose the error information from the linear term if we assume a mean zero stochastic beam error. Since the real world measurement is made only once, the linear term perturbation exists and we want to preserve some of the effect when estimating the power spectrum error. Taking the standard deviation of the linear perturbation term is a reasonable choice.

To better characterize the contamination of the beam uncertainty on the 21 cm PS, it is necessary to compute the data components in 3-dimensional Fourier space. Theoretically, we should formulate the statistical measurements in the comoving coordinates. However, since the mathematical form appears much simpler in observational coordinates $(\hat{\mathbf{n}}, \nu)$, and also considering that the 3D survey regions of these experiments can usually be regarded as thin redshift shells. Therefore, it is sufficient to use the Fourier conjugate of the observational coordinates, the transverse and radial wavenumbers, for a general characterization of the contamination.

However, a continuous Fourier transform with respect to ν could lead to a rather complicated expression with non-negligible finite integration limits and boundary effects. Therefore, we use the discrete cosine transform (DCT) instead. A power spectrum is transformed from frequency coordinates to DCT modes as[3]

$$\mathbf{P}_{ab}(\mathbf{q}) = \int \mathrm{d}\nu \int \mathrm{d}\nu'\; f_a(\nu) f_b(\nu') \mathbf{P}(\mathbf{q}, \nu, \nu') \tag{18.22}$$

where the sky component in $\mathbf{P}$, $\mathbf{P}_{sky}(\mathbf{q}, \nu, \nu')$ approximates the angular power spectrum in the flat-sky approximation:

$$\mathbf{P}_{sky}(\mathbf{q}, \nu, \nu') \approx C_{l=|\mathbf{q}|}(\nu, \nu'), \tag{18.23}$$

which are discussed in detail in Chapter 19.

[3]Note the difference between $\mathbf{q}$ and $\mathbf{k}_\perp$. $\mathbf{q}$ is the transverse component of the wavevector in "observational" coordinates, whereas scaling down $\mathbf{q}$ with the comoving distance at the observed redshift gives $\mathbf{k}_\perp = \mathbf{q}/\chi$, the transverse wavevector in "cosmological" coordinates.

Putting all the previous results together, we now define a single Figure-of-Merit (FoM) allowing us to rank reflector-feed systems according to the factor by which the total SNR for detecting a cosmological signal is decreased after beam uncertaity error has been included. The total SNR for detecting a cosmological signal may be defined as a sum of SNR over all Fourier modes[4]

$$(\mathrm{SNR})_{tot} = \sum_{\mathbf{q},a} \mathrm{SNR}(\mathbf{q}, a) \tag{18.24}$$

In the absence of beam uncertainty, the per mode SNR is given by

$$\mathrm{SNR}(\mathbf{q}, a) = \frac{\mathbf{P}_{21\mathrm{cm},aa}(\mathbf{q})}{\mathbf{P}_{\mathrm{fg},aa}(\mathbf{q}) + \mathbf{P}_{\mathbf{n},aa}(\mathbf{q})} \tag{18.25}$$

When beam uncertainty comes into consideration, this becomes modified to

$$\overline{\mathrm{SNR}}(\mathbf{q}, a) = \frac{\mathbf{P}_{21\mathrm{cm},aa}(\mathbf{q})}{\mathbf{P}_{\mathrm{fg},aa}(\mathbf{q}) + \mathbf{P}_{\mathbf{n},aa}(\mathbf{q}) + \overline{\mathbf{P}_{\delta\mathbf{M},aa}}(\mathbf{q})} \tag{18.26}$$

Then the FoM is obtained as

$$\mathrm{F.o.M.} = \frac{\sum_{\mathbf{q},a} \overline{\mathrm{SNR}}(\mathbf{q}, a)}{\sum_{\mathbf{q},a} \mathrm{SNR}(\mathbf{q}, a)}. \tag{18.27}$$

18.4 Single-dish analysis

Most 21cm experiments of interest are interferometric, largely owing to the long wavelengths involved. However, because of the less complicated data path from collection to map making, single dish experiments (e.g., FAST) and the single-dish mode of an array (e.g., MeerKAT) are also considered means for HI intensity mapping.

Dropping the indices, the polarized beam patterns of the auto-correlation can be associated to the far fields in the form of matrix multiplications,

$$B^{\mathcal{S}}_{pp}(\hat{\mathbf{n}}) = \mathbf{E}^{\dagger}_{p}(\hat{\mathbf{n}}) \sigma^{\mathcal{S}} \mathbf{E}_{p}(\hat{\mathbf{n}}) \tag{18.28}$$

where $\mathcal{S} = \mathbf{I}, \mathbf{Q}, \mathbf{U}, \mathbf{V}$ and the far-field vector reads $\mathbf{E}_x(\hat{\mathbf{n}}) = (E^{\theta}_x(\hat{\mathbf{n}}),\ E^{\phi}_x(\hat{\mathbf{n}}))^T$. Equation (18.6) allows us to get E/B beams out of Q/U beams. Then the beam matrices for the unpolarized measurement are given by

$$\mathbf{M}_{\mathcal{S}'} = B^{\mathcal{S}'}_{xx} + B^{\mathcal{S}'}_{yy}, \tag{18.29}$$

[4]In a more complete analysis, one may instead want to look at the ability to measure certain cosmological parameters, which is effectively a reweighting of the modes in the definition of total SNR.

for $\mathcal{S}' = \mathbf{I}, \mathbf{E}, \mathbf{B}, \mathbf{V}$. Ideally, $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{V}$ beams would vanish. Adopting a continuous notation, we may write a single-dish measurement explicitly as

$$\mathbf{d}(\hat{\mathbf{n}}', \nu) = \int \mathrm{d}^2\hat{\mathbf{n}}\, \mathbf{M_I}(\hat{\mathbf{n}} - \hat{\mathbf{n}}', \nu)\, \mathbf{T_I}(\hat{\mathbf{n}}, \nu) + \mathbf{n}(\hat{\mathbf{n}}', \nu). \tag{18.30}$$

In the flat-sky approximation, the above equation can be regarded as the convolution of the Stokes-I sky and the beam profile on the 2D plane. As the consequence of convolution theorem, the above equation can be deconvolved into the Fourier conjugate coordinates of $\Delta\hat{\mathbf{n}}$, i.e., the tangential wavenumbers $\mathbf{q}$:

$$\mathbf{d}(\mathbf{q}, \nu) = \mathbf{M_I}(\mathbf{q}, \nu)\, \mathbf{T_I}(\mathbf{q}, \nu) + \mathbf{n}(\mathbf{q}, \nu), \tag{18.31}$$

which is seen in the discrete linear system of measurement equations as $\mathbf{M_I}$ being block diagonal in $\mathbf{q}$. The Stokes I sky estimator in $\mathbf{q}$ coordinates then becomes

$$\mathbf{T}_{est}(\mathbf{q}, \nu) = [\mathbf{M_I}(\mathbf{q}, \nu)]^{-1}\mathbf{d}(\mathbf{q}, \nu). \tag{18.32}$$

Now if we consider a general uncertainty term $\delta\mathbf{E}_p$ in the feed far-field pattern, the primary beam uncertainty can be expressed as

$$\delta B^{\mathcal{S}}_{pp} = (\delta\mathbf{E}_p)^\dagger \sigma^{\mathcal{S}} \mathbf{E}_p + \mathbf{E}_p^\dagger \sigma^{\mathcal{S}} (\delta\mathbf{E}_p) + (\delta\mathbf{E}_p)^\dagger \sigma^{\mathcal{S}} (\delta\mathbf{E}_p), \tag{18.33}$$

which also determines the form of $\delta\mathbf{M}_{\mathcal{S}'}$.

The reconstructed sky is shifted by the beam errors according to

$$\delta\mathbf{T} = \sum_{\mathcal{S}'} \mathbf{M_I}^{-1} \delta\mathbf{M}_{\mathcal{S}'} \mathbf{T}_{\mathcal{S}'} \tag{18.34}$$

For any "stochastic beam error," we have $\delta\mathbf{M}_{\mathcal{S}'} = \delta\mathbf{M}_{\mathcal{S}'}(\hat{\mathbf{n}} - \hat{\mathbf{n}}')$ so that $\delta\mathbf{M}_{\mathcal{S}'}$ is also diagonal in $\mathbf{q}$. Equation (18.34) in $\mathbf{q}$ coordinate then reads

$$\delta\mathbf{T}(\mathbf{q}, \nu) = \sum_{\mathcal{S}'} \frac{\delta\mathbf{M}_{\mathcal{S}'}(\mathbf{q}, \nu)}{\mathbf{M_I}(\mathbf{q}, \nu)}\, \mathbf{T}_{\mathcal{S}'}(\mathbf{q}, \nu). \tag{18.35}$$

However, for the time varying beam errors, the error propagation appears more complicated; because of the dependence on the absolute position of the beam center, $\delta\mathbf{M}_{\mathcal{S}'}$ is not diagonal in $\mathbf{q}$ and results in a mixing of modes. In this case, the induced error in the estimator is given by

$$\delta\mathbf{T}(\mathbf{q}, \nu) = \sum_{\mathcal{S}'} \sum_{\mathbf{q}'} \frac{\delta\mathbf{M}_{\mathcal{S}'}(\mathbf{q}, \mathbf{q}', \nu)}{\mathbf{M_I}(\mathbf{q}, \nu)}\, \mathbf{T}_{\mathcal{S}'}(\mathbf{q}', \nu) \tag{18.36}$$

18.5 Interferometric analysis

In traditional radio astronomy, interferometric measurements are more used in the "tracking" mode, where an antenna array is pointed to a specific patch of the sky and the Earth rotation continuously rotates the antenna baselines used for synthesis imaging. The entire observation can be understood as a set of varying baselines measuring different modes of the sky brightness constant.

However, many interferometric experiments nowadays drift scan the sky. All dishes are pointed at the meridian at a given declination, and the sky overhead moves as the Earth rotates. Since BAO measurements usually require a large survey area, drift scan observations are favored by HI intensity mapping experiments over the traditional tracking mode which requires more costly hardware. For this reason, we consider only drift scan interferometric arrays.

Drift-scan interferometric observations are more closely related to single-dish experiments in that the sky is convolved with a beam in both cases. In the flat-sky approximation, the beam of a baseline may be considered the primary beam modulated by $\exp\left(-2\pi i\frac{\Delta\hat{\mathbf{n}}\cdot\mathbf{b}_\perp}{\lambda}\right)$. In a drift-scan survey, $\mathbf{b}_\perp$ does not change with the rotation of the Earth. However, $\mathbf{b}_\perp$, and therefore the modulated beam, will change when dish is pointed from one declination to another. A drift-scan survey only convolves the constant beam with the sky along the east-west direction. This is different from the two-dimensional convolution in the single-dish mode. Only when the various pointings are limited in a narrow range of declination can we approximate a drift-scan survey as a 2D convolution along the longitude and latitude of the celestial sphere.

18.5.1 2D convolution scenario

Let us first consider a survey strategy for which the 2D convolution viewpoint is a good approximation. By Fourier transforming equation (18.7) with respect to $\mathbf{x_0}$, we deconvolve a single measurement as

$$V_{Ip,Jq}(\mathbf{q}) = \sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} B^{\mathcal{S}'}_{pq,\mathbf{b}_{IJ}}\left(\mathbf{q}+2\pi\frac{\mathbf{b}_{IJ,\perp}}{\lambda}\right)T_{\mathcal{S}'}(\mathbf{q}). \tag{18.37}$$

The perturbed primary beams for $\mathcal{S}=\mathbf{I},\mathbf{Q},\mathbf{U},\mathbf{V}$ are given by

$$\delta B^{\mathcal{S}}_{pq,\mathbf{b}_{IJ}} = (\delta\mathbf{E}_{Ip})^\dagger\sigma^{\mathcal{S}}\mathbf{E}_q + \mathbf{E}_p^\dagger\sigma^{\mathcal{S}}(\delta\mathbf{E}_{Jq}) + (\delta\mathbf{E}_{Ip})^\dagger\sigma^{\mathcal{S}}(\delta\mathbf{E}_{Jq}), \tag{18.38}$$

from which the uncertainty primary beams $\delta B^{\mathcal{S}'}_{pp,\mathbf{b}_{IJ}}$ in E/B mode can be derived. In the real space, the estimated beam matrix of the unpolarized measurement is given by

$$\mathbf{M}_{\mathcal{S}',\mathbf{b}_{IJ}}(\mathbf{x}) = \left[B^{\mathcal{S}'}_{xx,\mathbf{b}_{IJ}}(\mathbf{x}) + B^{\mathcal{S}'}_{yy,\mathbf{b}_{IJ}}(\mathbf{x})\right]\exp\left(-2\pi i\frac{\mathbf{x}\cdot\mathbf{b}_{IJ,\perp}}{\lambda}\right), \tag{18.39}$$

which as pointed out before is diagonal in $\mathbf{q}$ coordinates, so that

$$\mathbf{M}_{\mathcal{S}',\mathbf{b}_{IJ}}(\mathbf{q}) = B^{\mathcal{S}'}_{xx,\mathbf{b}_{IJ}}\left(\mathbf{q} + 2\pi\frac{\mathbf{b}_{IJ,\perp}}{\lambda}\right) + B^{\mathcal{S}'}_{yy,\mathbf{b}_{IJ}}\left(\mathbf{q} + 2\pi\frac{\mathbf{b}_{IJ,\perp}}{\lambda}\right). \tag{18.40}$$

The perturbed beam is of a similar diagonal form only in cases of "stochastic beam errors"

$$\delta\mathbf{M}_{\mathcal{S}',\mathbf{b}_{IJ}}(\mathbf{q}) = \delta B^{\mathcal{S}'}_{xx,\mathbf{b}_{IJ}}\left(\mathbf{q} + 2\pi\frac{\mathbf{b}_{IJ,\perp}}{\lambda}\right) + \delta B^{\mathcal{S}'}_{yy,\mathbf{b}_{IJ}}\left(\mathbf{q} + 2\pi\frac{\mathbf{b}_{IJ,\perp}}{\lambda}\right). \tag{18.41}$$

Otherwise it would become some general non-diagonal mixing matrix, $\delta\mathbf{M}_{\mathcal{S}',\mathbf{b}_{IJ}}(\mathbf{q},\mathbf{q}')$.

An important observation is that all the matrices in equation (18.11) are block diagonal in baseline $\mathbf{b}$, wavenumber $\mathbf{q}$ and frequency ν, which brings about a significant simplification in calculating the sky estimator:

$$\mathbf{T}_{est}(\mathbf{q},\nu) = \left[\sum_{\mathbf{b}}\frac{|\mathbf{M}_{\mathbf{I},\mathbf{b}}(\mathbf{q},\nu)|^2}{n_{\mathbf{b}}(\mathbf{q},\nu)}\right]^{-1}\left(\sum_{\mathbf{b}'}\frac{\mathbf{M}_{\mathbf{I},\mathbf{b}'}(\mathbf{q},\nu)\mathbf{d}_{\mathbf{b}'}(\mathbf{q},\nu)}{n_{\mathbf{b}'}(\mathbf{q},\nu)}\right). \tag{18.42}$$

where $n_{\mathbf{b}}$ is the noise power scale. Considering that the different measurements are independent and the sky measurement points are uniformly distributed, we can ignore the dependence of the noise on the $\mathbf{q}$ coordinate. In addition, since we assume that all feeds have similar noise power, $n_{\mathbf{b}}$ is actually inversely proportional to the redundancy of the baseline, or the actual number of antenna pairs with the same spatial separation. The actual noise scale of an individual feed is unimportant, since it appears in both the numerator and denominator and thus cancels out. For these reasons, we define a function, $\mathcal{N}(\mathbf{b})$, which gives the number of redundant baselines of $\mathbf{b}$ in an array. Then the equation (18.42) can be rewritten as

$$\mathbf{T}_{est}(\mathbf{q},\nu) = \frac{\sum_{\mathbf{b}'}\mathcal{N}(\mathbf{b}')\mathbf{M}_{\mathbf{I},\mathbf{b}'}(\mathbf{q},\nu)\mathbf{d}_{\mathbf{b}'}(\mathbf{q},\nu)}{\sum_{\mathbf{b}}\mathcal{N}(\mathbf{b})|\mathbf{M}_{\mathbf{I},\mathbf{b}}(\mathbf{q},\nu)|^2}. \tag{18.43}$$

For an instance beam, we have

$$\mathbf{d}_{\mathbf{b}'}(\mathbf{q},\nu) = \mathbf{M}_{\mathbf{I},\mathbf{b}'}(\mathbf{q},\nu)\mathbf{T}_{\mathbf{I}}(\mathbf{q},\nu) + \sum_{\mathcal{S}'}\delta\mathbf{M}_{\mathcal{S}',\mathbf{b}'}(\mathbf{q},\nu)\mathbf{T}_{\mathcal{S}'}(\mathbf{q},\nu) \tag{18.44}$$

so the error map caused by the stochastic beam error is

$$\delta\mathbf{T}_{est}(\mathbf{q},\nu) = \frac{\sum_{\mathcal{S}'}\sum_{\mathbf{b}'}\mathcal{N}(\mathbf{b}')\mathbf{M}_{\mathbf{I},\mathbf{b}'}(\mathbf{q},\nu)\delta\mathbf{M}_{\mathcal{S}',\mathbf{b}'}(\mathbf{q},\nu)\mathbf{T}_{\mathcal{S}'}(\mathbf{q},\nu)}{\sum_{\mathbf{b}}\mathcal{N}(\mathbf{b})|\mathbf{M}_{\mathbf{I},\mathbf{b}}(\mathbf{q},\nu)|^2}. \tag{18.45}$$

For time varying beam errors,

$$\mathbf{d}_{\mathbf{b}'}(\mathbf{q},\nu) = \mathbf{M}_{\mathbf{I},\mathbf{b}'}(\mathbf{q},\nu)\mathbf{T}_{\mathbf{I}}(\mathbf{q},\nu) + \sum_{\mathbf{q}'}\sum_{\mathcal{S}'}\delta\mathbf{M}_{\mathcal{S}',\mathbf{b}'}(\mathbf{q},\mathbf{q}',\nu)\mathbf{T}_{\mathcal{S}'}(\mathbf{q}',\nu) \tag{18.46}$$

and the aliasing effect emerges:

$$\delta\mathbf{T}_{est}(\mathbf{q},\nu) = \frac{\sum_{\mathcal{S}'}\sum_{\mathbf{b}'}\sum_{\mathbf{q}'}\mathcal{N}(\mathbf{b}')\mathbf{M}_{\mathbf{I},\mathbf{b}'}(\mathbf{q},\nu)\delta\mathbf{M}_{\mathcal{S}',\mathbf{b}'}(\mathbf{q},\mathbf{q}',\nu)\mathbf{T}_{\mathcal{S}'}(\mathbf{q}',\nu)}{\sum_{\mathbf{b}}\mathcal{N}(\mathbf{b})|\mathbf{M}_{\mathbf{I},\mathbf{b}}(\mathbf{q},\nu)|^2}. \tag{18.47}$$

18.5.2 1D convolution scenario

In situations where the 2D planar convolution isn't suited for a survey strategy, a one-dimensional convolution, analogous to m-mode analysis (Shaw et al., 2014), might be more applicable. This approach analyzes each declination observation separately. Using the flat-sky approximation, the deconvolution of an unpolarized measurement for a given declination is given by:

$$\mathbf{d}_{\mathbf{b}}(x_{\theta_0},q_\phi) = \int \frac{\mathrm{d}q_\theta}{2\pi}\,\mathbf{M}_{\mathbf{I},\mathbf{b}}(q_\theta,q_\phi)\,\mathbf{T}_{\mathbf{I}}(q_\theta,q_\phi)\exp\left(iq_\theta x_{\theta_0}\right) + \mathbf{n}_{\mathbf{b}}(q_\phi), \tag{18.48}$$

where x_{θ_0} denotes the reference point inclination under the flat sky approximation, reflecting the declination of the pointing center. Unlike the 2D convolution case, this representation is not significantly simplified computationally, requiring matrix operations for error computation, as detailed in Section 18.3.

Chapter 19

Input power spectra

In this section, we describe the statistical models used for simulating the radio sky. We generate the angular power spectrum of the 21cm signal and the unpolarized component of the galactic foreground using **cora** (Shaw et al., 2020). However, lacking better motivated physical model, we derive an analytical covariance model for the polarized foregrounds based on a synchrotron emission model with specific statistical properties explored by Alonso et al. (2014) and Shaw et al. (2015). Below we briefly describe these input power spectrum models of the cosmological signal, unpolarized and polarized galactic foregrounds and the instrumental noise.

19.1 The 21cm signal

The HI brightness temperature field can be factored into a homogeneous temperature field and an overdensity field, $T_b = \overline{T}_b + \delta T_b = \overline{T}_b(1+\delta_{\rm HI})$, where $\overline{T}_b(z)$ can be estimated using Chang et al. (2008), which is about $300\,\mu$K, and $\delta_{\rm HI}$ is expected to be a biased tracer of the dark matter distribution such that

$$\delta_{\rm HI}(\mathbf{k}, z) = \mathrm{F}(\mu, z) D_+(z) \delta_{\rm M}(\mathbf{k}). \tag{19.1}$$

where $\mathrm{F}(\mu, z) = b_{\rm HI} + f\mu^2$ is introduced to account for effects including the bias of the tracer and the redshift space distortion (Kaiser, 1987) on large linear scales; we have ignored the "Fingers of God" effect as it is trivial on large scales. $\mu = k_\parallel/k$ is the cosine of the angle between the line of sight and $\mathbf{k}$, and $D_+(z)$ describing the growth of the perturbations is given by

$$D_+(z) = \frac{5\Omega_{m0}H_0^2 H(a(z))}{2} \int_0^{a(z)} da' [a' H(a')]^{-3} \tag{19.2}$$

which is normalised such that $D_+(0) = 1$. The linear growth rate $f(z)$ is related to the growth function by $f = d\ln D_+ / d\ln a$. $\delta_{\rm M}$ is the current total matter density

perturbation field in the universe, and it defines the matter power spectrum

$$\langle \delta_{\rm M}(\mathbf{k}) \delta^*_{\rm M}(\mathbf{k}') \rangle \equiv (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') P_M(k). \tag{19.3}$$

We shall establish the FoM formalism in observational coordinates, including frequency (ν), 2D angular direction ($\hat{\mathbf{n}}$) and its Fourier dual ($\mathbf{q}$) in the flat-sky limit, which approximates the multipole $l \approx |\mathbf{q}|$. The quantity of interest is the angular power spectrum of the 21 cm brightness temperature in frequency coordinates, which can be calculated in a computationally efficient way in the flat-sky approximation (Datta et al., 2007)

$$C_l(\nu, \nu') = \frac{\overline{T}_b(z)\overline{T}_b(z')D_+(z)D_+(z')}{\pi \chi \chi'} \int_0^\infty {\rm d}k_\parallel \cos(k_\parallel \Delta\chi) {\rm F}(\mu, z) {\rm F}(\mu, z') P_M(k) \tag{19.4}$$

where redshift z and z' correspond to the 21cm redshifted frequency ν and ν'; χ and χ' are the correponding comoving distances and $\Delta\chi = \chi - \chi'$. The angular wavenumber l is related to the wavevector component by $l = \overline{\chi}|\mathbf{k}_\perp|$ and $\overline{\chi}$ is the mean of χ and χ'.

19.2 Unpolarized galactic foregrounds

Our statistical model for the unpolarized galactic foregrounds is based on Santos et al. (2005) and Shaw et al. (2015), while we shall only include the dominant galactic synchrotron emission. The angular power spectrum at a given frequency is

$$C_l^{\rm fg,I}(\nu) = A \left(\frac{l}{100}\right)^{-\alpha} \left(\frac{\nu}{\nu_0}\right)^{-2\beta}. \tag{19.5}$$

The covariance across two different frequencies ν and ν' is assumed to follow a smooth cut-off,

$$C_l^{\rm fg,I}(\nu, \nu') = e^{-\frac{\ln^2(\nu/\nu')}{2\xi^2}} \sqrt{C_l(\nu,\nu) C_l(\nu',\nu')}, \tag{19.6}$$

so that

$$C_l^{\rm fg,I}(\nu, \nu') = A \left(\frac{l}{100}\right)^{-\alpha} \left(\frac{\nu\nu'}{\nu_0^2}\right)^{-\beta} e^{-\frac{\ln^2(\nu/\nu')}{2\xi^2}}. \tag{19.7}$$

Shaw et al. (2015) calibrated the parameter values for low frequencies and high galactic latitudes, $A = 6.6 \times 10^{-3}\,{\rm K}^2, \beta = 2.8, \alpha = 2.8$, and $\xi = 4.0$.

19.3 Polarized galactic foregrounds

As the synchrotron emissions are highly polarized and undergo Faraday rotations in propagation, the observed polarization angle of the signal emitted by the source at distance s is different from its initial polarization angle ϕ_0 by $\Delta\phi = \psi(s, \hat{\mathbf{n}})(c/\nu)^2$,

where ψ is the Faraday rotation measure, given as

$$\psi(s,\hat{\mathbf{n}}) = \frac{e^3}{2\pi(m_e c^2)^2}\int_0^s n_e(s',\hat{\mathbf{n}})B_{\parallel}(s',\hat{\mathbf{n}})\,\mathrm{d}s'. \tag{19.8}$$

Thus, the linear polarized synchrotron intensity written in the form of complex intensity ($I_p = |I_P|e^{i2\phi} = Q + iU$) is:

$$I_p(\nu,\hat{\mathbf{n}}) = \Pi_0\int_0^\infty j_I(s,\hat{\mathbf{n}},\nu)e^{2i\phi_0(s,\hat{\mathbf{n}})}e^{i\psi(s,\hat{\mathbf{n}})x_\nu}\,\mathrm{d}s, \tag{19.9}$$

where $x_\nu = 2(c/\nu)^2$. $j_I(s,\hat{\mathbf{n}},\nu)$ is the emission coefficient and Π_0 is the degree of linear polarization (Rybicki et al., 1986).

Using the Faraday depth $\psi(s,\hat{\mathbf{n}})$ as the line-of-sight (LOS) coordinate, equation (19.9) can be rewritten as

$$I_p(\nu,\hat{\mathbf{n}}) = \int k(\psi,\hat{\mathbf{n}},\nu)e^{i\psi x_\nu}\,\mathrm{d}\psi, \tag{19.10}$$

where $k(\psi_0) = \int \delta^D(\psi(s)-\psi_0)j_I(s)e^{2i\phi_0(s)}\,\mathrm{d}s$ is the collective emission from regions with Faraday depth ψ, and δ^D is Dirac delta function.

The model we shall use is based on Alonso et al. 2014, where the collective emission $k(\psi,\hat{\mathbf{n}},\nu)$ is assumed to be proportional to the number of regions with the Faraday depth ψ; and we also assume the Faraday depths of regions along a LOS are normally distributed with zero mean and some variance $\sigma^2(\hat{n})$. This variance is estimated from an improved map of $\psi_\infty(\hat{n})$, the Faraday depth to the edge of the Milky Way (Oppermann et al. (2012)), by smoothing ψ_∞^2 on a large angular scale. This model also requires that the power-law spectral dependence in any direction be independent of Faraday depth. Then, the model of $k(\psi,\hat{\mathbf{n}},\nu)$ follows

$$k(\psi,\hat{\mathbf{n}},\nu) \propto \left(\frac{\nu}{\nu_{\mathrm{ref}}}\right)^{\alpha(\hat{\mathbf{n}})} e^{-\frac{\psi^2}{2\sigma^2(\hat{\mathbf{n}})}}\mu(\psi,\hat{\mathbf{n}}), \tag{19.11}$$

where $\alpha(\hat{\mathbf{n}})$ is the spectral index, which can be estimated from the Planck Sky Model (Delabrouille et al. (2013)), and ν_{ref} is the reference frequency we shall use for normalization; the angular structure term $\mu(\psi,\hat{\mathbf{n}})$ is assumed to have the same angular structure as the unpolarized emission and it's correlated in Faraday space on scales smaller than some correlation length ξ_ψ:

$$\langle\mu_{lm}(\psi)\mu^*_{l'm'}(\psi')\rangle \propto \delta_{ll'}\delta_{mm'}l^{-\beta}e^{-\frac{1}{2}\left(\frac{\psi-\psi'}{\xi_\psi}\right)^2}, \tag{19.12}$$

where δ is Kronecker delta function.

A useful observation is that the Fourier transform $\tilde{\mu}(x) \equiv \int \mu(\psi)e^{-i\psi x}\mathrm{d}\psi$ is uncor-

related for different values of x and thus the covariance is diagonal

$$\langle \tilde{\mu}_{lm}(x)\tilde{\mu}^*_{l'm'}(x')\rangle \propto \delta_{ll'}\delta_{mm'}\delta^D(x-x')l^{-\beta}e^{-\frac{x^2{\xi_\psi}^2}{2}}. \tag{19.13}$$

The intensity of linear polarization now can be written in terms of $\tilde{\mu}(x)$

$$\mathrm{I_p}(\nu,\hat{\mathbf{n}}) \propto \sigma(\hat{\mathbf{n}})\left(\frac{\nu}{\nu_{\mathrm{ref}}}\right)^{\alpha(\hat{\mathbf{n}})}\int \tilde{\mu}(x,\hat{\mathbf{n}})e^{-\frac{(x_\nu-x)^2}{2\sigma^{-2}(\hat{\mathbf{n}})}}\mathrm{d}x. \tag{19.14}$$

Next we shall derive the angular power spectrum of this model. Because the spectral index map and the Faraday depth map are homogeneous on the scales of interest, we shall factor out the corresponding smooth components for the simple estimates. In other words, we assume statistically homogeneous random field for a limited patch in the sky whose amplitude depends on the spectral index and the Faraday width at the position of the observed patch in the sky. The spherical harmonic coefficients of the polarized intensity are then estimated as

$$\mathrm{I}_{\mathrm{p}lm}(\hat{\mathbf{n}},\nu) \approx \sigma(\hat{\mathbf{n}})\left(\frac{\nu}{\nu_{\mathrm{ref}}}\right)^{\alpha(\hat{\mathbf{n}})}\int \tilde{\mu}_{lm}(x)e^{-\frac{(x_\nu-x)^2}{2\sigma^{-2}(\hat{\mathbf{n}})}}\,\mathrm{d}x. \tag{19.15}$$

Using equation (19.13), we can calculate the power spectrum of polarized intensity for a patch of the sky centred at $\hat{\mathbf{n}}$:

$$C_l^{\mathrm{Ip}}(\nu_1,\nu_2;\hat{\mathbf{n}}) \propto \sigma(\hat{\mathbf{n}})^2\left(\frac{\nu_1\nu_2}{\nu_{\mathrm{ref}}^2}\right)^{\alpha(\hat{\mathbf{n}})} l^{-\beta}\int \exp\left[-\frac{x^2{\xi_\psi}^2}{2}-\frac{(x_{\nu_1}-x)^2}{2\sigma^{-2}(\hat{\mathbf{n}})}-\frac{(x_{\nu_2}-x)^2}{2\sigma^{-2}(\hat{\mathbf{n}})}\right]\mathrm{d}x. \tag{19.16}$$

Note that $k(\psi,\hat{\mathbf{n}},\nu)$ is a complex intensity field with the real and imaginary components corresponding to Stokes Q and U of galactic emissions. We shall further decompose the foreground covariance into coordinate-invariant B-modes and E-modes by assuming

$$C_l^{EE}(\nu_1,\nu_2) = C_l^{BB}(\nu_1,\nu_2) = \frac{1}{2}C_l^{\mathrm{Ip}}(\nu_1,\nu_2) \tag{19.17}$$

and $C_l^{BE}(\nu_1,\nu_2) = C_l^{TE}(\nu_1,\nu_2) = C_l^{TB}(\nu_1,\nu_2) = 0$.

Finally, we have

$$C_l^{PP}(\nu_1,\nu_2;\hat{\mathbf{n}}) \propto l^{-\beta}\sigma^2\left(\frac{\nu_1\nu_2}{\nu_{\mathrm{ref}}^2}\right)^{\alpha}\exp\left[-\frac{\sigma^4((c/\nu_1)^2+(c/\nu_2)^2)^2}{\left(\xi_\psi^2/2+\sigma^2\right)^2}-\sigma^2\left(\frac{c}{\nu_1}\right)^2-\sigma^2\left(\frac{c}{\nu_2}\right)^2\right] \tag{19.18}$$

where $P = E, B$, and remember that the Faraday width σ and the spectral index α are dependent on the specific patch of the sky. A direct observation from the above expression is that the galactic synchrotron emissions at low frequencies are stronger, as shown by the power law term, but also suffering larger Faraday rotation effects which reduces the correlation strength between low frequencies, which is implied by

the exponential term. This general property is consistent with our empirical knowledge.

The proportionality constant in equation (19.18) is still to be determined. Assuming a statistical polarized fraction $f_{\rm pol}$ of the unpolarized emission, the normalization can be found by solving $C_l^{EE}(\nu_{\rm ref},\nu_{\rm ref}) = \frac{1}{2}f_{\rm pol}^2 C_l^{TT}(\nu_{\rm ref},\nu_{\rm ref})$. The value of $f_{\rm pol}$ can be finely calibrated using measured average polarized fraction at particular frequency and galactic latitude, for example using Kogut et al. (2007). However, for our purpose of formalizing the F.o.M formalism, it suffices to assume a fiducial polarization fraction $f_{\rm pol} = 0.5$.

19.4 Instrumental noise

The time series of antenna voltages induced by incident radiation field initially has the form of **incoherent** noise with rapid fluctuations. Although after the bandpass filter the voltage becomes temporally correlated from point to point,[1] it is still a random variable with zero mean in nature. One usual way to measure the power in the fluctuating signal is to average the square of the voltage for a long period $\mathcal{T}$, which is termed *integration time*. Measurement of this kind is called *auto-correlation*. Since the integration time is finite, the averaged output power from the integrator is only an estimate of the true value, with an associated uncertainty, which contributes as *noise*. In the Rayleigh-Jeans regime the rms fluctuations, or noise scale, in the output from the integrator is expressed in temperature by the radiometer equation:

$$\Delta T_{\rm rms} = \frac{T_{\rm sys}}{\sqrt{\Delta\nu\mathcal{T}}} \tag{19.19}$$

where $T_{\rm sys}$ is the system noise temperature and $\Delta\nu$ is the frequency channel width.

While for visibility measurements, one considers the *cross-correlation* between different feeds by digitizing the voltage signals of each feed and then integrating the product of the digital signals over time. The mean squared fluctuations in the amplitude of the integrator output are

$$\Delta T^2_{\rm rms,ij} = \frac{T_{\rm sys,i}T_{\rm sys,j}}{\Delta\nu\mathcal{T}} \tag{19.20}$$

where i, j represent the two elements of the antenna pair. Since the noise at each feed and each frequency channel is independent, the noise between different baseline pairs or different frequencies is uncorrelated. Just like what we have worked out for sky sources, we define the noise power spectrum in visibility correlation:

$$\mathbf{N}_{ij}(\nu,\nu') = \delta_{\nu\nu'}\frac{T_{sys,i}(\nu)T_{sys,j}(\nu')}{t_{tot}\Delta\nu} \tag{19.21}$$

[1]This is manifested as a low frequency modulation of the voltage envelope whose time scale is estimated as the inverse of the frequency channel width, $1/\Delta\nu$.

where δ is Kronecker delta function and $t_{\rm tot}$ is the total integration time for a baseline measurement, which is dependent on the particular survey strategy.

Chapter 20

Discussion

20.1 Generalizing Stokes I Strategies

In this polarized FoM formalism we use the most common Stokes I extraction strategy, i.e. $XX + YY$. This should be easily generalized to other linear Stokes I extraction strategies, such as those discussed in Part IV.

20.2 The aliasing effects

In the previous discussion we categorized general beam uncertainties into 'stochastic beam errors' (constant) and 'time varying beam errors.' The former leads to leakage of the same spatial (transverse) mode polarizations into the Stokes I signal, as shown in equation (18.45). The latter induces not only such polarization leakage, but also leakage of other modes, especially the galactic dominated modes of all polarizations, as shown in equation (18.47). This mode-mixing effect coupled with polarization leakage can be generally referred to as the aliasing effects in mode analysis.

20.2.1 Numerical example

This section provides a numerical example that demonstrates the mixing of transverse modes and the effect of polarization leakage due to beam simulation error.

Toy model of beam error

In our model, we propose that the beam error has separable time and beam angle dependence, such that

$$\delta\mathbf{E}_{Ip}(\mathbf{x}, \mathbf{x_0}) = \mathbf{E}_p(\mathbf{x})\Delta_{Ip}(\mathbf{x_0}). \tag{20.1}$$

Consequently, the resulting Stokes beam error is given by

$$B^{\mathcal{S}}_{pq,\mathbf{b}_{IJ}}(\mathbf{x}, \mathbf{x_0}) = [1 + f_{Ip,Jq}(\mathbf{x_0})]B^{\mathcal{S}}_{pq}(\mathbf{x}), \tag{20.2}$$

with

$$f_{Ip,Jq}(\mathbf{x_0}) = \Delta^*_{Ip}(\mathbf{x_0}) + \Delta_{Jq}(\mathbf{x_0}) + \Delta^*_{Ip}(\mathbf{x_0})\Delta_{Jq}(\mathbf{x_0}). \tag{20.3}$$

We further assume that $\Delta(\mathbf{x_0})$ is a Gaussian function with a diameter of approximately 25°. This represents the duration when the beam error is notably substantial during drift scanning.

Measurement Equation in Fourier Coordinate

Assuming a sufficiently large survey area, such that we can neglect the windowing effect, the Fourier-coordinate measurement equation with beam error becomes:

$$V_{Ip,Jq}(\mathbf{q_0}) = \sum_{\mathcal{S}'=\mathbf{I},\mathbf{E},\mathbf{B},\mathbf{V}} B^{\mathcal{S}'}\left(\mathbf{q_0} + 2\pi\frac{\mathbf{b}_{IJ,\perp}}{\lambda}\right) T_{\mathcal{S}'}(\mathbf{q_0}) + \int \mathrm{d}^2\mathbf{q} f(\mathbf{q_0}-\mathbf{q}) B^{\mathcal{S}'}\left(\mathbf{q} + 2\pi\frac{\mathbf{b}_{IJ,\perp}}{\lambda}\right) T_{\mathcal{S}'}(\mathbf{q}). \tag{20.4}$$

Impact of Beam Uncertainties

We consider an array of four identical HIRAX prime focus reflectors (as shown in Figure 20.1), each equipped with a single feed. Figure 20.2 shows the error in the angular power spectrum of the reconstructed Stokes I sky at 400 MHz due to beam uncertainties. This is illustrated by the equation (18.19). The results reveal a mix of different transverse modes, or ℓ modes, of Stokes I. Moreover, a significant leakage from various polarization modes to the reconstructed Stokes I is evident.

20.3 The convolution approximation

Our discussion, both single-dish analysis and interferometric analysis, are rooted in the convolutional view of measurement, which greatly simplifies the formalism of FoM. This view has limitations. For Stokes I and V (primary) beams, which are invariant as the beams drift in any direction, the idea of a constant scalar field convolving a sky field applies. On the other hand, the primary beams for Stokes Q and U could change from pointing to pointing due the the curvature of the sperical coordinates system. The idea of two-dimensional convolution can only be used if it is assumed that such changes are negligible. The assumption works well under the flat sky approximation, where the observed sky patch is treated as a two-dimensional plane (see Figure 6.5). This ensures a consistent relationship between antenna and celestial sphere coordinates across the survey area. Thus, if the flat-sky approximation is valid, our 2D convolution scenario of the primary beam remains applicable. Conversely, if the beam is not narrow enough, the 2D convolution scenario is limited to a narrow strip.

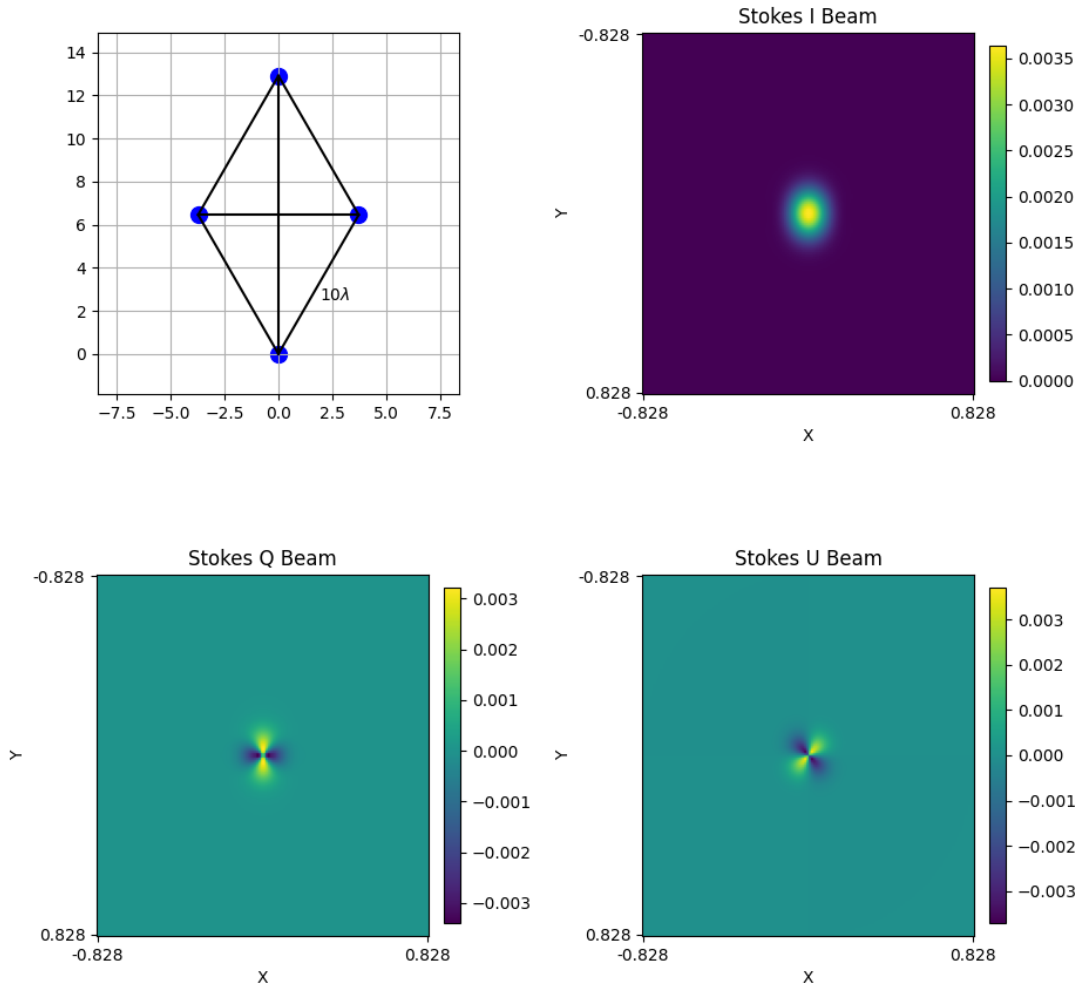

Figure 20.1: Antenna baselines and Stokes beams. Top left: Antenna Array. Top right: Stokes I beam. Bottom left: Stokes Q beam. Bottom right: Stokes U beam. X and Y are beam coordinates projected on the flat sky plane.

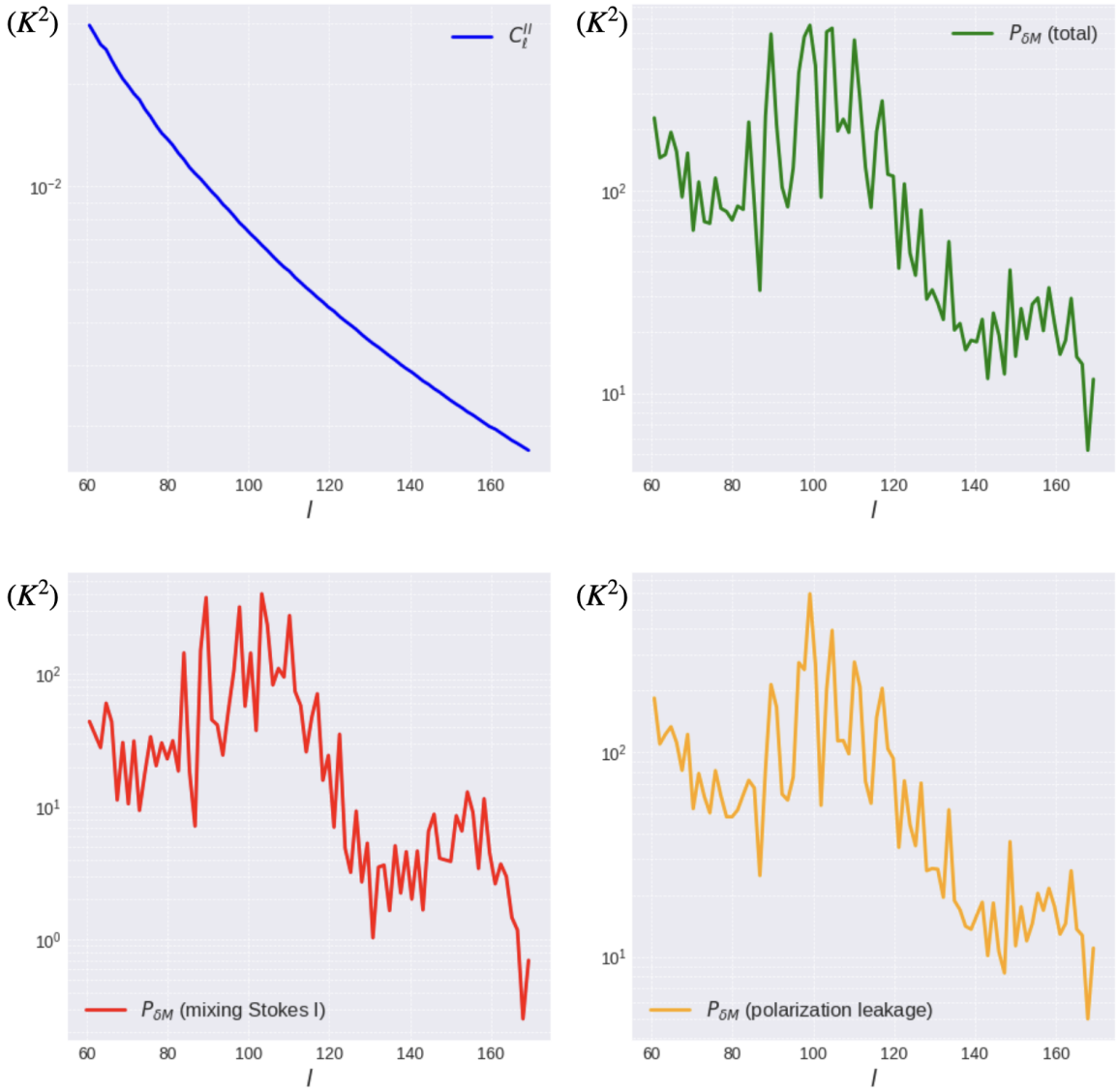

Figure 20.2: Angular power spectrum error of the reconstructed Stokes I sky due to beam uncertainties, as detailed in equation (18.19). *Top Left:* Simulated Stokes I sky angular power spectrum. *Top Right:* Total error in the angular power spectrum caused by beam error. *Bottom Left:* Error contribution from other Stokes I modes. *Bottom Right:* Error contribution from other polarizations.

However, the modulated beam in radio interferometric analysis is a different story. Reorienting the antenna at a different pointing or declination results in a change in the sine or cosine modulation function. In this case, the 2D convolutional view is only applicable to a narrow strip survey. Near the zenith this effect slows down. However, we can also perform the convolution along only one direction, as discussed in section 18.5.2.

20.4 Importance of beam error data

Mitigating or avoiding the effects discussed in this work is critical for 21cm intensity mapping experiments. The FoM formalism is one such beam evaluation strategy, based on the statistical models of stochastic beam uncertainties, that can be used to diagnose and thus minimize the polarization leakage and aliasing effects caused by beam uncertainties. Ideally, there would be a wealth of data on these issues that could be used as inputs to the construction of the FoM. However, such data are currently lacking. We would expect to see more data and simulations on these issues in the future, since the basic strategy for precision observations is that even if we do not have absolute knowledge of the systematics, we at least establish an accurate statistical description of it, which is essential for understanding the confidence in the experimental results.

Part VI

Conclusion

This thesis has examined the role of polarization in 21-cm intensity mapping experiments from several perspectives. We began by reviewing the formalism of polarization measurements and the properties of polarized foregrounds. These discussions set the stage for our primary original work discussed in part IV and part V of this thesis. Specifically, in part II, we formalized the polarization measurements. Traditional approaches employ Jones matrices to link field points of the Stokes sky and the feed far field (refer to Thompson et al. (2017) for details). However, our method in part II provides an analytical function based on the antenna coordinates and pointing. This function can handle the transformation of any field degree of freedom.

In the discussion of polarization foregrounds, we qualitatively describe the polarization and spectral properties of diffuse foregrounds. This provides an understanding of foreground mitigation strategies in 21-centimeter detections based on Stokes I spectral properties, as well as an understanding of why polarization foregrounds can be problematic. Specifically, we provide an analytical angular power spectrum of polarization derived entirely from the polarized foreground model presented in Alonso et al. (2014) (see section 19.3). This provides a convenient choice for related fast simulations.

With this groundwork laid out, our original contributions in part IV and part V deal with the optimal extraction of the Stokes I signal from radio measurements (in part IV) and the quantification of the impact of beam uncertainty on intensity mapping experiments (in part V). In the following sections, we will summarize the role of polarization in each of these topics.

Optimal Stokes I extraction

When considering the Stokes I extraction approach, we frame the measurement within a linear formalism wherein Stokes I extraction involves identifying the right linear combinations of data (or 'data projection' in technical terms). This is typical for many Stokes I measurements. For example, the $XX + YY$ strategy commonly used is essentially a special linear combination of all independent measurements. A more general application can be seen in e.g. Shaw et al. (2015), where data is projected onto the cokernel of the beam of polarizations.

In part IV of this thesis, we argue that the desired procedure is more complicated than a simple projection. Theoretically, we have the flexibility to adjust the role of polarization and set up Stokes I extraction strategies under different polarization scenarios. For example, we can map both the Stokes I and the polarizations using flat priors. We can also choose linear combinations so that there is no polarization leakage, provided the beam matrix is perfectly known. To further reduce variance of Stokes I, additional information is essential. We can use polarization leakage to reduce the variance. This requires knowledge of $\mathbf{C}^{\mathbf{PP}}$, the *a priori* power spectrum of the polarizations.

However, to effectively achieve minimal variance in Stokes I extraction, one must understand the expected, or prior, values for both the Stokes I and Stokes Q/U sky. We have shown how to find Stokes I dominated modes by integrating out the polarization, which is also known as a Wiener filtered map. This is biased but in a way that is completely understood.

Beam figure-of-merit

The polarized FoM formalism characterizes, in a broader sense, the mode mixing and polarization leakage caused by the stochastic time-varying beam uncertainty, as well as the coupling between these two effects, and their impact on intensity mapping experiments. The concepts of 'mode mixing' and 'polarization leakage' are well recognized in the field of 21-cm cosmology, although their interpretations vary between different studies. As an example, Hazelton et al. (2013) discusses a type of mode mixing that occurs when measurements from non-identical baselines are combined. Some other work, e.g. Thyagarajan et al. (2016), studies the mode mixing effects in the context of antenna beam chromaticity. On the other hand, polarization leakage is mostly discussed in the context of a certain but unknown error due to construction tolerances of polarimeters, which could be understood as the polarization mismatch error in interferometry (Thompson et al., 2017). The effect of polarization leakage has been widely studied in EoR tomography (e.g. Jelić et al. (2010); Nunhokee et al. (2017); Martinot et al. (2018)). In the context of the 21-cm global signal, the effect of polarized foreground leakage is studied in Spinelli et al. (2019).

In the polarized FoM framework, we have addressed a more general characterization of instrument errors by perturbing the far-field of the antenna. By describing the beam uncertainty as either a stochastic beam error or a time-varying error, the formalism encapsulates the dependence of the instrument systematics on direction, polarization, frequency, and time (or pointing). By evaluating the intensity map error and the associated power spectrum error due to the beam uncertainty, the formalism calculates the ratio of the cosmic signal to the combined "noise + error" for each Fourier mode. Aggregation of the signal-to-noise ratios for all modes yields a single figure-of-merit. This approach provides an isolated assessment of the beam effectiveness in achieving scientific goals.

Bibliography

N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont et al. Planck 2018 Results. *Astronomy Astrophysics*, 641, 2020.

N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi et al. Planck 2018 Results. *Astronomy Astrophysics*, 641, 2020.

A. Albrecht, P. J. Steinhardt, M. S. Turner, and F. Wilczek. Reheating an Inflationary Universe. *Physical Review Letters*, 48(20), 1982.

D. Alonso, P. G. Ferreira, and M. G. Santos. Fast Simulations for Intensity Mapping Experiments. *Monthly Notices of the Royal Astronomical Society*, 444(4), 2014.

R. A. Alpher, R. Herman, and G. Gamow. Thermonuclear Reactions in the Expanding Universe. *Physical Review*, 74(9), 1948.

M. Amiri, K. Bandura, A. Boskovic, T. Chen, J.-F. Cliche et al. An Overview of CHIME, the Canadian Hydrogen Intensity Mapping Experiment. *Astrophysical Journal Supplement Series*, 261(2), 2022.

C. J. Anderson, N. J. Luciw, Y. C. Li, C. Y. Kuo, J. Yadav et al. Low-amplitude Clustering in Low-redshift 21-cm Intensity Maps Cross-correlated with 2dF Galaxy Densities. *Monthly Notices of the Royal Astronomical Society*, 476(3), 2018.

W. Baade. The Absolute Photographic Magnitude of Supernovae. *Astrophysical Journal*, 88, 1938.

D. J. Bacon, R. A. Battye, P. Bull, S. Camera, P. G. Ferreira et al. Cosmology with Phase 1 of the Square Kilometre Array Red Book 2018: Technical specifications and performance forecasts. *Publications of the Astronomical Society of Australia*, 37, 2020.

N. A. Bahcall and S. P. Oh. The Peculiar Velocity Function of Galaxy Clusters. *Astrophysical Journal*, 462(2), 1996.

N. A. Bahcall, M. Gramann, and R. Cen. The motions of clusters and group of galaxies. *Astrophysical Journal*, 436, 1994.

K. Bandura, G. E. Addison, M. Amiri, J. R. Bond, D. Campbell-Wilson et al. Canadian Hydrogen Intensity Mapping Experiment (CHIME) Pathfinder. In *Ground-based and Airborne Telescopes V*, volume 9145. SPIE, 2014.

N. Barry, G. Bernardi, B. Greig, N. Kern, and F. Mertens. SKA-low Intensity Mapping Pathfinder Updates: deeper 21 cm power spectrum limits from improved analysis frameworks. *Journal of Astronomical Telescopes, Instruments, and Systems*, 8(1), 2022.

B. A. Bassett and R. Hlozek. Baryon Acoustic Oscillations, 2009.

D. Baumann. *Cosmology*. Cambridge University Press, 2022.

R. Beck. Magnetic Fields in the Milky Way and Other Spiral Galaxies. In *How does the Galaxy Work? A Galactic Tertulia with Don Cox and Ron Reynolds*. Springer, 2004.

R. Beck. Magnetic Fields in Spiral Galaxies. *Astronomy and Astrophysics Review*, 24 (1), 2016.

R. Brout, F. Englert, and E. Gunzig. The Creation of the Universe as a Quantum Phenomenon. *Annals of Physics*, 115(1), 1978.

M. Bucher, K. Moodley, C. Pieterse, D. de Villiers, and Z. Zhang. Toward a Beam Quality Figure of Merit for Radio Intensity Mapping. in prep.

P. Bull, P. G. Ferreira, P. Patel, and M. G. Santos. Late-time Cosmology with 21 cm Intensity Mapping Experiments. *Astrophysical Journal*, 803(1), 2015.

S. M. Carroll, V. Duvvuri, M. Trodden, and M. S. Turner. Is Cosmic Speed-up Due to New Gravitational Physics? *Phys. Rev. D*, 70, 2004.

T.-C. Chang, U.-L. Pen, J. B. Peterson, and P. McDonald. Baryon Acoustic Oscillation Intensity Mapping of Dark Energy. *Physical Review Letters*, 100(9), 2008.

T.-C. Chang, U.-L. Pen, K. Bandura, and J. B. Peterson. An Intensity Map of Hydrogen 21-cm Emission at Redshift z ≈0.8. *Nature*, 466(7305), 2010.

E. Chapman and V. Jelić. Foregrounds and Their Mitigation. In *The Cosmic 21-cm Revolution: Charting the first billion years of our universe*. IOP Publishing Bristol, UK, 2019.

X. Chen, F. Wu, H. Shi, J. Zhang, and Y. Wang. The Tianlai 21cm Intensity Mapping Experiment. *IAU General Assembly*, 22, 2015.

E. Churchwell, M. Povich, D. Allen, M. Taylor, M. Meade et al. The Bubbling Galactic Disk. *Astrophysical Journal*, 649(2), 2006.

J. J. Condon and S. M. Ransom. *Essential Radio Astronomy*. Princeton University Press, 2016.

R. J. Cooke, M. Pettini, and C. C. Steidel. One Percent Determination of the Primordial Deuterium Abundance. *Astrophysical Journal*, 855(2), 2018.

C. J. Copi, D. N. Schramm, and M. S. Turner. Big-bang Nucleosynthesis and the Baryon Density of the Universe. *Science*, 267(5195), 1995.

D. Crichton, M. Aich, A. Amara, K. Bandura, B. A. Bassett et al. Hydrogen Intensity and Real-Time Analysis Experiment: 256-element array status and overview. *Journal of Astronomical Telescopes, Instruments, and Systems*, 8(1), 2022.

S. Cunnington, M. O. Irfan, I. P. Carucci, A. Pourtsidou, and J. Bobin. 21-cm Foregrounds and Polarization Leakage: cleaning and mitigation strategies. *Monthly Notices of the Royal Astronomical Society*, 504(1), 2021.

K. K. Datta, T. R. Choudhury, and S. Bharadwaj. The Multifrequency Angular Power Spectrum of the Epoch of Reionization 21-cm Signal. *Monthly Notices of the Royal Astronomical Society*, 378(1), 2007.

M. Davis, J. A. Newman, and S. M. F. A. C. Phillips. The DEEP2 Redshift Survey. In *Deep Fields*. Springer-Verlag, 2003.

J. Delabrouille, M. Betoule, J.-B. Melin, M.-A. Miville-Deschê nes, J. Gonzalez-Nuevo et al. The Pre-launch Planck Sky Model: a model of sky emission at submillimetre to centimetre wavelengths. *Astronomy Astrophysics*, 553, 2013.

C. Dickinson. BINGO-A Novel Method to Detect BAOs Using a Total-power Radio Telescope. *arXiv preprint arXiv:1405.7936*, 2014.

S. Dodelson and F. Schmidt. *Modern Cosmology*. Academic Press, 2020.

B. T. Draine. *Physics of the Interstellar and Intergalactic Medium*. Princeton University Press, 2010.

G. Dvali, G. Gabadadze, and M. Porrati. 4D Gravity on a Brane in 5D Minkowski Space. *Physics Letters B*, 485(1-3), 2000.

G. Efstathiou, J. Bond, and S. D. White. COBE Background Radiation Anisotropies and Large-scale Structure in the Universe. *Monthly Notices of the Royal Astronomical Society*, 258(1), 1992.

A. Einstein and W. de Sitter. On the Relation between the Expansion and the Mean Density of the Universe. *Proceedings of the National Academy of Science*, 18(3), 1932.

D. Eisenstein and M. White. Theoretical Uncertainty in Baryon Oscillations. *Physical Review D*, 70(10), 2004.

D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton et al. Detection of the Baryon Acoustic Peak in the Large-scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophysical Journal*, 633(2), 2005.

G. E. Evans. Antenna Measurement Techniques. *Norwood*, 1990.

H. I. Ewen and E. M. Purcell. Observation of a Line in the Galactic Radio Spectrum: Radiation from galactic hydrogen at 1,420 Mc./sec. *Nature*, 168(4270), 1951.

K. M. Ferriere. The Interstellar Environment of Our Galaxy. *Reviews of Modern Physics*, 73(4), 2001.

W. L. Freedman and L. L. Feng. Determination of the Hubble Constant. *Proceedings of the National Academy of Sciences*, 96(20), 1999.

A. Friedmann. Über die krümmung des raumes. *Z. Phys.*, 10, 1922.

S. R. Furlanetto, S. P. Oh, and F. H. Briggs. Cosmology at Low Frequencies: The 21cm transition and the high-redshift Universe. *Physics Reports*, 433(4), 2006.

A. H. Guth and S.-Y. Pi. Fluctuations in the New Inflationary Universe. *Physical Review Letters*, 49(15), 1982.

C. Haslam, C. Salter, H. Stoffel, and W. Wilson. A 408 MHz All-sky Continuum Survey. II-The atlas of contour maps. *Astronomy and Astrophysics Supplement Series*, 47, 1982.

B. J. Hazelton, M. F. Morales, and I. S. Sullivan. The Fundamental Multi-baseline Mode-mixing Foreground in 21 cm Epoch of Reionization Observations. *Astrophysical Journal*, 770(2), 2013.

M. L. Heger. The Spectra of Certain Class B stars in the Regions 5630A-6680A and 3280A-3380A. *Lick Observatory Bulletin*, 10(337), 1922.

C. M. Hirata and K. Sigurdson. The Spin-resolved Atomic Velocity Distribution and 21-cm Line Profile of Dark-age Gas. *Monthly Notices of the Royal Astronomical Society*, 375(4), 2007.

C. J. Hogan. Big Bang Nucleosynthesis and the Observed Abundances of Light Elements. In N. Turok, editor, *Critical Dialogues in Cosmology*. World Scientific, 1997.

D. W. Hogg. Distance Measures in Cosmology, 2000.

F. Hoyle and R. J. Tayler. The Mystery of the Cosmic Helium Abundance. *Nature*, 203(4950), 1964.

E. Hubble. A Relation Between Distance and Radial Velocity Among Extra-galactic Nebulae. *Proceedings of the national academy of sciences*, 15(3), 1929.

V. Jelić, S. Zaroubi, P. Labropoulos, G. Bernardi, A. G. de Bruyn et al. Realistic Simulations of the Galactic Polarized Foreground: Consequences for 21-cm Reionization Detection Experiments. *Monthly Notices of the Royal Astronomical Society*, 409(4), 2010.

A. Joyce, B. Jain, J. Khoury, and M. Trodden. Beyond the Cosmological Standard Model. *Physics Reports*, 568, 2015.

N. Kaiser. Clustering in Real Space and in Redshift Space. *Monthly Notices of the Royal Astronomical Society*, 227(1), 1987.

D. Kazanas. Dynamics of the Universe and Spontaneous Symmetry Breaking. *Astrophysical Journal*, 241, 1980.

F. Kerr, J. Hindman, and B. Robinson. Observations of the 21 cm Line from the Magellanic Clouds. *Australian Journal of Physics*, 7(2), 1954.

R. Knop et al. Measurements of Ω and Λ from 42 High-redshift Supernovae. *The Astrophysical Journal*, 517(2), 1999.

H. Kodama and M. Sasaki. Cosmological Perturbation Theory. *Progress of Theoretical Physics Supplement*, 78, 1984.

A. Kogut, J. Dunkley, C. L. Bennett, O. Doré, B. Gold et al. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP)* Observations: Foreground Polarization. *Astrophysical Journal*, 665(1), 2007.

T. L. Landecker and R. Wielebinski. The Galactic Metre Wave Radiation: A two-frequency survey between declinations+ 25o and-25o and the preparation of a map of the whole sky. *Australian Journal of Physics, Astrophysical Supplement, Vol. 16, p. 1 (1970)*, 16, 1970.

A. D. Linde. A New Inflationary Universe Scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108(6), 1982.

A. Liu and J. R. Shaw. Data Analysis for Precision 21 cm Cosmology. *Publications of the Astronomical Society of the Pacific*, 132(1012), 2020.

A. Loeb and S. R. Furlanetto. *The First Galaxies in the Universe*, volume 21. Princeton University Press, 2013.

A. Ludwig. The definition of cross polarization. *IEEE Transactions on Antennas and Propagation*, 21(1), 1973.

Z. E. Martinot, J. E. Aguirre, S. A. Kohn, and I. Q. Washington. Ionospheric Attenuation of Polarized Foregrounds in 21 cm Epoch of Reionization Measurements: A demonstration for the HERA experiment. *Astrophysical Journal*, 869(1), 2018.

K. W. Masui, E. R. Switzer, N. Banavar, K. Bandura, C. Blake et al. Measurement of 21 cm Brightness Fluctuations at z $\sim$ 0.8 in Cross-correlation. *Astrophysical Journal*, 763(1), 2013.

A. Meiksin. Detecting the Epoch of First Light in 21-CM Radiation. In *Perspectives on Radio Astronomy: Science with Large Antenna Arrays*, 2000.

P. W. Merrill. Unidentified Interstellar Lines. *Publications of the Astronomical Society of the Pacific, Vol. 46, No. 272, p. 206-207*, 46, 1934.

B. S. Meyer and D. N. Schramm. General Constraints on the Age and Chemical Evolution of the Galaxy. *Astrophysical Journal*, 311, 1986.

P. Mezger and A. Henderson. Galactic H II Regions. I. Observations of their continuum radiation at the frequency 5 GHz. *Astrophysical Journal*, 147, 1967.

V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger. Theory of Cosmological Perturbations. *Physics Reports*, 215(5-6), 1992.

L. Newburgh, K. Bandura, M. Bucher, T.-C. Chang, H. Chiang et al. HIRAX: A Probe of Dark Energy and Radio Transients. In *Ground-based and Airborne Telescopes VI*, volume 9906. SPIE, 2016.

C. D. Nunhokee, G. Bernardi, S. A. Kohn, J. E. Aguirre, N. Thyagarajan et al. Constraining Polarized Foregrounds for EoR Experiments. II. Polarization Leakage Simulations in the Avoidance Scheme. *Astrophysical Journal*, 848(1), 2017.

N. Oppermann, H. Junklewitz, G. Robbers, M. R. Bell, T. Enßlin et al. An Improved Map of the Galactic Faraday Sky. *Astronomy Astrophysics*, 542, 2012.

S. Paul, M. G. Santos, Z. Chen, and L. Wolz. A First Detection of Neutral Hydrogen Intensity Mapping on Mpc Scales at $z \approx 0.32$ and $z \approx 0.44$. *arXiv preprint arXiv:2301.11943*, 2023.

P. J. E. Peebles. Primeval Helium Abundance and the Primeval Fireball. *Phys. Rev. Lett.*, 16, 1966.

P. J. E. Peebles. *The Large-scale Structure of the Universe*, volume 98. Princeton University Press, 1980.

P. J. E. Peebles. *Principles of Physical Cosmology*, volume 27. Princeton University Press, 1993.

W. J. Percival and M. White. Testing Cosmological Structure Formation Using Redshift-space Distortions. *Monthly Notices of the Royal Astronomical Society*, 393 (1), 2009.

S. Perlmutter, G. Aldering, M. D. Valle, S. Deustua, R. Ellis et al. Discovery of a Supernova Explosion at Half the Age of the Universe. *Nature*, 391(6662), 1998.

S. Perlmutter et al. Measuring the Acceleration of the Cosmic Expansion Using Supernovae. *Nobel Lecture*, 2011.

J. R. Pritchard and A. Loeb. 21 cm Cosmology in the 21st Century. *Reports on Progress in Physics*, 75(8), 2012.

V. Radhakrishnan, J. Murray, P. Lockhart, and R. Whittle. The Parkes Survey of 21-centimeter Absorption in Discrete-source Spectra. II. Galactic 21-centimeter observations in the direction of 35 extragalactic sources. *Astrophysical Journal Supplement*, 24, 1972.

S. Riemer-Sørensen and E. S. Jenssen. Nucleosynthesis Predictions and High-Precision Deuterium Measurements. *Universe*, 3(4), 2017.

A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astronomical Journal*, 116(3), 1998.

G. Rybicki, A. Lightman, and H. Paul. Book-Review-Radiative Processes in Astrophysics. *Astronomische Nachrichten*, 307, 1986.

G. B. Rybicki and A. P. Lightman. *Radiative Processes in Astrophysics*. John Wiley & Sons, 2008.

B. Ryden and R. W. Pogge. *Interstellar and Intergalactic Medium*. Cambridge University Press, 2021.

M. Santos, P. Bull, D. Alonso, S. Camera, P. Ferreira et al. Cosmology from a SKA HI Intensity Mapping Survey. In *Advancing Astrophysics with the Square Kilometre Array*, volume 215. SISSA Medialab, 2015.

M. G. Santos, A. Cooray, and L. Knox. Multifrequency Analysis of 21 Centimeter Fluctuations from the Era of Reionization. *Astrophysical Journal*, 625(2), 2005.

S. Sarkar. Big Bang Nucleosynthesis and Physics Beyond the Standard Model. *Reports on Progress in Physics*, 59(12), 1996.

K. Sato. First-order Phase Transition of a Vacuum and the Expansion of the Universe. *Monthly Notices of the Royal Astronomical Society*, 195(3), 1981.

D. N. Schramm. The Age of the Universe-Concordance. *Astrophysical Ages and Dating Methods*, 1990.

J. R. Shaw, K. Sigurdson, U.-L. Pen, A. Stebbins, and M. Sitwell. All-sky Interferometry with Spherical Harmonic Transit Telescopes. *Astrophysical Journal*, 781(2), 2014.

J. R. Shaw, K. Sigurdson, M. Sitwell, A. Stebbins, and U.-L. Pen. Coaxing Cosmic 21 cm Fluctuations from the Polarized Sky Using M-mode Analysis. *Physical Review D*, 91(8), 2015.

J. R. Shaw, R. Nitsche, S. Foreman, and A. Kefala. radiocosmology/cora: 20.5.0, 2020.

O. M. Smirnov. Revisiting the Radio Interferometer Measurement Equation-IV. A generalized tensor formalism. *Astronomy Astrophysics*, 531, 2011.

G. F. Smoot. Galactic Free-free and H-alpha Emission. *arXiv preprint astro-ph/9801121*, 1998.

G. F. Smoot and I. Debono. 21 cm Intensity Mapping with the Five Hundred Metre Aperture Spherical Telescope. *Astronomy Astrophysics*, 597, 2017.

M. Spinelli, G. Bernardi, and M. G. Santos. On the Contamination of the Global 21-cm Signal from Polarized Foregrounds. *Monthly Notices of the Royal Astronomical Society*, 489(3), 2019.

M. Spite and F. Spite. Lithium Abundance at the Formation of the Galaxy. *Nature*, 297(5866), 1982.

A. A. Starobinsky. A New Type of Isotropic Cosmological Models without Singularity. *Physics Letters B*, 91(1), 1980.

M. Strauss. The Large-Scale Velocity Field. In N. Turok, editor, *Critical Dialogues in Cosmology*. World Scientific, 1997.

A. W. Strong, I. V. Moskalenko, and V. S. Ptuskin. Cosmic-ray Propagation and Interactions in the Galaxy. *Annu. Rev. Nucl. Part. Sci.*, 57, 2007.

E. R. Switzer, K. W. Masui, K. Bandura, L.-M. Calin, T.-C. Chang et al. Determination of z $\sim$ 0.8 Neutral Hydrogen Fluctuations Using the 21 cm Intensity Mapping Autocorrelation. *Monthly Notices of the Royal Astronomical Society: Letters*, 434 (1), 2013.

A. R. Thompson, J. M. Moran, and G. W. Swenson. *Interferometry and Synthesis in Radio Astronomy*. Springer Nature, 2017.

N. Thyagarajan, A. R. Parsons, D. R. DeBoer, J. D. Bowman, A. M. Ewall-Wice et al. Effects of Antenna Beam Chromaticity on Redshifted 21 cm Power Spectrum and Implications for Hydrogen Epoch of Reionization Array. *Astrophysical Journal*, 825 (1), 2016.

J. Tinbergen. Astronomical Polarimetry. *Astronomical Polarimetry*, 2005.

G. Torres del Castillo. Spin-weighted Spherical Harmonics and Their Applications. *Revista Mexicana de Física*, 53, 2007.

N. Turok. *Critical Dialogues in Cosmology: Princeton, New Jersey, USA, 24-27 June 1996, in Celebration of the 250th Anniversary of Princeton University*. World Scientific, 1997.

H. Van de Hulst. Radiogolven Uit Het Wereldruim: II. Herkomst der radiogolven; Radiogolven uit het wereldruim: II. Herkomst der radiogolven; Radio waves from space. *Nederlandsch Tijdschrift voor Natuurkunde*, 11, 1945.

F. Villaescusa-Navarro, D. Alonso, and M. Viel. Baryonic Acoustic Oscillations from 21 cm Intensity Mapping: the Square Kilometre Array case. *Monthly Notices of the Royal Astronomical Society*, 466(3), 2017.

F. Villaescusa-Navarro, S. Genel, E. Castorina, A. Obuljen, D. N. Spergel et al. Ingredients for 21 cm Intensity Mapping. *Astrophysical Journal*, 866(2), 2018.

R. V. Wagoner, W. A. Fowler, and F. Hoyle. On the Synthesis of Elements at Very High Temperatures. *Astrophysical Journal*, 148, 1967.

S. D. White, G. Efstathiou, and C. Frenk. The Amplitude of Mass Fluctuations in the Universe. *Monthly Notices of the Royal Astronomical Society*, 262(4), 1993.

R. Wielebinski. A History of Radio Astronomy Polarisation Measurements. *Journal of Astronomical History and Heritage*, 15(2), 2012.

J. P. Wild. The Radio-Frequency Line Spectrum of Atomic Hydrogen and its Applications in Astronomy. *Astrophysical Journal*, 115, 1952.

S. A. Wouthuysen. On the Excitation Mechanism of the 21-cm (Radio-frequency) Interstellar Hydrogen Emission Line. *Astrophysical Journal*, 57, 1952.

E. L. Wright, S. S. Meyer, C. L. Bennett, N. W. Boggess, E. Cheng et al. Interpretation of the Cosmic Microwave Background Radiation anisotropy detected by the COBE Differential Microwave Radiometer. *Astrophysical Journal*, 396, 1992.

J. S. B. Wyithe, A. Loeb, and P. M. Geil. Baryonic Acoustic Oscillations in 21-cm Emission: a probe of dark energy out to high redshifts. *Monthly Notices of the Royal Astronomical Society*, 383(3), 2008.

Z. Zhang and M. Bucher. Optimal Stokes-I Extraction for 21cm Intensity Mapping Experiments. in prep.

Z. Zhang, M. Bucher, C. Pieterse, and D. de Villiers. Polarization Rejection Beam Quality Figure-of-Merit for Radio Intensity Mapping. in prep.

Résumé

Au cours de la dernière décennie, on a assisté à un regain d'intérêt pour l'utilisation de la cartographie d'intensité (IM) à 21 cm comme sonde de la structure à grande échelle. Cette idée tire parti de la raie spectrale optiquement mince de 21 cm qui provient de la transition hyperfine de l'hydrogène atomique neutre (HI), qui a envahi l'univers après la recombinaison. Bien que dans l'univers tardif la réionisation ait détruit la plus grande partie de l'hydrogène neutre, dans les galaxies le HI surdense a un taux de recombinaison, proportionnel à n^2, qui est suffisant pour surmonter le taux d'ionisation, proportionnel à n, dû au rayonnement ionisant de fond. De plus, pour les nuages denses et de grande taille, le blindage doit être pris en compte. Lorsque la température de spin est supérieure à la température du fond diffus cosmologique (CMB), l'émission stimulée surdense domine l'absorption, ce qui conduit à l'émission nette de 21 cm et peut être observé comme la différence entre la température observée et la température du CMB. L'observation de la raie d'émission 21cm décalée vers le rouge trace la distribution tridimensionnelle des galaxies et permet d'étudier le regroupement des halos effondrés sur une large gamme de décalages vers le rouge et d'échelles spatiales. Ces données constituent un moyen puissant de contraindre les modèles cosmologiques.

Comme l'avant-plan galactique est spectralement lisse, il est théoriquement possible d'extraire le signal extragalactique désiré. Cependant, les systématiques instrumentales peuvent entraîner une fuite d'une partie de cet avant-plan spectralement lisse vers des composantes variant rapidement avec la fréquence et ainsi être confondues avec un signal extragalactique. La fuite de polarisation est l'un des défis les plus difficiles à relever car l'avant-plan fortement polarisé subit une rotation de Faraday lorsqu'il se propage dans le MIS en présence du champ magnétique galactique. Bien que le signal cosmologique souhaité soit supposé non polarisé, les observations radio étant basées sur la polarimétrie, une représentation inexacte des mesures de polarisation conduit à une fuite de l'avant-plan polarisé dans le signal. Les avant-plans polarisés ayant une structure spectrale complexe compliqueraient l'élimination des avant-plans qui repose sur la régularité spectrale de l'avant-plan galactique.

Dans cette thèse, nous examinons plusieurs aspects du rôle de la polarisation dans la cartographie de l'intensité à 21 cm. Nous passons en revue les mesures de polarisation du ciel radio, nous formalisons la mesure du signal de tension en termes de diagramme de champ lointain de l'antenne, et nous présentons un modèle de mesures de puissance polarisée.

Nous examinons également la polarisation de l'avant-plan galactique. Les expériences à 21 cm tirent parti de la régularité spectrale du signal de l'avant-plan galactique. Mais ce n'est pas tout. D'une part, les défis liés à l'avant-plan peuvent être couplés à des systématiques instrumentales, ce qui peut conduire à des fuites d'émission polarisée

dans les mesures de Stokes I. Les avant-plans polarisés ont généralement une structure spectrale complexe due à la rotation de Faraday pendant la propagation. Cela crée de sérieuses difficultés pour les stratégies d'atténuation des avant-plans basées sur la régularité spectrale. D'autre part, les expériences à 21 cm ont généralement une gamme dynamique élevée, c'est-à-dire que nous nous intéressons à des contrastes beaucoup plus faibles que l'intensité moyenne. Cependant, la douceur connue du spectre de l'avant-plan, soit par extrapolation à partir du succès de stratégies similaires dans les expériences CMB, soit par suite d'une modélisation théorique, ne fournit pas de preuve directe que l'hypothèse s'appliquerait à la gamme dynamique des expériences à 21 cm. Par conséquent, une étude détaillée de l'émission d'avant-plan est essentielle pour les observations à 21 cm. Le ciel radio diffus à basse fréquence est dominé presque exclusivement par le rayonnement synchrotron relativiste, le rayonnement des électrons accélérés dans le champ magnétique galactique, et l'émission libre, le rayonnement produit par les électrons accélérés dans le potentiel de Coulomb des ions. Les deux processus physiques sont rigoureusement calculés. Comme les ions dans un plasma peuvent apparaître dans n'importe quelle direction du faisceau d'électrons incident, la polarisation zéro est une bonne approximation pour l'émission libre dans le cadre d'une moyenne d'ensemble. La polarisation du rayonnement synchrotron relativiste est plus complexe. Les ondes monochromatiques sont par définition entièrement polarisées. Nous voyons que le rayonnement synchrotron à un seul électron a en général une densité spectrale de puissance non nulle dans les deux directions de polarisation, de sorte que l'onde monochromatique du rayonnement synchrotron à un seul électron est polarisée elliptiquement. Le champ de rayonnement synchrotron à un seul électron est symétrique par rapport au plan de vitesse et d'accélération. Cela indique que deux observateurs situés de part et d'autre du plan observent des polarisations circulaires opposées. Cela suggère que les polarisations circulaires observées pourraient être annulées dans une certaine mesure. La situation idéale est celle où toutes les polarisations circulaires de gauche et de droite s'annulent. Dans ce cas, l'observateur ne reçoit qu'un rayonnement partiellement polarisé linéairement. Nous abordons également les effets de la propagation des ondes électromagnétiques dans la région HII. Dans le milieu interstellaire (MIS) de la Voie lactée, il existe des champs magnétiques omniprésents, connus sous le nom de champs magnétiques galactiques, qui servent de champ magnétique de fond exercé pour le plasma. En raison de l'effet de rotation de Faraday, les ondes à polarisation linéaire sont sujettes à une rotation du plan de polarisation dans les régions HII fermées par le champ magnétique galactique à grande échelle. Nous construisons également un spectre de puissance angulaire pour le ciel polarisé basé sur un modèle du ciel radio polarisé. Dans ce modèle, l'émission collective est supposée être proportionnelle au nombre de régions ayant la profondeur de Faraday ψ ; et nous supposons également que les profondeurs de Faraday des régions le long d'une LOS sont normalement distribuées avec une moyenne nulle et une certaine variance $\sigma^2(\hat{n})$.

Cette variance est estimée à partir d'une carte améliorée de $\psi_\infty(\hat{n})$, la profondeur de Faraday jusqu'au bord de la Voie Lactée, en lissant ψ_∞^2 sur une grande échelle angulaire. Comme la carte de l'indice spectral et la carte de la profondeur de Faraday sont homogènes sur les échelles d'intérêt, nous factorisons simplement les composantes lisses correspondantes pour les estimations simples et nous obtenons un modèle de spectre de puissance angulaire dépendant de la direction.

En utilisant l'approximation linéaire, nous décrivons le processus de mesure et d'analyse des données. Nous définissons les quatre sous-espaces fondamentaux, chacun pouvant être couvert par un ensemble de vecteurs singuliers. Nous montrons qu'après avoir introduit l'a priori du ciel et le spectre de puissance du bruit, on peut caractériser l'intensité du signal et l'échelle du bruit pour chaque "sous-espace". Étant donné le formalisme linéaire de la mesure, nous pouvons l'appliquer à la projection des données, c'est-à-dire obtenir la partie du vecteur de données dans un sous-espace linéaire de l'espace des données. Nous définissons des opérateurs de projection de données qui réalisent un compromis optimal entre plusieurs objectifs de réduction ou d'extraction de données, y compris le rejet de la polarisation, la réduction du bruit et la sélection et l'évitement de degrés de liberté spécifiques du faisceau. Nous proposons également une stratégie optimale d'extraction de Stokes I. En introduisant des spectres de puissance a priori sur le ciel polarisé, nous montrons que trouver les modes dominants de Stokes I appropriés est une question d'intégration de la polarisation.

Enfin, nous développons une figure de mérite pour quantifier l'impact d'une caractérisation inexacte et imprécise du faisceau et de sa complexité sur l'extraction de la carte d'intensité. Il existe plusieurs sources d'erreurs d'estimation du faisceau, telles que les erreurs ou incertitudes dans les simulations CEM, les variations des facteurs environnementaux, etc. En général, nous définissons deux types différents d'incertitudes sur les faisceaux, les "erreurs stochastiques sur les faisceaux" ou les "erreurs sur les faisceaux variables dans le temps", qui se distinguent par le fait qu'elles varient ou non dans le temps. Pour le type constant, à savoir les erreurs stochastiques du faisceau, les ciels polarisés avec tous les paramètres de Stokes des mêmes modes de nombres d'ondes transversales s'infiltreront dans le signal cosmologique. Pour le type variable dans le temps, appelé erreurs de faisceau variables dans le temps, quelque chose d'encore plus grave se produit en plus de la fuite de polarisation : les effets de repliement peuvent provoquer un mélange dans les modes de nombre d'ondes transversales, ce qui peut conduire à une fuite des modes dominés par la galaxie de toutes les polarisations dans les modes à haut RSB pour le signal cosmologique. Il est essentiel d'atténuer ou d'éviter les effets décrits ci-dessus pour les expériences de cartographie de l'intensité à 21 cm. En introduisant une description statistique des systématiques du faisceau, nous modélisons l'impact des fuites de polarisation, nous modélisons l'impact de la fuite de polarisation et du mélange de modes dans la reconstruction I de Stokes. En utilisant des modèles statistiques du ciel et du bruit, nous évaluons le rapport signal/bruit

(RSB) de chaque mode de Fourier avant et après la prise en compte de l'incertitude du faisceau. Une figure de mérite (FoM) est ensuite définie pour classer les systèmes d'alimentation par réflecteur en fonction du facteur par lequel le rapport signal/bruit total pour la détection d'un signal cosmologique est réduit après prise en compte de l'erreur d'incertitude du faisceau. Le formalisme du FoM peut être utilisé pour diagnostiquer et donc minimiser les fuites de polarisation et les effets de repliement causés par les incertitudes du faisceau. Notre discussion, à la fois l'analyse d'une seule coupelle et l'analyse interférométrique, est ancrée dans la vision convolutionnelle de la mesure, ce qui simplifie grandement le formalisme du FoM. Mais comme nous l'avons précisé au début, ce point de vue a des limites. Pour les faisceaux de Stokes I et V (primaires), qui sont invariants lorsque les faisceaux dérivent dans n'importe quelle direction, l'idée d'un champ scalaire constant convoluant un champ du ciel s'applique. En revanche, les faisceaux primaires de Stokes Q et U, et donc E et B, restent constants lorsque les faisceaux sont balayés le long d'une ligne de déclinaison, mais changent lorsque la déclinaison change, dans le système de coordonnées sphériques équatoriales. L'idée de convolution bidimensionnelle ne peut être utilisée que si l'on suppose que ces changements sont négligeables. Mais heureusement, cette hypothèse n'est pas plus forte que l'approximation du ciel plat. Les vecteurs de base des coordonnées de l'antenne à toutes les positions sur ce plan sont liés aux vecteurs de base des coordonnées de la sphère céleste de la même manière linéaire qu'au centre du faisceau. Le balayage du faisceau le long de thêta (colatitude) ou de phi (ascension droite) ne modifie pas l'angle entre la base de l'antenne et la base céleste au centre du faisceau, tant que l'antenne ne tourne pas autour de son propre axe. Ainsi, si l'approximation du ciel plat fonctionne bien, notre vue convolutionnelle bidimensionnelle du faisceau primaire peut toujours être utilisée. Cependant, si le faisceau primaire n'est pas si étroit que l'approximation du ciel plat échoue, nous ne pouvons utiliser la vue convolutionnelle en 2D que pour une bande d'observation très étroite avec une faible largeur en déclinaison. Le faisceau modulé dans l'analyse interférométrique radio est une autre histoire. La réorientation de l'antenne vers un autre pointage ou une autre déclinaison entraîne une modification de la fonction de modulation sinusoïdale ou cosinusoïdale. Dans ce cas, la vue convolutionnelle en 2D n'est applicable qu'à une bande étroite. Au voisinage du zénith, cet effet ralentit. Cependant, nous pouvons également effectuer la convolution dans une seule direction. L'idéal serait de disposer d'une multitude de données sur ces questions d'erreur de faisceau, qui pourraient être utilisées pour la construction du FoM. Toutefois, ces données font actuellement défaut. Nous espérons voir davantage de données et de simulations sur ces questions à l'avenir, car la stratégie de base pour les observations de précision est que même si nous n'avons pas une connaissance absolue des systématiques, nous établissons au moins une description statistique précise de celles-ci, ce qui est essentiel pour comprendre la confiance dans les résultats expérimentaux.