COSMIC TENSIONS

Theoretical Aspects of Astroparticle Physics, Cosmology and Gravitation – 2023

José Luis Bernal
Instituto de Física de Cantabria, CSIC-UC, and Departamento de Física Teórica, Universidad de Cantabria, Av. de Los Castros s/n, 39005 Santander, Spain

Sarah Libanore
Department of Physics, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel

A University of Cantabria publication

These are the lectures notes for the “Theoretical Aspects of Astroparticle Physics, Cosmology and Gravitation” school in 2023 at Galileo Galilei Institute for Theoretical Physics in Florence. This course, taught by José Luis Bernal and Sarah Libanore, focuses on cosmic tensions, and it is structured in 5 lectures and 2 applied sessions. The Applied Sessions are mainly meant for discussion and brain storming, so to increase the students’ awareness and understanding of the physical quantities, processes and implications of what the main lectures present from a more theoretical point of view. The notes represent a guideline through the reasoning, but they can possibly be not exhaustive with respect to what is discussed during the meeting.

These notes contain significantly more detail than what was covered in the lectures, especially regarding derivations and expressions in synchronous gauge. Throughout, natural units $c = \hbar = 1$ will be used unless otherwise stated and we will use the *mostly positive* signature, i.e., for which the Minkowski space is determined by a diagonal metric given by $\{-1, 1, 1, 1\}$. As a disclaimer, the references explicitly cited in these notes are far from being a representative sample of the community’s work in the field. Always that it is possible, reviews will be referenced so that the reader can find the relevant references in them.

CONTENTS IN BRIEF

CONTENTS

CHAPTER 1

LECTURE 0: BASICS

In this initial chapter we provide a long discussion on some of the basics of cosmological perturbation theory, which are required to fully understand the content of the most formal bits of the course. Of course, this preliminary chapter provides a lot of detail, in particular for some derivations, but we prefer to include it in case some of the participants in the school want to deepen in the study of cosmological perturbations or to have as a focused reference which uses the same notation as the rest of the course.

Cosmology deals with the nature and evolution of the Universe and its components in a statistical manner, therefore it is at its core the application of general relativity and statistical mechanics, combined with the astrophysics that drive the physics of the tracers that we can observe. However, we will ignore the latter for the time being and focus on the statistical properties of matter and radiation in the Universe, and how they affect and are affected by gravity,[1] which is the only relevant long-range force that we will consider.

Some approaches prefer to treat some of the components of the Universe as fluids. Instead, we will treat each component from a statistical point of view: since we do not care about the behavior of individual particles, but their collective properties, all the information that we need is enclosed in the distribution function f of the number of particles N in an infinitesimal phase-space element around position $\boldsymbol{x}$ and momentum $\boldsymbol{p}$, such as

$$N(\boldsymbol{x}, \boldsymbol{p}, t) = f(\boldsymbol{x}, \boldsymbol{p}, t)\mathrm{d}^3\boldsymbol{x}\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\,, \tag{1.1}$$

where we assume that the number of particles is large enough for f to approach the continuous limit. The $(2\pi)^3$ factor appears because by Heisenberg's principle, no particle can be localized into a region of phase space smaller than $(2\pi\hbar)^3$, which makes it the size of the fundamental element. The equations describing the evolution of f as function of time and phase-space coordinates are the Boltzmann equations.

As we will see, the Boltzmann Equations already include the continuity and Euler equation that are usually applied to describe the dynamics of fluids for cosmological perturbation theory, but in addition provide a framework to straightforwardly include any additional interaction between the particles of the fluid or between different components of matter and radiation. Furthermore, some components impact cosmological perturbations beyond their density, velocity and anisotropic stress (the monopole, dipole and quadrupole of the phase-space distribution), and higher-order moments, not considered in the continuity and Euler equations, must be taken into account. This is why we prefer to develop the cosmological perturbation theory with full generality, and then specify the properties for each component.

1.1 The FLRW metric and the Einstein equations

For most of the derivations and discussions in these lectures we will assume a flat spatial sector in our metric, i.e., an Euclidean Universe. Special relativity is described by the Minkowski metric $g_{\mu\nu} = \gamma_{\mu\nu} = \mathrm{Diag}\{-1, 1, 1, 1\}$. For an expanding Universe, if we limit our description to the background, we have the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, given by

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\mathrm{d}x^i\mathrm{d}x_i\,, \tag{1.2}$$

where $a(t)$ is the scale factor, which evolves with time.

[1]Studies of general relativity usually refer as *matter* to all components that appear in the right-hand side of the Einstein equations, i.e., including also radiation. We may use that language in certain situations.

The shortest (or extremal) path between two points in space time is defined by the geodesic. General relativity states that this is the path that a particle follows in the absence of any force apart from gravity. Therefore, we can understand the geodesic as a generalization of Newton's law with no forces to general relativity. Defining the parameter λ as the evolution parameter of the system, which monotonically increases along the particle's path, the geodesic equation can be written as

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{dx^\beta}{\mathrm{d}\lambda} = 0\,. \tag{1.3}$$

$\Gamma^\mu_{\alpha\beta}$ in the equation above is the Christoffel symbol, which is transformation tensor (more exactly, the affine connection) between the Newtonian, Euclidean case and its relativistic generalization. Expressed as function of the metric, they are given by

$$\Gamma^\mu_{\alpha\beta} = \frac{g^{\mu v}}{2}\left[\frac{\partial g_{\alpha v}}{\partial x^\beta} + \frac{\partial g_{\beta v}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^v}\right]\,, \tag{1.4}$$

where you can note that the Christoffel symbols are symmetric for the low indices. Therefore, it is evident that calculating the Christoffel symbols is a basic step for any calculation.

However, the Universe is not empty, and it contains matter and radiation which determines its evolution. The relation between the metric and the constituents of the Universe is contained in the Einstein equations:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}\,, \tag{1.5}$$

where $G_{\mu\nu} = R_{\mu\nu} - Rg_{\mu\nu}/2$ is the Einstein tensor, which depends on the Ricci tensor and the Ricci scalar, Λ is the cosmological constant, G is the gravitational constant and $T_{\mu\nu}$ is the stress-energy tensor (or ten energy-momentum tensor). The Ricci tensor is given by

$$R_{\mu\nu} = \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\lambda} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\rho}\,, \tag{1.6}$$

and the Ricci scalar is the contraction of this tensor: $R \equiv g^{\mu\nu}R_{\mu\nu}$. For the FLRW metric, all Christoffel symbols are 0 except for

$$\Gamma^0_{ij} = \delta_{ij}\dot{a}a\,, \qquad \Gamma^i_{0j} = \delta_{ij}H\,, \tag{1.7}$$

where the dot denotes a time derivative, δ_{ij} is the Kronecker delta, and we have defined the Hubble parameter $H \equiv \dot{a}/a$. Similarly, the non-vanishing components of the Ricci tensor and the Ricci scalar are given by

$$R_{00} = -3\frac{\ddot{a}}{a}\,, \qquad R_{ij} = \delta_{ij}(2\dot{a}^2 + a\ddot{a})\,, \qquad R = 6\left(\frac{\ddot{a}}{a} + H^2\right)\,. \tag{1.8}$$

Focusing again only in the background and mean quantities, the stress-energy tensor only features the mean density $\bar{\rho}$ and pressure $\bar{P}$: there cannot be any mean net momentum or velocity since that would break the isotropy of the Universe postulated by the cosmological principle. Therefore, in the isotropic smooth Universe the stress-energy tensor is a diagonal tensor given by $T^\mu_\nu = \mathrm{Diag}\{-\bar{\rho}, \bar{P}, \bar{P}, \bar{P}\}$. As this tensor describes the energy and momentum of the components of the Universe, it must respect the local energy and momentum conservation:

$$\nabla_\mu T^\mu_\nu \equiv \frac{\partial T^\mu_\nu}{\partial x^\mu} + \Gamma^\mu_{\alpha\mu} T^\alpha_\nu - \Gamma^\alpha_{\nu\mu} T^\mu_\alpha = 0. \tag{1.9}$$

Now we have all the pieces to write the Einstein equations correspondent to the Euclidean FLRW metric. First, we take the time-time component of the equation and find

$$H^2 = \frac{8\pi G}{3}\bar{\rho}\,, \tag{1.10}$$

where $\bar{\rho}$ is the mean energy density of all components (including the cosmological constant). This equation is known as the first Friedmann equation. However, it may be more familiar if we rewrite it in terms of the density parameter of each component s, $\Omega_s \equiv \bar{\rho}_s/\rho_{\rm crit}$, where $\rho_{\rm crit} \equiv 3H_0^2/8\pi G$ is the critical density of the Universe today. Thus,

$$H^2 = H_0^2 \sum_s \Omega_s (a)^{-3(1+w_s)}\,, \tag{1.11}$$

where $w_s \equiv \bar{P}_s/\bar{\rho}_s$ is the equation of state parameter of each component. If we consider a curved Universe, we can define $\Omega_K = 1 - \sum_s \Omega_s$ and add a term $\Omega_K a^{-2}$ to the right-hand side of the equation above.

Taking the trace of the Einstein equations we get to

$$-\frac{3\ddot{a}}{a} = 4\pi G\left[3\bar{P} + \bar{\rho}\right] - \Lambda, \tag{1.12}$$

which is known as the second Friedmann equation.

1.2 Boltzmann Equations

A system of particles is statistically determined by its distribution function f in phase space, so we just need an equation that describe its evolution. Neglecting for now any interaction between particles (e.g., scatter, decays, annihilation, etc), the total number of particles must be conserved. This case is referred to as 'collisionless' in the context of the Boltzmann equations. Therefore, the *total* (rather than *partial*) time derivative of the distribution function must vanish:

$$\frac{\mathrm{d}f(\boldsymbol{x},\boldsymbol{p},t)}{\mathrm{d}t} = 0\,; \qquad \text{where } \frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \dot{\boldsymbol{x}}\cdot\boldsymbol{\nabla}_x + \dot{\boldsymbol{p}}\cdot\boldsymbol{\nabla}_p\,, \tag{1.13}$$

where dot denote time derivatives and the subscript of the gradients denote the arguments they must be taken with respect to. The forces driving the problem at hand are included substituting the equations of motion in the expression above. But before that, we need to generalize this expression to the case of an expanding Universe.

One of the main benefits of working in terms of the distribution function f is that we can use it to derive all macroscopic properties of the particles under study. In all generality, the relativistic energy-momentum tensor is

$$T^{\mu}_{\nu}(\boldsymbol{x},t) = \frac{g_*}{\sqrt{-\det(g_{\alpha\beta})}} \int \frac{\mathrm{d}P_1\mathrm{d}P_2\mathrm{d}P_3}{(2\pi^3)} \frac{P^{\mu}P_{\nu}}{P^0} f(\boldsymbol{x},\boldsymbol{p},t)\,, \tag{1.14}$$

where g_* accounts for all the degenerate particle state that are described by f (e.g., $g_* = 2$ for a particle with spin 1/2) and P^{μ} is the four-momentum, defined in terms of the affine parameter of the geodesic with λ (to avoid confusion with the shear stress σ, which will be introduced at the end of this chapter) as

$$P^{\mu} \equiv \frac{\mathrm{d}x^{\mu}}{\mathrm{d}\lambda}\,. \tag{1.15}$$

Equation (1.14) shows that the energy-momentum tensor gives the current density of the four-momentum carried by the particles with distribution function f. The momentum integral over f gives you the number density; weighted by P_{ν} it gives you the four-momentum density; and additionally weighted by the four-velocity P^{μ}/P^0 gives you the current density of the four-momentum. Finally, the prefactor is a geometric factor to ensure the conservation of the energy-momentum tensor: $\nabla_{\mu}T^{\mu}_{\nu} = 0$.

We will first consider a smooth Universe, expanding according to the FLRW metric. However, f still depends on a six-dimensional phase space: we will track time separately as before, and we can express P^0 as function of $\boldsymbol{p}$ using the norm p of the three-momentum and the mass-shell constraint. Then, for the FLRW metric, we have

$$\left(P^0\right)^2 \equiv E^2 = p^2 + m^2\,. \tag{1.16}$$

Furthermore, it is convenient to separate the dependence on $\boldsymbol{p}$ into a dependence on its magnitude p and the unitary vector $\hat{p}^i = \hat{p}_i$ which determines its direction and satisfies $\delta_{ij}\hat{p}^i\hat{p}^j = 1$.[2] Since $\hat{p}^i$ is expected to be proportional to P^i, such as $P^i = \mathcal{C}\hat{p}^i$; then

$$p^2 = g_{ij}P^iP^j = g_{ij}\hat{p}^i\hat{p}^j\mathcal{C}^2 = a^2\delta_{ij}\hat{p}^i\hat{p}^j\mathcal{C}^2 \Rightarrow \mathcal{C} = \frac{p}{a} \Rightarrow P^i = \frac{p}{a}\hat{p}^i\,, \tag{1.17}$$

and we can interchange always P^i by p and $\hat{p}^i$. Therefore, we can generalize $f(\boldsymbol{x},\boldsymbol{p},t) = f(x^i,\hat{p}^i,p,t)$ and express the Boltzmann Equation as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}\frac{\mathrm{d}x^i}{\mathrm{d}t} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial f}{\partial \hat{p}^i}\frac{\mathrm{d}\hat{p}^i}{\mathrm{d}t} = 0\,. \tag{1.18}$$

[2] In general, we will use hats as the notation to denote unitary 3-vectors.

This is the most general expression of the Boltzmann equation in the absence of interactions between particles. In the rest of the section we will discuss a specific simple case and discuss generally the source term that encloses the microphysics determining the particle interactions.

1.2.1 Boltzmann Equation in FLRW

Let us specify the Boltzmann Equation for a smooth expanding Universe, as the one described by the FLRW. In this scenario, the direction of the momentum does not change, hence we can drop the last term in Eq. (1.18). The term than depends on $\partial f/\partial x^i$ could also be dropped (the background is homogenous and isotropic), but it is easy to handle and will be useful once we add perturbations. We need to obtain then the values of the total derivatives of x^i, and p with respect to time. Using Eq. (1.15), we get

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t} = P^i \frac{1}{P^0} = \frac{p}{E}\frac{\hat{p}^i}{a}\,. \tag{1.19}$$

In order to obtain the total derivative of p with respect to time we start from the time component of the geodesic,

$$\frac{\mathrm{d}P^0}{\mathrm{d}\lambda} = -\Gamma^0_{\alpha\beta}P^\alpha P^\beta = -a^2 H\delta_{ij}P^iP^j\,, \tag{1.20}$$

where the last equality relies on the Christoffel symbols for the FLRW metric. Since $P^0 = \mathrm{d}t/\mathrm{d}\lambda$, we have (multiplying and deriving by $\mathrm{d}t$)

$$P^0\frac{\mathrm{d}P^0}{\mathrm{d}t} = p\frac{\mathrm{d}p}{\mathrm{d}t} = -Hp^2 \rightarrow \frac{\mathrm{d}p}{\mathrm{d}t} = -Hp\,, \tag{1.21}$$

where the first equality is obtained from applying the time derivative to Eq. (1.16) in the form $\mathrm{d}(P^0)^2/\mathrm{d}t = 2P^0\mathrm{d}(P^0)/\mathrm{d}t = \mathrm{d}(E^0)^2/\mathrm{d}t = 2p\mathrm{d}p/\mathrm{d}t$. The equation above shows that the physical momentum of any particle decays as $1/a$ in a smooth expanding Universe. Then the collisionless Boltzmann equation in an unperturbed expanding Universe is given by

$$\frac{\partial f}{\partial t} + \frac{p}{E}\frac{\hat{p}^i}{a}\frac{\partial f}{\partial x^i} - Hp\frac{\partial f}{\partial p} = 0\,. \tag{1.22}$$

Similarly, we can derive the evolution of the number density, since it is the momentum integral of f, as discuss above. For a homogeneous Universe (i.e., $\partial f/\partial x^i = 0$), and integrating by parts the momentum component as

$$H\int\frac{\mathrm{d}^2\Omega_\mathrm{p}}{(2\pi)^3}\int_0^\infty p^2\mathrm{d}pp\frac{\partial f}{\partial p} = -3H\int\frac{\mathrm{d}^2\Omega_\mathrm{p}}{(2\pi)^3}\int_0^\infty p^2\mathrm{d}pf = -3Hn\,, \tag{1.23}$$

where Ω_p is the solid angle for the momentum vector and we have used that for any regular distribution function fp^3 vanishes at $p=0$ and $p=\infty$, we find

$$\frac{\mathrm{d}n}{\mathrm{d}t} + 3Hn = 0\,. \tag{1.24}$$

This is, in the absence of collisions, the number density decays as a^{-3} in a homogeneous expanding Universe. However, collisions can change this behaviour, as well as the evolution of the distribution function. We briefly introduce below the collision term.

1.2.2 Collision terms

So far we have studied the evolution of the distribution function for particles that do not interact between them or have any interaction with other components beyond long-range forces (e.g., gravity). However, when these conditions do not apply there is a source term in the Boltzmann equation, called the collision term.[3] As it is expected, the collision term depends on the actual phase-space distribution of the particles involved; hence, in general, the Boltzmann equation becomes

$$\frac{\mathrm{d}f}{\mathrm{d}t} = C[f]\,. \tag{1.25}$$

In order to show in a simple example how to derive the collision term, let us consider a reaction where particles of type (1) and (2) interact to form particles of type (3) and (4):

$$(1)_{\boldsymbol{p}} + (2)_{\boldsymbol{q}} \longleftrightarrow (3)_{\boldsymbol{p}'} + (4)_{\boldsymbol{q}'}\,, \tag{1.26}$$

where the subscripts denote each particle's momenta.[4] Of course, the reaction conserves energy and momentum, and each particle has its own distribution function f_s, with some states that can be degenerate (e.g., often in cosmology, the spin does not play an active role, hence instead of tracking it directly, we weight the distribution function with a suitable degeneracy weight g_*).

We assume that this reaction is local, e.g., the reaction occurs at $(\boldsymbol{x}, t)$ and we only need to determine the momenta. Furthermore, we need to compute the collision term for each independent particle type (which most likely will couple the evolution equations for the fours types of particles).

At the end of the day, the collision term (say, for particles of type 1), as a source term, accounts for all particles that get scattered away from $\boldsymbol{p}$ by the forward reaction (subtract them from $f_1(\boldsymbol{x}, \boldsymbol{p}, t)$) and all particles that get scattered to $\boldsymbol{p}$ by the backward reaction (add them to $f_1(\boldsymbol{x}, \boldsymbol{p}, t)$). The forward and backward reaction rates are determined by the scattering amplitude $|\mathcal{M}|^2$, which can be computed using Feynman diagrams, and the number of particles of each type with the momenta required. In this case we have the

[3]In the context of the Boltzmann equation, the effect of direct particle interactions is referred to as 'collisions', and it is a way to describe the microphysics driving the particle interaction in an effective statistical way.

[4]This reaction can be of scattering and annihilation, depending on the nature of particles 3 and 4 with respect to particles 1 and 2. The derivation of this subsection can straightforwardly be extended to other cases involving a different number of particles.

products $f_1(\boldsymbol{p})f_2(\boldsymbol{q})$ and $f_3(\boldsymbol{p}')f_4(\boldsymbol{q}')$ for the forward and backward reactions. We need to account for stimulated emission (i.e., Bose enhancement) and Pauli exclusion principle (i.e., Pauli blocking), too, which amounts to include factors of $(1 \pm f_3(\boldsymbol{p}))(1 \pm f_4(\boldsymbol{q}'))$ to the forward reaction (and equivalently to the backward reaction), depending on whether the particle involved is a fermion or a boson. What matters is the occupation of the phase-space element in the *result* state from each reaction; this is why they are interchanged. If the particle is a boson the reaction is enhanced, since bosons occupying the same state are favored, while if the particle is a fermion, if a specific state is occupied the reaction cannot happen. Finally, the conservation of momentum and energy is enforced using corresponding Dirac delta functions.

In order to consider the whole phase space in position $(\boldsymbol{x}, t)$ that affects particle 1 with momentum $\boldsymbol{p}$ we need to integrate over all momenta of particles 2, 3, and 4. However, there is a small subtlety: in a relativistic setting, phase-space integrals are four-dimensional (three momentum components and the energy), but energy and momentum are related by the mass-shell constraint, therefore

$$\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} \int \mathrm{d}E \delta_{\mathrm{D}}^{(1)}(E^2 - p^2 - m^2) = \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} \int \mathrm{d}E \frac{\delta_{\mathrm{D}}^{(1)}\left(E - \sqrt{p^2 + m^2}\right)}{2E} , \tag{1.27}$$

which adds a factor of $1/2E$ after integrating over the energy.

Taking all these considerations into account, the collision term becomes

$$\begin{aligned} C[f_1(\boldsymbol{p})] = & \frac{1}{2E_1(p)} \int \frac{\mathrm{d}^3\boldsymbol{q}}{(2\pi)^3 2E_2(q)} \int \frac{\mathrm{d}^3\boldsymbol{p}'}{(2\pi)^3 2E_3(p')} \int \frac{\mathrm{d}^3\boldsymbol{q}'}{(2\pi)^3 2E_4(q')} |\mathcal{M}|^2 \times \\ & \times (2\pi)^4 \delta_{\mathrm{D}}^{(3)}(\boldsymbol{p} + \boldsymbol{q} - \boldsymbol{p}' - \boldsymbol{q}') \delta_{\mathrm{D}}^{(1)} \left[E_1(p) + E_2(q) - E_3(p') - E_4(q')\right] \times \\ & \times \{f_3(\boldsymbol{p}')f_4(\boldsymbol{q}') \left[1 \pm f_1(\boldsymbol{p})\right] \left[1 \pm f_2(\boldsymbol{q})\right] - \\ & \quad - f_1(\boldsymbol{p})f_2(\boldsymbol{q}) \left[1 \pm f_3(\boldsymbol{p}')\right] \left[1 \pm f_4(\boldsymbol{q}')\right]\} . \end{aligned} \tag{1.28}$$

1.3 Perturbed Universe

So far, we have considered only a smooth expanding Universe described by the FLRW metric. This is enough to study the background expansion and thermal history of the Universe, but the Universe has small inhomogeneities that grow over time and host the galaxies and large scale structure that we observe today. Fortunately to us, these inhomogeneities are very small, which allows us to treat them perturbatively. In particular, the linear approximation will be enough except for the small scales at late times.

As we have discussed previously in the context of a smooth Universe, the metric perturbations and the perturbations in the phase-space distributions of the matter components are coupled. Metric perturbations produce perturbations in the background properties of matter and radiation, which in turn affect the metric perturbations, as expected from the energy-momentum tensor term in the Einstein equations and the presence of gravity (through the time derivatives of position and momentum) in the Boltzmann equations. Therefore, we need to treat perturbatively both set of equations to derive the system that describes the evolution of cosmological perturbations.

In what follows we will assume a flat Universe described by the ΛCDM model and where general relativity is an accurate description of gravity, with dark matter, baryons, neutrinos and photons. Also, for convenience, we will work with the conformal time τ, which is related with the physical time through $\mathrm{d}t = a\mathrm{d}\tau$. Derivatives with respect to conformal time are denoted with a prime, and $a'/a = \mathcal{H}$ is the conformal Hubble parameter. Using the conformal time, the FLRW metric satisfies

$$\mathrm{d}s^2 = a^2(\tau)\left[-\mathrm{d}\tau^2 + \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\right] . \tag{1.29}$$

Consider now a small perturbation to this metric, which can be written in full generality as $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$, where the bar denotes background values. Since all these perturbations are very small, we can restrict our work to linear order in perturbation theory. Therefore, we can consider the perturbations a three-tensors and raise and lower indices in the spatial indices always with Kronecker delta (this is not the case for four-vector indices).

The perturbation (actually, any tensor) can be decomposed in scalar, vector and tensor contributions. The decomposition theorem is a very important result in general relativity (which we will not prove here, can be found in Ref. (1)), which states that perturbations of each type evolve independently at linear order. Taking this into account, let us express the perturbed metric as

$$\begin{aligned} g_{00} &= -a^2(\tau)\left\{1 + 2\Psi(\boldsymbol{x},\tau)\right\} , \qquad g_{0i} = a^2(\tau) w_i(\boldsymbol{x},\tau) , \\ g_{ij} &= a^2(\tau)\left\{\left[1 + 2\Phi(\boldsymbol{x},\tau)\right]\delta_{ij} + \chi_{ij}(\boldsymbol{x},\tau)\right\} , \end{aligned} \tag{1.30}$$

where Ψ and Φ are scalars, w_i is a vector and χ_{ij} is a symmetric trace-free tensor ($\delta^{ij}\chi_{ij} = 0$); χ_{ij} can be taken to be traceless since any trace can be reabsorbed into Φ.[5] We will make use of the scalar-vector-tensor decomposition theorem and consider only scalar perturbations.

w_i can be decomposed in a curl-free and divergence-free components in such a way that it depends on a scalar and a transverse vector. In turn, χ_{ij} can be decomposed in a longitudinal, solenoidal and traceless-transverse parts, which involve a scalar, a transverse vector and a trace-free divergence-free tensor. Therefore, we have 4 scalar perturbations (4), 2 transverse vector perturbations ($2\times 2 = 4$) and a trace-free divergence-free tensor (2) for a total

[5]A warning regarding different references is in place regarding this decomposition and the derivation below. Different references use different conventions (signature of the metric, signs of the perturbations, notation of the perturbations, etc), which may impact final expressions in factors and signs for each of the perturbations.

of 10 (4+4+2) degrees of freedom. This means that the metric perturbations from Eq. (1.30) are not uniquely defined (metric perturbations must have 6 degrees of freedom and we counted 10 above) and depend on the coordinate choice. In general relativity, the choice of coordinates and fixing of degrees of freedom is called *gauge choice*. Any time a metric is written, a time slicing is chosen and specific spatial coordinates are defined within it.

A suitable gauge eases significantly the computations but a poor choice may complicate them and even introduce spurious perturbations can arise. This is why gauge-invariant quantities are so important: actually any cosmological observable must be gauge invariant, since physics cannot depend on the choice of coordinates. In any case, transforming between gauge with ease is very important for almost any cosmological computation.

Two of the most used gauges for cosmological perturbation theory are the synchronous and the conformal Newtonian gauges (also called in some contexts longitudinal gauge). The conformal Newtonian gauge (we will refer to this gauge as simply Newtonian from now on) is a particularly simple gauge, and we will use it throughout this study. This gauge has the drawback that is limited to only scalar perturbations, although it has been generalized to tensor perturbations too. In this gauge the metric is given by

$$\mathrm{d}s^2 = a^2(\tau)\left\{-(1+2\Psi)\mathrm{d}\tau^2 + (1+2\Phi)\,\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\right\}\,, \tag{1.31}$$

which leaves the metric tensor diagonal. In this case, the metric perturbations are Ψ and Φ.

1.3.1 Fourier-space computations

Before going deeper in the derivation of the evolution equations, let us step back to discuss the benefits of working in Fourier space (rather than in configuration space), determined by wavevectors $\boldsymbol{k}$. As reference, we follow the Fourier convention

$$\tilde{f}(\boldsymbol{k}) = \int \mathrm{d}^3\boldsymbol{x}\; f(\boldsymbol{x})e^{-i\boldsymbol{k}\boldsymbol{x}}, \qquad f(\boldsymbol{x}) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\;\tilde{f}(\boldsymbol{k})e^{i\boldsymbol{k}\boldsymbol{x}}\,, \tag{1.32}$$

where the tilde denotes Fourier-space functions.[6] Spatial derivatives simplify significantly in Fourier space:

$$\partial_i F(\boldsymbol{x},t) = ik_i\tilde{F}(\boldsymbol{k},t)\,, \tag{1.33}$$

where $k_i = k^i$ is a 3-vector in Euclidean space. We will drop the tilde later for convenience, but the arguments and the presence of k avoids any confusion between configuration-space and Fourier-space quantities.

[6]As above, Fourier conventions usually lead to confusion and missing factors. Usually conventions differ in how the $(2\pi)^3$ factor is distributed between the expressions above.

As an example of how working in Fourier space simplifies computations, let us consider a linear partial differential equation

$$\frac{\partial^2}{\partial t^2}\delta + A(t)\frac{\partial}{\partial t}\delta + B(t)\nabla^2 C = 0\,, \tag{1.34}$$

which in Fourier space becomes

$$\frac{\partial^2}{\partial t^2}\delta + A(t)\frac{\partial}{\partial t}\delta - B(t)k^2 C = 0\,, \tag{1.35}$$

a set of decoupled ordinary differential equations: we can solve the equation independently for each $\boldsymbol{k}$ mode, which means that every Fourier model evolves independently of the rest (instead of solving an infinite number of coupled equations in configuration space). At linear order in cosmology, each mode evolves independently of the rest, hence cosmological perturbation theory is solved in Fourier space; non linearities couple different Fourier modes, which significantly complicates the computations.

1.3.2 Perturbed stress-energy tensor

The Einstein equations relate the Einstein tensor and the stress-energy tensor (i.e., gravity with matter). Therefore, before deriving the equations describing the evolution of the metric perturbations, we need to find the form of the linear perturbations for the stress-energy tensor.

For a perfect fluid, the stress-energy tensor is $T^{\mu\nu} = (\rho + P)u^\mu u^\nu + Pg^{\mu\nu}$, where ρ and P are the total proper energy density and pressure in the rest frame and $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\lambda$ is the 4-velocity.[7] In locally flat coordinates in the fluid rest frame, $T^{00} = \rho$ is the energy density, $T^{i0} = 0$ is the momentum density, and $T^{ij} = P\delta^{ij}$ is the spatial stress tensor. However, an imperfect fluid may have additional components describing shear, bulk viscosity or thermal conduction. The most general stress tensor is defined as

$$T^\mu_\nu = (\rho + P)u^\mu u_\nu + Pg^\mu_\nu + \Sigma^\mu_\nu\,, \tag{1.36}$$

where $\Sigma^{\mu\nu}$ can be taken traceless and flow orthogonal ($\Sigma^\mu_\nu u^\nu = 0$). In locally flat coordinates in the fluid rest frame, only the spatial coordinates are nonzero. Under these definitions, ρu^μ is the energy-current 4-vector, including heat conduction, while P includes the bulk viscosity and $\Sigma^{\mu\nu}$ (called the shear stress), includes the shear viscosity.

This situation is very similar for a perturbed system, where we have the perturbed metric from Eq. (1.31). Let us express the perturbed stress energy tensor as $T^\mu_\nu = \bar{T}^\mu_\nu + \delta T^\mu_\nu$. The total stress energy tensor is

$$\begin{aligned} T^0_0 &= -\rho\,, & T^i_0 &= -(\rho + P)v^i\,, \\ T^0_j &= (\rho + P)v_j\,, & T^i_j &= P\delta^i_j + \Sigma^i_j\,. \end{aligned} \tag{1.37}$$

[7]Do not confuse the notation for the pressure and the 4-momentum. The latter will always have an index.

The perturbation term is to linear order

$$\delta T^{\mu}_{\nu} = (\delta\rho + \delta P)\bar{u}^{\mu}\bar{u}_{\nu} + (\bar{\rho} + \bar{P})(\delta u^{\mu}\bar{u}_{\nu} + \bar{u}^{\mu}\delta u_{\nu}) + \delta P \delta^{\mu}_{\nu} + \Sigma^{\mu}_{\nu}\,, \tag{1.38}$$

where we are only taking into account that the shear stress is a perturbation, and we also consider the 4-velocity as its mean value plus a perturbation δu^{μ}. Starting from the normalization of the 4-velocity $g_{\mu\nu}u^{\mu}u^{\nu} = -1$, at linear order its perturbation is

$$\delta g_{\mu\nu}\bar{u}^{\mu}\bar{u}^{\nu} + 2\bar{u}_{\mu}\delta u^{\mu} = 0\,. \tag{1.39}$$

Since $\bar{u}^{\mu} = a^{-1}\delta^{\mu}_{0}$, $\bar{u}_{\mu} = -a\delta^{0}_{\mu}$, and $\delta g_{00} = -2a^2\Psi$, we find that $\delta u^0 = -\Psi/a$. On the other hand, δu^i is proportional to the *coordinate velocity* $v^i \equiv \mathrm{d}x^i/\mathrm{d}\tau$, finding $\delta u^i = v^i/a$. Then, at linear order,

$$u^{\mu} = a^{-1}\left[1 - \Psi, v^i\right]\,, \qquad\qquad u_{\mu} = a\left[-(1+\Psi), v_i\right]\,. \tag{1.40}$$

Substituting this expression and Eq. (1.31) in Eq. (1.38) we find at linear order

$$\begin{aligned} \delta T^{0}_{0} &= -\delta\rho\,, & \delta T^{i}_{0} &= -(\bar{\rho} + \bar{P})v^i\,, \\ \delta T^{0}_{j} &= (\bar{\rho} + \bar{P})v_i\,, & \delta T^{i}_{j} &= \delta P \delta^{i}_{j} + \Sigma^{i}_{j}\,. \end{aligned} \tag{1.41}$$

If there are several matter components, each of the quantities appearing above is the sum of all of the component contributions, except for the velocities, for which the momentum densities $(\bar{\rho} + \bar{P})v^i$ is the quantity that is additive. Finally, a similar scalar-vector-tensor decomposition can be applied to the stress-energy tensor.

1.3.3 Evolution of metric perturbations

We are now ready to derive the Einstein equations at linear order. It is a straightforward exercise, but with very cumbersome tensor manipulations. As a reference, the metric is

$$\begin{aligned} g_{00} &= -a^2(1 + 2\Psi)\,, & g_{i0} &= 0\,, & g_{ij} &= a^2\delta_{ij}(1 + 2\Phi)\,, \\ g^{00} &= -a^{-2}(1 - 2\Psi)\,, & g^{i0} &= 0\,, & g^{ij} &= a^{-2}\delta^{ij}(1 - 2\Phi)\,, \end{aligned} \tag{1.42}$$

and the Einstein equations are

$$R^{\mu}_{\nu} - \frac{1}{2}R g^{\mu}_{\nu} = 8\pi G T^{\mu}_{\nu}\,. \tag{1.43}$$

To evaluate the left-hand side we need to compute the Christoffel symbols for the perturbed metric, use them to obtain the Ricci tensor and contract this one to form the Ricci scalar. We will work in Fourier space (changing spatial derivatives to $i\boldsymbol{k}$ factors) directly to ease the computations. We need two independent equations (one for Ψ and another one for Φ), which are easily identifiable with the 00 and scalar ij components of the Einstein equations.

1.3.3.1 Computing the pieces for the perturbed Einstein tensor

The Christoffel symbols are given by

$$\Gamma^{\mu}_{\nu\rho} = \frac{1}{2} g^{\mu\lambda} (\partial_\nu g_{\lambda\rho} + \partial_\rho g_{\lambda\nu} - \partial_\lambda g_{\nu\rho}). \tag{1.44}$$

The components $\Gamma^0_{\mu\nu} = -(1-2\Psi)/2a^2 \left[\partial_\mu g_{0\nu} + \partial_\nu g_{0\mu} - \partial_0 g_{\mu\nu}\right]$. For the 00 component the three elements in the brackets are identical, which leaves $\Gamma^0_{00} = \mathcal{H} + \Psi'$. Using a similar approach, the Christoffel symbols at linear order are

$$\begin{aligned}
\Gamma^0_{00} &= \mathcal{H} + \Psi' \,, \\
\Gamma^0_{0i} &= i k_i \Psi \,, \\
\Gamma^0_{ij} &= \delta_{ij} \left(\mathcal{H} + 2\mathcal{H}\left[\Phi - \Psi\right] + \Phi'\right) \,, \\
\Gamma^i_{00} &= i \delta^i_j k_j \Psi \,, \\
\Gamma^i_{j0} &= \delta^i_j \left(\mathcal{H} + \Phi'\right) \,, \\
\Gamma^i_{jk} &= \left[\delta^i_j k_k + \delta^i_k k_j - \delta_{jk} \delta^i_l k_l\right] i\Phi \,.
\end{aligned} \tag{1.45}$$

The Ricci tensor is given by

$$R_{\mu\nu} = \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\lambda} + \Gamma^\lambda_{\lambda\rho} \Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\rho} \,. \tag{1.46}$$

As we did before, let us explore the time-time component: $R_{00} = \partial_\alpha \Gamma^\alpha_{00} - \partial_0 \Gamma^\alpha_{0\alpha} + \Gamma^\alpha_{\beta\alpha} \Gamma^\beta_{00} - \Gamma^\alpha_{\beta 0} \Gamma^\beta_{\alpha 0}$. When $\alpha = 0$ all terms cancel directly. For the rest we have (remember that $\partial_\tau \mathcal{H} = a''/a - \mathcal{H}^2$)

$$\begin{aligned}
R_{00} &= -k^2\Psi - 3\left(\frac{a''}{a} - \mathcal{H}^2 + \Phi''\right) + 3\mathcal{H}\left(\mathcal{H} + \Psi' + \Phi'\right) - 3\mathcal{H}\left(\mathcal{H} + 2\Phi'\right) = \\
&= -k^2\Psi - 3\left(\frac{a''}{a} - \mathcal{H}^2 + \Phi''\right) + 3\mathcal{H}\left(\Psi' - \Phi'\right) \,.
\end{aligned} \tag{1.47}$$

We will skip the $0i$ component for reasons that will be apparent later, and finally the space-space part is

$$\begin{aligned}
R_{ij} = \delta_{ij} \Bigg[&\left(\frac{a''}{a} + \mathcal{H}^2\right)\left(1 + 2\Phi - 2\Psi\right) + \\
&+ \mathcal{H}\left(5\Phi' - \Psi'\right) + \Phi'' + k^2\Phi \Bigg] + k_i k_j \left(\Phi + \Psi\right) \,.
\end{aligned} \tag{1.48}$$

Now we can contract the Ricci tensor to obtain the Ricci scalar, $R \equiv g^{\mu\nu} R_{\mu\nu} = g^{00} R_{00} + g^{ij} R_{ij}$, as

$$\begin{aligned}
a^2 R =& -(1-2\Psi)\left[-k^2\Psi - 3\left(\frac{a''}{a} - \mathcal{H}^2 + \Phi''\right) + 3\mathcal{H}\left(\Psi' - \Phi'\right)\right] + \\
&+ (1-2\Phi)\left[3\left\{\left(\frac{a''}{a} + \mathcal{H}^2\right)(1 + 2\Phi - 2\Psi) + \right.\right. \\
&\left.\left. + \mathcal{H}\left(5\Phi' - \Psi'\right) + \Phi'' + k^2\Phi\right\} + k^2\left(\Phi + \Psi\right)\right] = \\
=& 6\frac{a''}{a} + 2k^2(\Psi + 2\Phi) + 6\Phi'' - 12\frac{a''}{a}\Psi + 6\mathcal{H}(3\Phi' - \Psi'),
\end{aligned} \tag{1.49}$$

where we have separated the background and linear-order terms.

Now we have all the pieces to compute the perturbed Einstein tensor. Remember that

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R. \tag{1.50}$$

For the time-time component we have then

$$G_{00} = 3\mathcal{H}^2 + 6\mathcal{H}\Phi' + 2k^2\Phi, \tag{1.51}$$

where many terms cancel between the Ricci tensor and the Ricci scalar, and others are neglected due to being second order. We can skip the time-space component, since we only need two equations. For the space-space component,

$$\begin{aligned}
G_{ij} = \delta_{ij}\left[-2\frac{a''}{a} + \mathcal{H}^2\right] + 2\delta_{ij}&\left[2\frac{a''}{a}\left(\Psi - \Phi\right) + \mathcal{H}^2\left(\Phi - \Psi\right) + \right. \\
&\left. + \mathcal{H}\left(\Psi' - 2\Phi'\right) - 2\Phi'' - k^2\left(\Phi + \Psi\right)\right] + k_i k_j\left(\Phi + \Psi\right).
\end{aligned} \tag{1.52}$$

1.3.3.2 Perturbed Einstein Equations

Now we just need to manipulate the elements listed above to solve the Einstein equations. Note that substituting the elements derived in the previous subsection, we find the background solution for the Einstein Equations (i.e., the Friedman equations), hence we can cancel all background terms below, but we cannot do this directly, since some multiplicative terms may survive.

Let us consider first the trace-free part of the space-space component of the equations. The longitudinal trace-free part of the space components of a tensor A_{ij} can be obtained contracting it with $(\hat{k}^i\hat{k}^j - \delta^{ij}/3)$, therefore: $(\hat{k}^i\hat{k}^j - \frac{1}{3}\delta^{ij})\delta G_{ij} = 8\pi G(\hat{k}^i\hat{k}^j - \frac{1}{3}\delta^{ij})\delta T_{ij})$, such as

$$k^2\left(\Phi + \Psi\right) = -12\pi G a^2(\bar{\rho} + \bar{P})\sigma, \tag{1.53}$$

where we have defined $(\bar{\rho} + \bar{P})\sigma = -\left(\hat{k}_i\hat{k}_j - \delta_{ij}/3\right)\Sigma^i_j = \sum_s(\bar{\rho}_s + \bar{P}_s)\sigma_s$ as the shear (note that Σ^i_j is by definiton the traceless component of T^i_j). This is a very important result in cosmology: at linear order, and in the absence of shear, $\Psi = -\Phi$. In the standard cosmological model, there is a very small shear that is generated by photons and neutrinos, as we will see in the next section.

Now let us consider the time-time component. In this case $T^0_0 = -\rho$, and in the Newtonian gauge $T_{00} = g_{00}T^0_0 = a^2(1+2\Psi)\rho$, so that we have

$$3\mathcal{H}^2 + 6\mathcal{H}\Phi' + 2k^2\Phi = 8\pi G a^2(1+2\Psi)(\bar{\rho}+\delta\rho) = 8\pi G a^2\bar{\rho}(1+2\Psi+\delta)\,, \quad (1.54)$$

where we have defined $\delta = \delta\rho/\bar{\rho}$ and the last equality retains only linear terms. Then, canceling the background solution $3\mathcal{H}^2 = 8\pi G a^2\bar{\rho}$, we find

$$k^2\Phi + 3\mathcal{H}\left(\Phi' - \mathcal{H}\Psi\right) = 4\pi G a^2\bar{\rho}\delta\,. \quad (1.55)$$

Therefore, the two Einstein equations of interest for scalar perturbations in the Newtonian gauge are given by

$$\begin{aligned} k^2\left(\Phi+\Psi\right) &= -12\pi G a^2(\bar{\rho}+\bar{P})\sigma\,, \\ k^2\Phi + 3\mathcal{H}\left(\Phi' - \mathcal{H}\Psi\right) &= 4\pi G a^2\bar{\rho}\delta\,. \end{aligned} \quad (1.56)$$

The second equation is the generalization of the Poisson equation to an expanding and perturbed Universe. We recover the behavior of Newtonian gravity in the Newtonian regime, which is achieved for very small scales, in which expansion can be neglected, and $k \gg \mathcal{H}$.

There are two other possible equations, obtained from the $0i$ component and the trace of the spatial sector of the Einstein equation. These are

$$\begin{aligned} &-k^2\left(\Phi' - \mathcal{H}\Psi\right) = 4\pi G a^2(\bar{\rho}+\bar{P})\theta\,, \\ &-\Phi'' + \mathcal{H}(\Psi' - 2\Phi') + \left(2\frac{a''}{a} - \mathcal{H}^2\right)\Psi - \frac{k^3}{3}(\Phi+\Psi) = \frac{4\pi}{3}G a^2\delta T^i_i\,, \end{aligned} \quad (1.57)$$

where $(\bar{\rho}+\bar{P})\theta = \sum_s(\bar{\rho}_s+\bar{P}_s)\theta_s$ and $\theta = ik_i v^i$ is the divergence of the coordinate (or fluid) velocity.

1.3.4 Perturbed Boltzmann equations

At the beginning of this chapter we discussed the Boltzmann equations in a smooth expanding Universe. Now it is time to introduce metric perturbations in the formalism. Metric perturbations affect how particles move, which in turn affect the phase-space distribution. We need to know how the position, momentum and direction of the momentum change with time.

The mass-shell constrain $g_{\mu\nu}P^\mu P^\nu = -m^2$, accounting for metric perturbations, is now given by

$$a^2(1+2\Psi)(P^0)^2 + p^2 = -m^2\,, \quad (1.58)$$

where as always $p^2 = g_{ij}P^iP^j$. Defining still the energy as in the unperturbed case, $E = \sqrt{p^2+m^2}$, the time component of P^μ is determined by the energy and the metric perturbation. At linear order,

$$P^\mu = [E(1-\Psi)/a, p_i(1-\Phi)/a] \ . \tag{1.59}$$

The conjugate momenta is $P_i = ap_i(1+\Phi)$.

Note that the phase-space distribution is a scalar and is invariant under canonical transformations. Its zeroth-order is either a Fermi-Dirac (+) or a Bose-Einstein (−) distribution:

$$f_0 = f_0(\epsilon) = \frac{g_*}{h_P^3}\frac{1}{\exp\{\epsilon/k_\mathrm{B}T_0\} \pm 1} \ , \tag{1.60}$$

where we have defined $\epsilon = aE = a\sqrt{p^2+m^2}$ and $T_0 = aT$ as the temperature of the particles today for convenience, and h_P and k_B are the Planck and the Boltzmann constants, respectively. ϵ is related to the time component of the 4-momentum by $P_0 = -\epsilon(1+\Psi)$.

Also, for convenience, let us replace the conjugate momentum P_i by the comoving momentum $q_i \equiv ap_i$ in order to eliminate the metric perturbations from the definition of the momenta, and as we have done before we separate $q_i = q\hat{q}_i$ on its magnitude and direction. Then, $f(x^i, P_j, \tau) \to f(x^i, q, \hat{q}_j, \tau)$. q_j is not the conjugate momentum and we cannot consider the phase-space volume element to be $\mathrm{d}^3\boldsymbol{x}\mathrm{d}\boldsymbol{q}/(2\pi)^3$. In practice we have moved the impact of the perturbations from the variable to the phase-space volume element.

Remember that the general expression for the stress-energy tensor can be written as

$$T_{\mu\nu} = \frac{1}{\sqrt{-|g_{\alpha\beta}|}}\int \mathrm{d}^3\boldsymbol{P}\frac{P_\mu P_\nu}{P^0} f(x^i, P_j, \tau) \, , \tag{1.61}$$

and, as done with all quantities so far, we can treat the distribution function perturbatively,

$$f(x^i, P_j, \tau) = f_0(q, m)\left(1 + \varphi(x^i, q, \hat{q}_j, \tau)\right) \ , \tag{1.62}$$

such as the only thing left is to transform the geometric factors from Eq. (1.61). First, $(-|g_{\alpha\beta}|)^{-1/2} = (1-\Psi-3\Phi)/a^4$ and $\mathrm{d}^3\boldsymbol{P} = (1+3\Phi)q^2\mathrm{d}q\mathrm{d}\Omega_\mathrm{q}$ at linear order, where Ω_q is the solid angle for $\hat{q}_j$. Now we can express the components of the stress-energy tensor in terms of the perturbed phase-space distribution (substituting the 4-momenta in Eq. (1.61)):

$$\begin{aligned}
T^0_0 &= -a^{-4}\int \mathrm{d}q\mathrm{d}\Omega_\mathrm{q} q^2\sqrt{q^2+m^2a^2} f_0(q,m)(1+\varphi) \, , \\
T^0_i &= a^{-4}\int \mathrm{d}q\mathrm{d}\Omega_\mathrm{q} q^2 q\hat{q}_i f_0(q,m)\varphi \, , \\
T^i_j &= a^{-4}\int \mathrm{d}q\mathrm{d}\Omega_\mathrm{q} q^2\frac{q^2\hat{q}^i\hat{q}_j}{\sqrt{q^2+m^2a^2}} f_0(q,m)(1+\varphi) \, ,
\end{aligned} \tag{1.63}$$

where we have used that $\int \mathrm{d}\Omega_{\mathrm{q}}\hat{q}_i = \int \mathrm{d}\Omega_{\mathrm{q}}\hat{q}_i\hat{q}_j\hat{q}_k = 0$ (which cancels the unperturbed f_0 term in T_i^0) and $\int \mathrm{d}\Omega_{\mathrm{q}}\hat{q}_i\hat{q}_j = 4\pi\delta_{ij}/3$ (which makes that the term in T_j^i survives).

The general Boltzmann equation is, in terms of the variables discussed now,

$$\frac{\mathrm{d}f}{\mathrm{d}\tau} = \frac{\partial f}{\partial \tau} + \frac{\partial f}{\partial x^i}\frac{\mathrm{d}x^i}{\mathrm{d}\tau} + \frac{\partial f}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}\tau} + \frac{\partial f}{\partial \hat{q}_i}\frac{\mathrm{d}\hat{q}_i}{\mathrm{d}\tau} = C[f]\,. \tag{1.64}$$

Then, we need to obtain the total derivatives as function of τ to obtain the expression in each gauge. The total derivatives (how position and momentum change with time in the absence of collisions) is where gravity (through the determination of the geodesics) chimes in. Remember that $P^i \equiv \mathrm{d}x^i/\mathrm{d}\lambda$ and $P^0 \equiv \mathrm{d}\tau/\mathrm{d}\lambda$, such as at linear order

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}\tau} = \frac{P^i}{P^0}\,, \tag{1.65}$$

which corresponds to

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = q\hat{q}^i(1-\Phi+\Psi)/\epsilon\,. \tag{1.66}$$

Remember the geodesic equation

$$\frac{\mathrm{d}P^\mu}{\mathrm{d}\lambda} = -\Gamma^\mu_{\alpha\beta}P^\alpha P^\beta\,. \tag{1.67}$$

Using that $\mathrm{d}/\mathrm{d}\lambda = (\mathrm{d}x^\mu/\mathrm{d}\lambda)(\mathrm{d}/\mathrm{d}x^\mu) = P^\mu \mathrm{d}/\mathrm{d}x^\mu$, we have for the space component,

$$P^0\frac{\mathrm{d}P^i}{\mathrm{d}\tau} + P^j\frac{\mathrm{d}P^i}{\mathrm{d}x^j} = -\Gamma^i_{\alpha\beta}P^\alpha P^\beta\,. \tag{1.68}$$

Now we can straightforwardly obtain the derivative with respect q^i doing some algebra: we start from P^i and go to p^i, and from this to p and q. Then, we need to use Eq. (1.59) in the equation above and propagate. To start, note that

$$\begin{aligned}\frac{\mathrm{d}P^i}{\mathrm{d}\tau} &= \frac{1-\Phi}{a}\frac{\mathrm{d}p^i}{\mathrm{d}\tau} - \frac{p^i}{a}\left(\mathcal{H}[1-\Phi]+\Phi'\right)\,,\\ \frac{\mathrm{d}P^i}{\mathrm{d}x^j} &= -\frac{ip^ik_j\Phi}{a}\,.\end{aligned} \tag{1.69}$$

Substituting this in Eq. (1.68) and isolating $\mathrm{d}p^i/\mathrm{d}\tau$, we find that the geodesic equation is transformed to

$$\begin{aligned}\frac{\mathrm{d}p^i}{\mathrm{d}\tau} =& \frac{a^2(1+\Phi+\Psi)}{E}\times\\ &\times\left\{-\Gamma^i_{\alpha\beta}P^\alpha P^\beta + \frac{ip^ip^jk_j\Phi}{a^2} + \frac{Ep^i}{a^2}\left(\mathcal{H}[1-\Phi-\Psi]+\Phi'\right)\right\}\,.\end{aligned} \tag{1.70}$$

We have to compute now the term with the Christoffel symbols, using Eq. (1.45). Expanding the expression we have

$$-\Gamma^i_{\alpha\beta} = -\left(\Gamma^i_{00}P^0P^0 + 2\Gamma^i_{0j}P^0P^j + \Gamma^i_{jk}P^jP^k\right) . \tag{1.71}$$

Note that the Christoffel symbols in the first and last terms are already first order, so we can ignore the contributions from the perturbations in the momenta. Neglecting higher-order terms, we have

$$\begin{aligned} -\Gamma^i_{\alpha\beta} = &-i\frac{E^2}{a^2}k_i\Psi- \\ &-\frac{2Ep^i}{a^2}\left(\mathcal{H}[1-\Psi-\Phi]+\Phi'\right)- \\ &-\frac{i\Phi}{a^2}\left(p^ip_kk_k+p^ip_kk_k-p^2k_i\right) . \end{aligned} \tag{1.72}$$

Therefore, adding all contributions, we have

$$\frac{\mathrm{d}p^i}{\mathrm{d}\tau} = -p^i(\mathcal{H}+\Phi') - iEk_i\Psi - i\frac{\Phi}{E}(p^ip^jk_j - p^2k_i)\,, \tag{1.73}$$

and using that

$$\frac{\mathrm{d}p}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\sqrt{\delta_{ij}p^ip^j} = \delta_{ij}\frac{p^i}{p}\frac{\mathrm{d}p^j}{\mathrm{d}\tau}\,, \tag{1.74}$$

we have

$$\frac{\mathrm{d}p}{\mathrm{d}\tau} = -p(\mathcal{H}+\Phi') - iE\hat{p}^ik_i\Psi\,. \tag{1.75}$$

However the cumbersome calculation, it would have been possible to qualitatively guess the result, since the first term corresponds to the loss of momentum due to the Hubble expansion (including cosmological redshift and decay of the peculiar velocity) and the second term encodes the effect of the particle traveling into a potential well. The last two terms in Eq. (1.73) cancel when taking the norm, and this is because they do not change the particle's momentum at linear order, but they do change its direction.

The expression above in the variables we want to use, $q = ap$ and $\epsilon = aE$, converts to

$$\frac{\mathrm{d}q}{\mathrm{d}\tau} = -q\Phi' - i\epsilon\hat{q}^ik_i\Psi\,. \tag{1.76}$$

Finally, since $\partial f/\partial\hat{q}$ is also a first-order quantity, the last term in the left-hand side of Eq. (1.64) can be neglected to first order. Joining all the terms and keeping only first-order quantities, the perturbed boltzmann equation is given by

$$\frac{\partial\varphi}{\partial\tau} + i\frac{q}{\epsilon}\boldsymbol{k}\hat{\boldsymbol{q}}\varphi + \frac{\partial\log f_0}{\partial\log q}\left(-\Phi' - i\frac{\epsilon}{q}\boldsymbol{k}\hat{\boldsymbol{q}}\Psi\right) = \frac{C[f]}{f_0}\,. \tag{1.77}$$

The Boltzmann equation only depends on the direction of the momentum through its angle with $\boldsymbol{k}$, so unless there is further dependence in the collision term, φ only depends on $\hat{\boldsymbol{q}}$ through the product $\hat{\boldsymbol{k}}\hat{\boldsymbol{q}}$.

In some cases, it is also useful to keep the perturbed Boltzmann equation for the whole distribution (assuming that the zero-th order is homogeneous and does not depend on the direction of the momentum), and using the proper momentum p and energy E. Then, we have

$$\frac{\partial f}{\partial \tau} + i\frac{p}{E}\boldsymbol{k}\hat{\boldsymbol{p}}f + \frac{\partial f}{\partial p}p\left(-\mathcal{H} - \Phi' - i\frac{E}{p}\boldsymbol{k}\hat{\boldsymbol{p}}\Psi\right) = C[f]\,. \tag{1.78}$$

1.4 Evolution of matter and radiation perturbations

Now we have all the tools to compute the perturbations of all components in the Universe that contribute to the stress-energy tensor.[8] We will consider cold dark matter, baryons, massless and massive neutrinos and photons.

1.4.1 Dark matter

Dark matter makes up for most of the non-relativistic matter in the Universe, and it is mostly cold. We will consider a completely collisionless cold dark matter, i.e., the dark matter does not interact with any other species in the Universe or itself other than gravitationally and it is completely non relativistic. This means $C[f] = 0$, and that factors $q/\epsilon = p/E \sim p/m$ will be very small: we will only retain up to linear-order terms in p/m, which accounts for the bulk motion of dark matter but not its velocity dispersion. These assumptions make that dark matter can be treated as a pressure-less effective fluid which is described by its density and velocity. We will derive the evolution equations taking moments of the Boltzmann equations.

From the phase-space distribution we can take the description of a collection of particles if we integrate over phase-space volume elements. Remember that the number density and the fluid velocity can be obtained by integrating over the proper momentum; denoting dark matter with a subscript 'c',

$$n_c = \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} f_c\,, \qquad n_c v_c^i = \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{p\hat{p}^i}{E} f_c\,. \tag{1.79}$$

[8] For effective fuilds, the fluid equations can be obtained by taking moments in q of the Boltzmann equations, similar to Eq. (1.61).

Then, if we multiply the Boltzmann equation in Eq. (1.78) for the whole distribution by the phase-space element and integrate we have

$$\begin{aligned}\frac{\partial}{\partial\tau}\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}f_c+i\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{p\boldsymbol{k}\hat{\boldsymbol{p}}}{E}f_c-(\mathcal{H}+\Phi')\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{\partial f_c}{\partial p}p-\\ -\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{\partial f_c}{\partial p}iE\boldsymbol{k}\hat{\boldsymbol{p}}\Psi=0\,,\end{aligned} \tag{1.80}$$

where we can substitute the first two terms by the number density and fluid velocity and integrating by parts the integral in the third term is

$$\frac{1}{(2\pi)^3}\int\mathrm{d}pp^3\frac{\partial}{\partial p}\int\mathrm{d}\Omega_{\mathrm{p}}f_c=\frac{-3}{(2\pi)^3}\int\mathrm{d}pp^2\frac{\partial}{\partial p}f_c=-3n_c\,, \tag{1.81}$$

and the fourth vanishes. Then, the zero-th moment of the Boltzmann equation returns

$$\frac{\partial n_c}{\partial\tau}+in_c\boldsymbol{k}\boldsymbol{v}_c+3(\mathcal{H}+\Phi')n_c=0\,, \tag{1.82}$$

which is the cosmological generalization of the continuity equation, including the last term to account for the perturbations of the metric and the expansion of the Universe. The zero-th order in perturbations above returns (remember that the velocity is already a first-order perturbation)

$$\frac{\partial\bar{n}_c}{\partial\tau}+3\mathcal{H}\bar{n}_c=0\,, \tag{1.83}$$

which shows that $n_c\propto a^{-3}$ for non-relativistic matter, as discussed before in the course. Perturbing this number density as $n_c=\bar{n}_c(1+\delta_c)$ (which also fulfills previous definitions of δ), and dividing by $a^3\bar{n}_c$ we find

$$\delta_c'=-\theta_c-3\Phi'\,, \tag{1.84}$$

where we have recovered the definition of θ as the velocity divergence. We still need another equation to determine the evolution of δ_c and θ_c, which we can get by using the first moment of the Boltzmann equation (weighting the integral with $p\hat{p}^j/E$):

$$\begin{aligned}\frac{\partial}{\partial\tau}\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{p\hat{p}^j}{E}f_c+i\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{p^2\hat{p}^j\boldsymbol{k}\hat{\boldsymbol{p}}}{E^2}f_c-(\mathcal{H}+\Phi')\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{\partial f_c}{\partial p}\frac{p^2\hat{p}^j}{E}-\\ -\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{\partial f_c}{\partial p}p\hat{p}^ji\boldsymbol{k}\hat{\boldsymbol{p}}\Psi=0\,.\end{aligned} \tag{1.85}$$

The first term is the time derivative of $a^3n_cv_c^j$ and the second can be neglected, since it is second order in p/E. Integrating by parts the third term we get

$$\begin{aligned}\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{\partial f_c}{\partial p}\frac{p^2\hat{p}^j}{E}&=\int\frac{\mathrm{d}\Omega_{\mathrm{p}}}{(2\pi)^3}\hat{p}^j\int\mathrm{d}p\frac{p^4}{E}\frac{\partial f_c}{\partial p}=\\ &=-\int\frac{\mathrm{d}\Omega_{\mathrm{p}}}{(2\pi)^3}\hat{p}^j\int\mathrm{d}pf_c\left(\frac{4p^3}{E}-\frac{p^5}{E^3}\right)\,.\end{aligned} \tag{1.86}$$

the first term in the brackets yields $-4a^3 n_c v_c^j$, while the second term is higher order in p/E and thus can be neglected. Applying the same approach to the last term in the weighted integral of the Boltzmann equation and considering that $\int \mathrm{d}\Omega_\mathrm{p} \hat{p}^i \hat{p}^j = \delta^{ij} 4\pi/3$, we get that the first moment of the Boltzmann equation is

$$\frac{\partial(n_c v_c^j)}{\partial\tau} + 4\mathcal{H} n_c v_c^j + i n_c k^j \Psi = 0\,. \tag{1.87}$$

Since all terms are already first order, we can directly write n_c as $\bar{n}_c$, and use its background evolution. After multiplying by ik_j, we find

$$\theta_c' = -\mathcal{H}\theta_c + k^2\Psi\,. \tag{1.88}$$

This is the momentum conservation, or Euler equation, although in this case it does not contain the standard $(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{v}$ term because it is second order. Even if the dark matter perturbations are the simplest ones, we can already see a common feature of integrating the Boltzmann equations: the integrated n-th moment always depends on the $(n+1)$-th moment. This leads to an infinite hierarchy of equations that needs to be closed at some moment. In the case of cold dark matter the hierarchy is closed setting the second moment to zero (which follows from the assumption that the dark matter is cold) and the drop of $(p/E)^2$ and higher terms. This will not be possible for relativistic species as neutrinos and photons, as we will see below.

To summarize, the Boltzmann equations for dark matter are

$$\delta_c' = -\theta_c - 3\Phi'\,, \qquad \theta_c' = -\mathcal{H}\theta_c + k^2\Psi\,. \tag{1.89}$$

1.4.2 Massless neutrinos

The perturbations of the stress-energy tensor corresponding to neutrinos are

$$\begin{aligned}
\delta\rho_\nu &= 3\delta P_\nu = a^{-4}\int q^2 \mathrm{d}q \mathrm{d}\Omega_\mathrm{q} q f_0(q)\varphi\,,\\
\delta T^0_{\nu i} &= a^{-4}\int q^2 \mathrm{d}q \mathrm{d}\Omega_\mathrm{q} q \hat{q}^i f_0(q)\varphi\,,\\
\Sigma^i_{\nu j} &= a^{-4}\int q^2 \mathrm{d}q \mathrm{d}\Omega_\mathrm{q} q (\hat{q}^i\hat{q}_j - \frac{1}{3}\delta^i_j) f_0(q)\varphi\,,
\end{aligned} \tag{1.90}$$

where, for massless particles, $q = \epsilon$. Since the quantities involved in the stress-energy tensor are weighted integrals of the phase space, we can further reduce the number of variables if we integrate out the q-dependence of φ and expand the angular dependence in a series of Legendre polynomials $\mathcal{P}_\ell(\mu)$, where $\mu = \hat{\boldsymbol{k}}\hat{\boldsymbol{q}}$. Let us define

$$\mathcal{F}_\nu(\boldsymbol{k},\hat{q},\tau) \equiv \frac{\int q^2 \mathrm{d}q q f_0 \varphi}{\int q^2 \mathrm{d}q q f_0} \equiv \sum (-i)^\ell (2\ell+1)\mathcal{F}_{\nu\ell}(\boldsymbol{k},\tau)\mathcal{P}_\ell(\mu)\,, \tag{1.91}$$

where the factor $(-i)^\ell(2\ell+1)$ has been chosen to simplify the expansion of a plane wave: $\mathcal{F}_\nu = \exp(-ikr\mu)$ has expansion coefficients $\mathcal{F}_{\nu\ell} = j_\ell(kr)$ given by the spherical Bessel functions. The purpose of the expansion in Legendre polynomials is to remove the explicit dependence in μ, which complicates the computations.

The fluid variables of interest ($\delta \equiv \delta\rho/\bar\rho$, $\theta \equiv ik^j\delta T^0_j/(\bar\rho+\bar P)$, and $\sigma \equiv -(\hat k_i \hat k^j - \delta^j_i/3)\Sigma^i_j/(\bar\rho+\bar P)$) can be expressed in terms of the expansion coefficients of the new variable $\mathcal{F}_\nu$ by performing the corresponding weighted angular integral to $\mathcal{F}_\nu$. From Eq. (1.90):

$$
\begin{aligned}
\delta_\nu &= \frac{1}{4\pi}\int \mathrm{d}\Omega_\mathrm{q}\mathcal{F}_\nu = \mathcal{F}_{\nu 0}\,, \\
\theta_\nu &= \frac{3i}{16\pi}\int \mathrm{d}\Omega_\mathrm{q}(\hat{\boldsymbol{k}}\hat{\boldsymbol{q}})k\mathcal{F}_\nu = \frac{3}{4}k\mathcal{F}_{\nu 1}\,, \\
\sigma_\nu &= -\frac{3}{16\pi}\int \mathrm{d}\Omega_\mathrm{q}\left[(\hat{\boldsymbol{k}}\hat{\boldsymbol{q}})^2 - \frac{1}{3}\right]\mathcal{F}_\nu = \frac{1}{2}\mathcal{F}_{\nu 2}\,,
\end{aligned}
\tag{1.92}
$$

where the division of the background quantities is already in the denominator of the definition of $\mathcal{F}_\nu$ and the numerical prefactors account for the angular integrals of the background, homogeneous distribution function and the match with the Legendre coefficients.

Applying the definition of $\mathcal{F}_\nu$ above to Eq. (1.77), the Boltzmann equation for massless neutrinos becomes

$$
\mathcal{F}_\nu' = -ik\mu\mathcal{F}_\nu - 4(\Phi' + ik\mu\Psi)\,,
\tag{1.93}
$$

where $\left[\int q^2\mathrm{d}q f_0 \mathrm{d}\log f_0/\mathrm{d}\log q\right] / \int q^2\mathrm{d}q f_0 = -4$ and $\mathcal{P}_2(\mu) = (3\mu^2-1)/2$.

Now we can mix Eqs. (1.93) and (1.91) to obtain the evolution for the coefficients. For instance, in the Newtonian gauge,

$$
\begin{aligned}
\sum(-i)^\ell(2\ell+1)\mathcal{F}'_{\nu\ell}\mathcal{P}_\ell(\mu) = &-k\sum(-i)^{\ell+1}(2\ell+1)\mathcal{F}_{\nu\ell}\mu\mathcal{P}_\ell(\mu) \\
&- 4(\Phi' + ik\mu\Psi)\,.
\end{aligned}
\tag{1.94}
$$

We can use the orthonormality of the Legendre polynomials and the recursion relation of $(2\ell+1)\mu\mathcal{P}_\ell(\mu) = \ell\mathcal{P}_{\ell-1}(\mu) + (\ell+1)\mathcal{P}_{\ell+1}(\mu)$, such as if we multiply each side of the equation above by $\mathcal{P}_{\ell'}$ and integrate over μ we can get the relations that we need:

$$
\begin{aligned}
\delta_\nu' &= -\frac{4}{3}\theta_\nu - 4\Phi'\,, \\
\theta_\nu' &= k^2\left(\frac{1}{4}\delta_\nu - \sigma_\nu\right) + k^2\Psi\,, \\
\mathcal{F}'_{\nu\ell} &= \frac{k}{2\ell+1}\left[\ell\mathcal{F}_{\nu\ell-1} - (\ell+1)\mathcal{F}_{\nu\ell+1}\right]\,, \qquad \ell \geq 2\,.
\end{aligned}
\tag{1.95}
$$

Note that for a given ℓ, $\mathcal{F}_{\nu\ell}$ is coupled to the two neighbouring modes, and that a priori the Boltzann hierarchy is infinite. Therefore, we need to truncate at some $\ell_{\rm max}$. One option is to set $\mathcal{F}_{\nu\ell} = 0$ for $\ell > \ell_{\rm max}$, but this is inaccurate because the error in the coupling at $\ell_{\rm max}$ propagates to smaller ℓ due to the coupling between modes.[9] An improved truncation scheme is based in the extrapolation of the behavior of $\mathcal{F}_{\nu\ell}$ at $\ell = \ell_{\rm max}+1$. More sophisticated schemes have been developed to improve the accuracy of the Boltzmann equations, including an exact solution transforming Eq. (1.93) into an integral equation, which allows to solve the system iteratively.

1.4.3 Massive neutrinos

Massive neutrinos are a very particular species in the Universe. Their mass, which sums to $0.06\,{\rm eV} \leq \sum m_\nu \lesssim 0.1$ eV implies that they were relativistic particles until $z \sim 100$, when they become non relativistic as the Universe expands and they get colder. They can be considered hot dark matter, and we will denote them with the subscript 'h'. The evolution of the perturbations to their distribution function is more complicated than in the case of massless neutrinos due to the finite mass.

Experimental and observational evidence cannot distinguish between the normal and the inverted hierarchy yet, and cannot determine whether any of the neutrinos is effectively massless or not. However, cosmological perturbations are practically sensitive only to the total neutrino mass, not able to distinguish between individual neutrino masses. Since the evolution of massless neutrinos is significantly simpler (and cheaper to compute), it is customary to consider a single massive neutrino and 2 massless neutrinos in the set of Boltzmann equations.

In this case, we cannot ignore the neutrino mass (i.e., $q \neq \epsilon = (q^2 + m_\nu^2 a^2)^{1/2}$). From Eqs. (1.63) and (1.90), the unperturbed energy density and pressure, and the corresponding perturbations, are

$$\begin{aligned}
\bar\rho_h &= a^{-4}\int q^2 {\rm d}q{\rm d}\Omega_{\rm q}\epsilon f_0\,, \quad \bar P_h = \frac{1}{3}a^{-4}\int q^2{\rm d}q{\rm d}\Omega_{\rm q}\frac{q^2}{\epsilon}f_0\varphi\,,\\
\delta\rho_h &= a^{-4}\int q^2{\rm d}q{\rm d}\Omega_{\rm q}\epsilon f_0\varphi\,, \quad \delta P_h = \frac{1}{3}a^{-4}\int q^2{\rm d}q{\rm d}\Omega_{\rm q}\frac{q^2}{\epsilon}f_0\varphi\,,\\
\delta T^0_{hi} &= a^{-4}\int q^2{\rm d}q{\rm d}\Omega_{\rm q} q\hat q_i f_0\,, \quad \Sigma^i_{hj} = a^{-4}\int q^2{\rm d}q{\rm d}\Omega_{\rm q}\frac{q^2}{\epsilon}\left(\hat q^i\hat q_j - \frac{1}{3}\delta^i_j\right)f_0\varphi\,.
\end{aligned} \tag{1.96}$$

Now we can proceed with the same philosophy as for the massless neutrinos, but note that here there is a critical difference. The energy-momentum relation depends *both* in time and the momentum (since ϵ does not completely describe the case), which prevents us to integrate out the q dependence as we did before. This forces us to expand φ in the Legendre polynomial series

[9] The error propagates to $\ell = 0$ in a time $\tau \approx \ell_{\rm max}/k$ and the reflects back to increasing ℓ, due again to the coupling, increasing the errors even more.

directly:

$$\varphi(\boldsymbol{k}, \hat{\boldsymbol{q}}, q, \tau) = \sum (-i)^\ell (2\ell+1) \varphi_\ell(\boldsymbol{k}, q, \tau) \mathcal{P}_\ell(\mu) \,, \tag{1.97}$$

which, after integration over the angular variables, leaves the perturbations of interest as

$$\begin{aligned} \delta\rho_h &= 4\pi a^{-4} \int q^2 \mathrm{d}q \epsilon f_0 \varphi_0 \,, \\ \delta P_h &= \frac{4\pi}{3} a^{-4} \int q^2 \mathrm{d}q \frac{q^2}{\epsilon} f_0 \varphi_0 \,, \\ (\bar{\rho}_h + \bar{P}_h)\theta_h &= 4\pi k a^{-4} \int q^2 \mathrm{d}q q f_0 \varphi_1 \,, \\ (\bar{\rho}_h + \bar{P}_h)\sigma_h &= \frac{8\pi}{3} a^{-4} \int q^2 \mathrm{d}q \frac{q^2}{\epsilon} f_0 \varphi_2 \,. \end{aligned} \tag{1.98}$$

We can then substitute the Legendre expansion in Eq. (1.77) and match the coefficients multiplying each Legendre polynomial (and the μ dependence on the metric perturbations). Following that approach and using the same recursion relation as above:

$$\begin{aligned} \varphi_0' &= -\frac{qk}{\epsilon}\varphi_1 + \Phi' \frac{\mathrm{d}\log f_0}{\mathrm{d}\log q} \,, \\ \varphi_1' &= \frac{qk}{3\epsilon}(\varphi_0 - 2\varphi_2) - \frac{\epsilon k}{3q} \Psi \frac{\mathrm{d}\log f_0}{\mathrm{d}\log q} \,, \\ \varphi_\ell' &= \frac{1k}{(2\ell+1)\epsilon} \left[\ell\varphi_{\ell-1} - (\ell+1)\varphi_{\ell+1}\right] \,, \qquad \ell \geq 2 \,. \end{aligned} \tag{1.99}$$

Note that in this case, the set of equations to solve is much larger, since due to the q dependence, we need to solve $\ell_{\max} \times N_{\mathrm{q}}$ equations, where $\ell_{\max}$ comes from the Boltzmann hierarchy and N_{q} comes from the number of evaluations in q used to approximate the q-integration for the phase-space distribution required to obtain the quantities that contribute to the stress-energy tensor, shown in Eq. (1.98).

1.4.4 Photons

Photons (which we will denote with γ) are massless particles that interact with baryons. Therefore, in this case we need to take into account the collision term in the Boltzmann equations, which describes the effects of the Compton scattering.[10] At zero-th order the distribution function follows an unperturbed Bose-Einstein distribution. This is because the collision term includes the forward and backward reactions and we assume photons are in equilibrium,

[10]Actually, electrons are not relativistic in the times of interest, so that we could talk about Thomson scattering.

hence both reactions cancel and we can assume a Bose-Einstein distribution with no collision term. However, the perturbations from the unperturbed phase distributions are going to be determined by the collision term.

The scattering process of interest is

$$e^-(\boldsymbol{q}) + \gamma(\boldsymbol{p}) \longleftrightarrow e^-(\boldsymbol{q'}) + \gamma(\boldsymbol{p'}) \,, \tag{1.100}$$

where the proper momentum of each particle is indicated between parentheses. We are interested in $f(\boldsymbol{p})$, so we need to integrate over the other three momenta.

We will skip the derivation of the collision term, but it is important to note that there are three main contributions. The dominant contribution comes from the isotropic, polarization-averaged scattering. In addition, there is a small correction coming from the anisotropic scattering that is proportional to the quadrupole of the of the photon phase-space distribution. Finally, since the quadrupole of the distribution generates linear polarization, photon perturbations and polarization perturbations are coupled. Denoting the difference between the two linear polarization components as $\mathcal{G}_\gamma$, the final Boltzmann equations for both $\mathcal{F}_\gamma$ and $\mathcal{G}_\gamma$, which satisfy Eq. (1.77) with a right-hand side given by

$$\begin{aligned} C[\mathcal{F}_\gamma] &= \left(\frac{\partial \mathcal{F}_\gamma}{\partial \tau}\right)_C = a n_e \sigma_{\mathrm{T}} \left[-\mathcal{F}_\gamma + \mathcal{F}_{\gamma 0} + 4\hat{\boldsymbol{q}}\boldsymbol{v}_b - \frac{\mathcal{F}_{\gamma 2} + \mathcal{G}_{\gamma 0} + \mathcal{G}_{\gamma 2}}{2}\mathcal{P}_2(\mu)\right] \,, \\ C[\mathcal{G}_\gamma] &= \left(\frac{\partial \mathcal{G}_\gamma}{\partial \tau}\right)_C = a n_e \sigma_{\mathrm{T}} \left[-\mathcal{G}_\gamma + \frac{\mathcal{F}_{\gamma 2} + \mathcal{G}_{\gamma 0} + \mathcal{G}_{\gamma 2}}{2}(1 - \mathcal{P}_2(\mu))\right] \,, \end{aligned} \tag{1.101}$$

where v_b is the baryon bulk velocity, n_e is the electron number density and σ_{T} is the Thomson scattering cross section. Note that $\mathcal{G}$ depends on the quadrupole of the photon distribution.

Now we can proceed as for the case in the massless neutrinos: we expand $\mathcal{F}_\gamma$ and $\mathcal{G}_\gamma$ in Legendre series and use the relations $\hat{\boldsymbol{q}}\boldsymbol{v}_b = -(i\theta_b/k)\mathcal{P}_1(\mu)$ and those analog to Eq. (1.92), we rewrite the collision terms as

$$\begin{aligned} C[\mathcal{F}_\gamma] = a n_e \sigma_{\mathrm{T}} \Bigg[& \frac{4i}{k}(\theta_\gamma - \theta_b)\mathcal{P}_1(\mu) + \left(9\sigma_\gamma - \frac{\mathcal{G}_{\gamma 0} + \mathcal{G}_{\gamma 2}}{2}\right)\mathcal{P}_2(\mu) \\ & - \sum_{\ell \geq 3} (-i)^\ell (2\ell + 1)\mathcal{F}_{\gamma \ell}\mathcal{P}_\ell(\mu)\Bigg] \,, \end{aligned} \tag{1.102}$$

and

$$\begin{aligned} C[\mathcal{G}_\gamma] = a n_e \sigma_{\mathrm{T}} \Bigg[& \frac{1}{2}\left(\mathcal{F}_{\gamma 2} + \mathcal{G}_{\gamma 0} + \mathcal{G}_{\gamma 2}\right)(1 - \mathcal{P}_2(\mu)) - \\ & - \sum_{\ell \geq 0} (-i)^\ell (2\ell + 1)\mathcal{G}_{\gamma \ell}\mathcal{P}_\ell(\mu)\Bigg] \,. \end{aligned} \tag{1.103}$$

Then, expanding the terms in the left-hand side of the Boltzmann equations in Legendre polynomials and matching the angular dependences, we find

$$\begin{aligned}
\delta_\gamma' &= -\frac{4}{3}\theta_\gamma - 4\Phi' \,, \\
\theta_\gamma' &= k^2\left(\frac{1}{4}\delta_\gamma - \sigma_\gamma\right) + k^2\Psi + a n_e \sigma_{\mathrm{T}}(\theta_b - \theta_\gamma) \,, \\
\mathcal{F}_{\gamma 2}' &= 2\sigma_\gamma' = \frac{8}{15}\theta_\gamma - \frac{3}{5}k\mathcal{F}_{\gamma 3} - \frac{9}{5}a n_e \sigma_{\mathrm{T}}\sigma_\gamma + \frac{1}{10}a n_e\sigma_{\mathrm{T}}(\mathcal{G}_{\gamma 0} + \mathcal{G}_{\gamma 2}) \,, \\
\mathcal{F}_{\gamma\ell}' &= \frac{k}{2\ell+1}\left[\ell\mathcal{F}_{\gamma\ell-1} - (\ell+1)\mathcal{F}_{\gamma\ell+1}\right] - a n_e\sigma_{\mathrm{T}}\mathcal{F}_{\gamma\ell} \,, \qquad \ell \geq 3 \,, \\
\mathcal{G}_{\gamma\ell}' &= \frac{k}{2\ell+1}\left[\ell\mathcal{G}_{\gamma\ell-1} - (\ell+1)\mathcal{G}_{\gamma\ell+1}\right] + \\
&\quad + a n_e\sigma_{\mathrm{T}}\left[-\mathcal{G}_{\gamma\ell} + \frac{1}{2}\left(\mathcal{F}_{\gamma 2} + \mathcal{G}_{\gamma 0} + \mathcal{G}_{\gamma 2}\right)\left(\delta_{\ell 0} + \frac{\delta_{\ell 2}}{5}\right)\right] \,.
\end{aligned} \tag{1.104}$$

Note that, as in the case for neutrinos, there is an infinite Boltzmann hierarchy that also needs to be closed, or solved using integral equations.

Let us take a short detour here. While we have preferred to decompose the phase-space distribution between the background unperturbed value (i.e., the Bose-Einstein distribution f_0 for photons) and a perturbation, we can also expand the Bose-Einstein distribution in terms of a temperature perturbation $\Theta \equiv (T - \bar{T})/\bar{T}$. Then, in this case we have

$$f = f_0\left(\frac{q}{1+\Theta}\right) \,, \tag{1.105}$$

such as, by definition,

$$\Theta = -\left(\frac{\mathrm{d}\log f_0}{\mathrm{d}\log q}\right)^{-1}\varphi \,. \tag{1.106}$$

Since both the gravitational source terms and the linearized collision term in the Boltzmann equation for φ are proportional to the logarithmic derivative of f_0, Θ is independent of q. This means that the photon perturbations still have a Planck spectrum with a temperature that only depends on the photon direction (and not its moment).[11] From the equation above, we see that $\Theta = \mathcal{F}_\gamma/4$, which also relates the photon density and temperature perturbations by the same factor.

1.4.5 Baryons

The last component that we will study are the baryons. This misnomer is motivated by the fact that most of the energy density is dominated by the proton

[11]This does not hold for nonlinear perturbations, which indeed change the spectrum of the CMB. This is the case of for instance the Sunyaev-Zeldovic effect, among many other processes, that generate what is known as spectral distortions: deviations from the blackbody spectrum of the CMB.

and neutron masses (since electrons are much lighter and heavier metals are much less abundant), and by the fact that Coulomb scattering (which couples protons and electrons) has a rate that is much larger than the expansion rate at all times of interest, which makes that the perturbations of all particles are the same. Hence, we will use the subscript 'b' for all of them collectively.

Baryons can be treated as cold and non relativistic, and therefore we will consider only the first two moments of their Boltzmann equations, as we did for dark matter. However, in this case we need to take also into account the coupling with photons due to the Compton scattering. Hence, the left hand side of the Boltzmann equations have the same form than for the dark matter. At the epochs of interest (around and after recombination), the reactions that change the number of electrons and nucleons (e.g., pair production, annihilation, etc.) are rare and therefore irrelevant. This means that there is no source term for the continuity equation, and thus the zero-th moment of the Boltzmann equation is as for cold dark matter,

$$\delta_b' = -\theta_b - 3\Phi' \,. \tag{1.107}$$

While the number of baryons is conserved, their momentum is not, since there is momentum transfer with the photons. The derivation of the second moment is similar than for the dark matter, but instead of weighting the integrals by $\boldsymbol{p}/E$, we use only $\boldsymbol{p}$, which makes the cold dark matter derivation correct if we multiply by a factor of mass m. Since the proton mass vastly dominates, we have

$$m_p \frac{\partial(n_b v_b^j)}{\partial \tau} + 4\mathcal{H} m_p n_b v_b^j + i m_p n_b k^j \Psi = F_{e\gamma}^j \,, \tag{1.108}$$

where m_p is the proton mass and the force density $\boldsymbol{F}_{e\gamma}$ encodes the momentum transfer between photons and electrons due to Compton scattering.[12] Dividing both sides by $\bar{\rho}_b = m_p \bar{n}_b$ we are left with

$$\frac{\partial v_b^j}{\partial \tau} + \mathcal{H} v_b^j + i k^j \Psi = \frac{F_{e\gamma}^j}{\bar{\rho}_b} \,. \tag{1.109}$$

We have left to compute the momentum transfer between photons and electrons. Since momentum is conserved, the force term has to be precisely equal and opposite to the force term in the photon analog of the baryon Euler equation. Therefore, momentum conservation introduces a term $(4\bar{\rho}_\gamma/3\bar{\rho}_b) a n_e \sigma_{\mathrm{T}} (\theta_\gamma - \theta_b)$, where the prefactors in the mean densities come from the different time dependence for each component.

In addition there is another term coming from the baryon sound speed $c_s^2 = \delta P_b/\delta\rho_b$. This is because baryons, although being non relativistic, are not completely cold as dark matter (which we assume it has zero temperature). The finite temperature of baryons introduces this non-zero (although very

[12]Electrons transfer the momentum to the nuclei immediately, and the nuclei-photon interaction is suppressed by a m_e^2/m_p^2 factor, hence neglected.

small) sound speed, which can be neglected in all terms except the acoustic term $c_s^2 k^2 \delta$. The sound speed for baryons depends on the gas temperature, the evolution of which can also be tracked using the first law of thermodynamics. The perturbations of the gas temperatures are therefore coupled to the baryon perturbations and therefore to the whole system to solve, although its effect is limited to very small scales and usually neglected in most studies that involve only linear scales and do not depend directly in the gas temperature.

Then, in Newtonian gauge, we have

$$
\begin{aligned}
\delta_b' &= -\theta_b - 3\Phi' \,, \\
\theta_b' &= -\mathcal{H}\theta_b + k^2\Psi + c_s^2 k^2 \delta_b + \frac{4\bar{\rho}_\gamma}{3\bar{\rho}_b} a n_e \sigma_{\mathrm{T}} (\theta_\gamma - \theta_b) \,.
\end{aligned}
\tag{1.110}
$$

1.4.6 Others

We have considered the standard components of the Universe in the ΛCDM model, but this does not mean that there may be other components and new physics. New components, or new interaction between the standard species can be included in the system of differential equations that describe the evolution of the matter, radiation and metric perturbations in the Universe, following a similar analysis that we have done in this section.

1.5 Initial conditions

So far we have discussed how perturbations evolve in an inhomogeneous expanding Universe, but we need to determine some initial condition to be able to solve the problem. These are determined by the physics of the primordial Universe, usually studied under the umbrella general theory of inflation (the most studied and plausible theory of the primordial Universe), although there are other alternatives such as ekpyrosis or bouncing Universes.

We will not enter to discuss inflation or the generation of the initial conditions, but discuss the essential pieces to our interest. Inflation proposes an exponential expansion of the primordial Universe, which solves some of the problems of the Big Bang, in particular the horizon problem and the flatness problem.

The accelerated expansion can be achieved with negative pressure, but we do not know any kind of matter that fulfills this requirement. The simplest possibility is the potential energy of a scalar field ϕ.[13] The scalar field has an

[13]Indeed, one of the simplest models of dark energy beyond a cosmological constant is quintessence, which is also based on the inclusion of scalar field(s). Note, however, that both quintessence and inflation cannot be trivially explained by the only scalar field we know, the Higgs boson, since its properties are too constrained by now for us to know that we cannot make it work for these purposes.

energy momentum tensor

$$T^{\alpha}_{\beta} = g^{\alpha}\nu \frac{\partial\phi}{\partial x^{\nu}} \frac{\partial\phi}{\partial x^{\beta}} - \delta^{\alpha}_{\beta} \left[\frac{1}{2} g^{\mu}\nu \frac{\partial\phi}{\partial x^{\mu}} \frac{\partial\phi}{\partial x^{\nu}} + V(\phi) \right] , \tag{1.111}$$

where $V(\phi)$ is the potential for the field. For its background behavior, the stress-energy tensor is the diagonal $\{-\rho, P, P, P\}$. For the time-time component $T^0_0 = -\rho$, so that

$$\rho = \frac{\dot{\phi}^2}{2} + V(\phi) \, , \tag{1.112}$$

which are the kinetic and potential energy densities of the field: a homogeneous scalar field has the same dynamics as a single particle in a potential. The pressure $P = T^i_i$ is

$$P = \frac{\dot{\phi}^2}{2} - V(\phi) \, . \tag{1.113}$$

Therefore, if the potential energy is larger than the kinetic energy, a negative pressure is possible. We can see the same in terms of the equation of state

$$w = \frac{P}{\rho} = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)} \, , \tag{1.114}$$

for which we approximate the behaviour (at background level) of a cosmological constant, e.g., $w = -1$, if $V(\phi) \gg \dot{\phi}^2$.

To quantify slow-roll, we can define variables that vanish in the limit in which ϕ is perfectly constant. There are many conventions, but we will use one of the most directly linked to observables. For scalar perturbations the most important inflation variable is

$$\epsilon_{\rm sr} \equiv \frac{\mathrm{d}}{\mathrm{d}t} H^{-1} = -\frac{\dot{H}}{H^2} = -\frac{H'}{aH^2} \, , \tag{1.115}$$

which yields the fractional change during an e-fold in the Hubble rate. Since H decreases, $\epsilon_{\rm sr}$ is always positive. Note that an alternative definition is

$$\epsilon_{\rm sr} - 1 = \frac{\mathrm{d}}{\mathrm{d}\tau} (aH)^{-1} \, . \tag{1.116}$$

1.5.0.1 Primordial perturbations

At any given point and time during inflation, there are small perturbations due to quantum fluctuations of the field against the uniform background. Statistically, the mean fluctuation is null because overdensity regions cancel with underdensities. However, the variance of these perturbations is not zero, and will be the main focus of our study. In principle we would have to specify the predicted perturbations for each species that results from inflation. In general, we can distinguish between two kind of perturbations: adiabatic and isocurvature perturbations.

Adiabatic perturbations fulfill that the local state of matter at some space time point of the perturbed Universe is the same as in the background at some slightly different time (where the time shift varies with the location). One way to understand adiabatic perturbations is to interpret that some regions of the Universe are *ahead* or *more evolved* than others. This local shift in time is *common* to all species involved, fulfilling that

$$\delta\rho(\tau, \boldsymbol{x}) = \bar{\rho}(\tau + \delta\tau(\boldsymbol{x}), \boldsymbol{x}) = \bar{\rho}'(\tau)\delta\tau(\boldsymbol{x}) \tag{1.117}$$

for all species, which means that

$$\frac{\delta\rho_x}{\bar{\rho}'_x} = \frac{\delta\rho_y}{\bar{\rho}'_y}\,. \tag{1.118}$$

Neglecting any energy transfer between fluid components at the background level, $\bar{\rho}'_x = -3\mathcal{H}(1+w_x)\bar{\rho}_x$, so that

$$\frac{\delta_x}{1+w_x} = \frac{\delta_y}{1+w_y}\,. \tag{1.119}$$

Thus, all matter species have the same fractional perturbations, while all radiation and relativistic species obey $\delta_\gamma = 4\delta_m/3$, since $w_\gamma = 1/3$. The relation for the velocity divergences is analog.

On the other hand, instead of corresponding to a change in the total energy density, isocurvature perturbations correspond to perturbations between different species that explicitly leave the total perturbations unchanged. Therefore, isocurvature perturbations can be defined as

$$S_{xy} = \frac{\delta_x}{1+w_x} - \frac{\delta_y}{1+w_y}\,. \tag{1.120}$$

There are different sets of isocurvature perturbations, usually defined with respect to the photon perturbations (e.g., neutrino isocurvature perturbations involve neutrino and photons in the expression above).

Single-field inflation, since it involves a single clock (scalar field), predicts only the generation of adiabatic perturbations. This is because any point during inflation is completely characterized by the value of the single scalar field involved. Actually, isocurvature perturbations are very constrained by current observations of the CMB anisotropies. Some exceptions are compensated dark-matter-baryon isocurvature perturbations (i.e., isocurvature perturbations involving *only* dark matter and baryons). Anyways, since we are focusing on single-field inflation, we restrict the discussion to adiabatic perturbations and we only need to derive $\delta\rho$. We can therefore specify the initial conditions in terms of a single metric perturbations.

The power spectrum of the perturbations of the scalar field within the horizon is

$$P_{\delta\phi} = \frac{1}{2k^3a^2\tau^2} = \left(\frac{H^2}{2k^3}\right)_{\text{hor. cross.}}\,. \tag{1.121}$$

where we have used that during inflation, assuming a constant Hubble rate, $\tau \simeq -1/Ha = -1/\mathcal{H}$. However, as inflation progresses, a connection between $\delta\phi$ and Ψ arises and freezes outside the horizon. Therefore, the primordial metric perturbations are determined by the scalar field perturbations at horizon crossing.

Let us define the curvature perturbation

$$\mathcal{R}(\boldsymbol{k},\tau) \equiv \frac{ik_i \delta T^i_0(\boldsymbol{k},\tau) a^2 H(\tau)}{k^2 \left[\bar{\rho} + \bar{P}\right](\tau)} - \Psi(\boldsymbol{k},\tau)\,. \tag{1.122}$$

During inflation, the first term dominates, and applying the energy-momentum tensor we have

$$\mathcal{R} = -\frac{aH}{\bar{\phi}'}\delta\phi\,, \qquad \text{(during inflation)}\,. \tag{1.123}$$

Enough time after inflation ends, when we are fully in the radiation dominated epoch, the stress-energy tensor is fully dominated by radiation and neutrinos (denoted by 'r' collectively). Using the equation of state of radiation in the denominator of $\mathcal{R}$, we have

$$\mathcal{R} = -\frac{3aH\mathcal{F}_{r1}}{4k} - \Psi = -\frac{3}{2}\Psi\,, \qquad \text{(post inflation ; rad. dom)}\,, \tag{1.124}$$

where we will discuss the last equality in a bit.

It can be demonstrated that $\mathcal{R}$ (which a gauge invariant quantiy) is conserved outside the horizon.[14] Therefore, the value of $\mathcal{R}$ is determined at horizon crossing, and we can therefore relate superhorizon values of Ψ with $\delta\phi$, finding

$$\left(P_\Psi\right)_{\text{post}}(k) = \left(P_\Phi\right)_{\text{post}}(k) = \frac{8\pi G}{9k^3}\left(\frac{H^2}{\epsilon_{\text{sr}}}\right)^2_{aH=k}\,, \tag{1.125}$$

where we also use the non-anisotropic stress quality of $\Phi = -\Psi$. Similarly, the curvature power spectrum is

$$P_{\mathcal{R}}(k) = \left(\frac{aH}{\bar{\phi}'}\right)^2 P_{\delta\phi}(k) = \left(\frac{2\pi G H^2}{\epsilon_{\text{sr}} k^3}\right)_{aH=k}\,. \tag{1.126}$$

In natural units, the Planck mass $M_{\text{P}} = G^{-1/2}$, and let us rephrase

$$P_{\mathcal{R}}(k) = \left(\frac{2\pi H^2}{\epsilon_{\text{sr}} M_{\text{P}}^2 k^3}\right)_{aH=k} \equiv 2\pi^2 \mathcal{A}_{\text{s}} k^{-3}\left(\frac{k}{k_{\text{p}}}\right)^{n_{\text{s}}-1}\,, \tag{1.127}$$

where $\mathcal{A}_{\text{s}}$ is the variance of curvature perturbations in a logarithmic wavenumber interval centered around the pivot scale k_{p} and n_{s} is the scalar spectra

[14]For this computation, the spatially flat gauge is very convenient.

index. For the Planck convention, $k_{\rm p} = 0.05\,{\rm Mpc}^{-1}$, $\mathcal{A}_{\rm s} = 2.1 \times 10^{-9}$, which corresponds to a perturbation amplitude $\sim 4.6 \times 10^{-5}$, of similar order of magnitude than the temperature fluctuations in the CMB.

We can describe the primordial power spectrum from the slope of $k^3 P_\Phi$. For instance, if it is constant, it is called a scale-invariant power spectrum. However, there is a small deviation from scale invariance, due to small changes in the slow-roll parameter. The field rolls down the potential slowly in such a way that the Hubble rate, nearly constant, decreases very slowly. This makes that the actual power spectrum is red-tilted, with the larger-scale perturbations (those that leave the horizon earlier) slightly larger than the smaller-scale ones. This feature has been confirmed by CMB observations, where the slope can be related to $\epsilon_{\rm sr}$ and another inflation parameter $\delta_{\rm sr}$.

1.5.0.2 Primordial matter and radiation perturbations

The only piece left is to relate the metric perturbations after inflation to the matter and radiation perturbations. These are the initial conditions for the system of differential equations derived before. Thanks to the fact that the primordial perturbations in single-field inflation are adiabatic, this derivation is significantly simplified.

We can start by taking the large-scale limit in the Boltzmann equation for the momentum-averaged perturbation of the phase-space distribution of photons from the previous chapter (Eqs. (1.93) and (1.102)). In this limit, $\mathcal{F}'_\gamma \sim \mathcal{F}_\gamma/\tau$, while $ik\mu\mathcal{F}_\gamma \sim k\mathcal{F}_\gamma$: the former is larger than the latter by a $1/k\tau$ factor, which in this limit is very large. This argument allows us to neglect any factor multiplied by k in the Boltzmann equation. Physically, this means that the scales under consideration are much larger than the size of the horizon and therefore are not causally connected. In this regime, only gravity is relevant: dark matter and baryons behave similarly, and their velocities are smaller than overdensity by the same factor $k\tau$. Furthermore, an observer within their causal horizon would only see a uniform sky, so that higher multipoles of the phase space distribution perturbation are negligible. Therefore, we have for radiation (photons and neutrinos alike),

$$\mathcal{F}_{r0} + 4\Phi' = 0\,, \tag{1.128}$$

and for the non-relativistic matter

$$\delta'_c = -3\Phi'\,. \tag{1.129}$$

Note that since we consider adiabatic perturbations and large scales, $\mathcal{F}_{\gamma 0} = \mathcal{F}_{\nu 0}$ and $\delta_c = \delta_b$.

Now we focus on the Einstein equations from Eq. (1.56). The k^2 term can be neglected in this limit and assuming that all energy density is given by radiation (radiation-domination epoch),

$$3\mathcal{H}\left(\Phi' - \mathcal{H}\Psi\right) = 4\pi G a^2 \bar{\rho}_r \mathcal{F}_{r0}\,. \tag{1.130}$$

During radiation domination, $a \propto \tau$, so that $\mathcal{H} = 1/\tau$, and

$$\frac{\Phi'}{\tau} - \frac{\Psi}{\tau^2} = 4\pi G a^2 \bar{\rho}_r \mathcal{F}_{r0} = \frac{\mathcal{F}_{r0}}{2\tau^2}\,, \tag{1.131}$$

where the last equality uses the Friedmann equation. Multiplying by τ^2, differentiating both sides and using $\mathcal{F}_{r0} = -4\Phi'$ we have

$$\Phi''\tau + \Phi' - \Psi' = -2\Phi' \Longrightarrow \Phi''\tau + 4\Phi' = 0\,, \tag{1.132}$$

where the last part neglects anisotropic stress, hence $\Phi = -\Psi$. Inserting the ansatz of $\Phi = \tau^p$ we have

$$p(p-1) + 4p = 0\,, \tag{1.133}$$

which has $p = -3$ and $p = 0$ as solutions. $p = -3$ is a decaying mode, so that it will quickly vanish without contributing to the growth of perturbations. Therefore, we focus on $p = 0$. In this case, from Eq. (1.131), after multiplying by τ^2 and under the same assumptions, for the initial time τ_i

$$\Phi = \frac{\mathcal{F}_{r0}}{2} = \frac{\mathcal{F}_{\gamma 0}}{2} = \frac{\mathcal{F}_{\nu 0}}{2}\,. \tag{1.134}$$

For dark matter and baryons we have $\delta_c(\boldsymbol{k}) = \delta_b(\boldsymbol{k}) = 3\mathcal{F}_{\gamma 0}(\boldsymbol{k})/4+\text{constant}(\boldsymbol{k})$. Since adiabatic perturbations must have a uniform matter-to-radiation ratio is

$$\frac{n_c}{n_\gamma} = \frac{\bar{n}_c}{\bar{n}_\gamma}\left[\frac{1+\delta_c}{1+3\mathcal{F}_{\gamma 0}/4}\right], \tag{1.135}$$

where the 3/4 factor for the photon perturbations comes from changing from energy density to number density (at linear order). The combination in the brackets linearizes to $1 + \delta_c - 3\mathcal{F}_{\gamma 0}/4$ must therefore be independent of the position, which forces the constant above to be null for the perturbations to sum up to zero.

From the space-time component of the Einstein equation (Eq. (1.57)) we can get the initial condition for the velocities, using that $\rho_r \gg \rho_m$ and neglecting the k^2 term. We find

$$\mathcal{F}_{\gamma 1} = \mathcal{F}_{\nu 1} = \frac{4\theta_c}{3k} = \frac{4\theta_b}{3k} = -\frac{k}{6aH}\Phi\,, \tag{1.136}$$

which returns the $\mathcal{R} = -3\Psi/2$ we used in the previous section.

CHAPTER 2

LECTURE 1: GROWTH OF STRUCTURES

The preliminary material discussed the equations describing the evolution of matter, radiation and metric perturbations in the Universe to linear order, as well as the primordial perturbations resulting from inflation that determine the initial conditions for the evolution equations. The evolution of perturbations consists of a system of coupled differential equations (metric perturbations depend on the total stress-energy tensor, which receive contributions from the perturbations of each independent component of the Universe, the evolution of which is in turn determined by the metric perturbations). Usually, this system is solved by Boltzmann codes like CLASS (2) or CAMB (3). In this chapter we want to get a qualiative understanding of the growth of matter perturbations, which are the ones that will determine the distribution of galaxies in the late Universe.[1]

In general, we can use the Poisson equation to relate the gravitational potential with the matter perturbations, which is correct for perturbations

[1]More details can be found in *Modern Cosmology* (Ref. (1)), which we closely follow in this chapter, homogenizing nomenclature and altering slightly the order of the discussion.

well within the horizon and in matter domination

$$k^2\Phi = 4\pi G a^2 \bar{\rho}_{\rm m}\delta_{\rm m}\,, \qquad (a \gg a_{\rm eq}, k \gg aH)\,. \tag{2.1}$$

Turning the background matter density using the density parameter and the critical density and the definition of the latter, we can express the matter overdensity as

$$\delta_{\rm m} = \frac{2k^2 a}{3\Omega_{\rm m} H_0^2}\Phi\,, \qquad (a \gg a_{\rm eq}, k \gg aH)\,. \tag{2.2}$$

This kind of conversion will appear many times in this chapter.

Here we attempt to get an approximate description of the growth of perturbations, hence we will reduce significantly the number of equations and limit ourselves to specific limits and regimes. Remember that before recombination, the photon distribution can be characterized by only the monopole and dipole of the momentum-averaged distribution, since all other moments are suppressed due to the tight coupling between photons and baryons. This breaks down after recombination, but at that time the photon perturbations play a negligible role in the growth of structures since the energy-density of the Univese is totally dominated by non-relativistic matter.[2] We will also neglect high multipoles of neutrinos. This is a bad approximation, since neutrinos free stream and are never tightly coupled, but it is better than neglecting them completely. Therefore we will consider the monopole and dipole of the whole relativistic species, photons and neutrinos, all together.

Tight-coupling also allows us to eliminate baryons from the Boltzmann equations, if we are only interested in the qualitative evolution of matter perturbations. This is because the collision term for photons can be neglected in the limit of small baryon density (with respect to photons).[3] Similarly, we will consider than matter perturbations are entirerly determined by cold dark matter.

In this limit, the photon distribution reduces to two equations for the monopole and dipole. Therefore, considering only cold dark matter and total radiation in this limit, we have the following set of differential equations for

[2]Following the evolution of the whole photon phase-space distribution is required to understand CMB observations, as primary anisotropies propagate through an evolving Universe, and also to model secondary anisotropies accurately. We will study this problem in the next chapter.

[3]First, since the quadrupole and the polarization are very small, we can neglect the terms multiplying $\mathcal{P}_2$. Then we can show the collision term is proportional to the baryon-to-photon energy ratio $R \equiv 3\bar{\rho}_b/4\bar{\rho}_\gamma$.

matter and radiation:

$$\begin{aligned}
&\mathcal{F}'_{r0} + k\mathcal{F}_{r1} = -4\Phi' \,, \\
&\mathcal{F}'_{r1} - \frac{k}{3}\mathcal{F}_{r0} = -\frac{4k}{3}\Phi \,, \\
&\delta'_c + \theta_c = -3\Phi' \,, \\
&\theta'_c + \mathcal{H}\theta_c = -k^2\Phi \,.
\end{aligned} \tag{2.3}$$

Under these approximations, there is no anisotropic stress, thus $\Phi = -\Psi$ (as used above). Then, we have the time-time component for the Einstein equations (Eq. (1.56)) and the redundant equation from the combination of this one and the time-space components (Eq. (1.57)) to describe metric perturbations and their relations with matter and radiation:[4]

$$\begin{aligned}
&k^2\Phi + 3\mathcal{H}\left(\Phi' + \mathcal{H}\Phi\right) = 4\pi G a^2 (\bar{\rho}_c\delta_c + \bar{\rho}_r\mathcal{F}_{r0}) \,, \\
&k^2\Phi = 4\pi G a^2 \left[\bar{\rho}_c\delta_c + \bar{\rho}_r\mathcal{F}_{r0} + \frac{3\mathcal{H}}{k}\left(\frac{\bar{\rho}_c\theta_c}{k} + \bar{\rho}_r\mathcal{F}_{r1})\right)\right] \,.
\end{aligned} \tag{2.4}$$

This set of 5 differential equations is very easy to solve numerically. Analytical solutions are harder to obtain since there is no analytic solution valid on all scales at all times. We need to take limits and specific regimes to study individual pieces of the cosmic evolution and patch them together afterwards.

We will study large scales (matter-radiation transition while outside the horizon and horizon crossing during matter domination) and small scales (horizon crossing during radiation-dominated era and matter-radiation transition within the horizon) analytically. We cannot treat analytically modes that enter the horizon around the epoch of equality, and numerical solutions solving the Boltzmann equations are required, but the physics are similar.

The approximations to obtain the equations above are rough. We have neglected the effects of baryons, which are $\sim 16\%$ of matter in the Universe, and the mass of neutrinos (as well as high multipoles of the neutrino and photon perturbations). We will indicate the impact of these additional components as we progress in the chapter.

2.1 Large scales

We can distinguish two different regimes for the large scales. First, the transition from radiation to matter domination takes place while the perturbations are outside the horizon. Second, perturbations enter the horizon already in the matter domination.

[4]Note that only one of them is needed to close the Boltzmann system, since we already fix $\Phi = -\Psi$.

2.1.1 Super-horizon solutions

Consider modes far outside the horizon, $k\tau \ll 1$: then we can drop all terms depending on k, which shows that velocities decouple from the system, leaving only three equations to solve.[5] We are left with

$$\begin{aligned} &\mathcal{F}'_{r0} = -4\Phi' \,, \qquad\qquad \delta'_c = -3\Phi' \,, \\ &3\mathcal{H}\left(\Phi' + \mathcal{H}\Phi\right) = 4\pi G a^2 (\bar{\rho}_c\delta_c + \bar{\rho}_r\mathcal{F}_{r0}) \,, \end{aligned} \tag{2.5}$$

with the first two equations showing that the combination $3\delta_c - 4\mathcal{F}_{r0}$ is constant, and zero (since they are adiabatic perturbations, see the discussion below Eq. (1.135)). Therefore we drop the equation for radiation. If we now introduce

$$y \equiv \frac{a}{a_{\rm eq}} = \frac{\bar{\rho}_m}{\bar{\rho}_r} \,, \tag{2.6}$$

and use it as evolution variable, rather than τ or a.[6] Then, the Einstein equations become

$$\begin{aligned} &3\mathcal{H}\left(\Phi' + \mathcal{H}\Phi\right) = 4\pi G a^2 \bar{\rho}_c\delta_c \left(1 + \frac{4}{3y}\right) \Longrightarrow \\ &y\frac{\mathrm{d}\Phi}{\mathrm{d}y} + \Phi = \frac{y}{2(y+1)}\delta_c\left(1+\frac{4}{3y}\right) = \frac{3y+4}{6(y+1)}\delta_c \,, \end{aligned} \tag{2.7}$$

where we have used $\mathrm{d}/\mathrm{d}\tau = \mathcal{H}y\mathrm{d}/\mathrm{d}y$, $a' = a\mathcal{H}$, and the last equality uses the Friedmann equation as function of y. Using the dark-matter equation we have $\mathrm{d}\delta_c/\mathrm{d}y = -3\mathrm{d}\Phi/\mathrm{d}y$. Then if we express the equation above as an equation for δ_c and derive with respect to y to get $\mathrm{d}\delta_c/\mathrm{d}y$ we have

$$\begin{aligned} &-3\frac{\mathrm{d}\Phi}{\mathrm{d}y} = \frac{\mathrm{d}}{\mathrm{d}y}\left[\frac{6(y+1)}{3y+4}\left\{y\frac{\mathrm{d}\Phi}{\mathrm{d}y} + \Phi\right\}\right] \Longrightarrow \\ &\frac{\mathrm{d}^2\Phi}{\mathrm{d}^2y} + \frac{21y^2+54y+32}{2y(y+1)(3y+4)}\frac{\mathrm{d}\Phi}{\mathrm{d}y} + \frac{\Phi}{y(y+1)(3y+4)} = 0 \,. \end{aligned} \tag{2.8}$$

Kodama and Sasaki found a solution to this equation in 1984 introducing a new variable

$$u \equiv \frac{y^3}{\sqrt{1+y}}\Phi \,, \tag{2.9}$$

which turns the equation above into

$$\frac{\mathrm{d}^2u}{\mathrm{d}y^2} + \frac{\mathrm{d}u}{\mathrm{d}y}\left[-\frac{2}{y} + \frac{3/2}{1+y} - \frac{3}{3y+4}\right] = 0 \,, \tag{2.10}$$

[5] Remember that $\theta = ikv$, hence we also neglect θ terms here.

[6] We could use $\bar{\rho}_c$ in the numerator to get a slightly more accurate solution, since we are ignoring baryons. But this is not necessary since we are aiming for a qualitative result anyways.

where there is no term proportional to u and leaves a first-order equation that is integrable. Denoting here $u' \equiv \mathrm{d}u/\mathrm{d}y$ to ease the notation, we have

$$\begin{aligned}
&\frac{\mathrm{d}u'}{u'} = \mathrm{d}\left[\frac{2}{y} - \frac{3/2}{1+y} + \frac{3}{3y+4}\right] \Longrightarrow \\
&\log u' = 2\log y - \frac{3}{2}\log(y+1) + \log(3y+4) + \text{constant} \Longrightarrow \\
&u' = A\frac{y^2(3y+4)}{(1+y)^{3/2}}\,,
\end{aligned} \tag{2.11}$$

where A is an integration constant to be found. Using the definition of u we can integrate the expression above to get

$$\frac{y^3}{\sqrt{y+1}}\Phi = A\int_0^y \mathrm{d}\tilde{y}\frac{\tilde{y}^2(3\tilde{y}+4)}{(1+\tilde{y})^{3/2}}\,, \tag{2.12}$$

where we have already eliminated the second integration constant since $y^3\Phi \to 0$ as $y \to 0$ (e.g., early times). The second constant can be obtained approximating the integrand in the small y limit, for which we obtain that $\Phi = 4A/3$, thus $A = 3\Phi(0)/4$. The integral above has an analytical solution, which leaves

$$\Phi = \frac{1}{10y^3}\left(16\sqrt{1+y} + 9y^3 + 2y^2 - 8y - 16\right)\Phi(0)\,. \tag{2.13}$$

Although it is not obvious, this expression fulfills that at small y, $\Phi = \Phi(0)$. At large y, in turn, once matter dominates, y^3 terms dominates and we find $\Phi = 9\Phi(0)/10$. This means that even at the largest scales, those which never enter the horizon, the gravitational potential drops a factor $9/10$ as the Universe undergoes the matter-radiation transition. Remembering that after inflation $\mathcal{R} = 3/2\Phi$, we obtain an important result for super-horizon scales

$$(\Phi(\boldsymbol{k},\tau))_{\mathrm{super-horizon}} = \begin{cases} \frac{2}{3}\mathcal{R}(\boldsymbol{k})\,, & \text{(radiation domination)}\,, \\ \frac{3}{5}\mathcal{R}(\boldsymbol{k})\,, & \text{(matter domination)}\,. \end{cases} \tag{2.14}$$

We have provided solutions in two limiting times, but the transition between pure radiation and pure matter domination epochs is very long.

Finally this analytic limit solution works reasonably well when compared with numerical results. The main difference is due to the neutrino quadrupole, which introduces a small anisotropic stress and therefore a small slip in the gravitional potentials (i.e., $\Phi \neq -\Psi$). Accounting for this effect drops the $9/10$ factor to $\simeq 0.86$.

2.1.2 Horizon crossing

Large scales enter the horizon already in the matter-domination epoch. We have studied their evolution outside the horizon, and now we want to show that also within the horizon the gravitational potential does not evolve over time.

Let us go back to our set of 5 differencial equations from Eqs. (2.3) and (2.4), and focus on scales within the horizon during matter domination. Therefore, we can neglect any role from radiation components, and we keep now the second of the two Einstein equations, which allow us to substitute Φ in the two differential equations for the cold dark matter.

Now we have a set of two differential equations, but we can also add some prior knowledge about the initial conditions: we know that deep in the matter-domination epoch, the gravitational potential on super horizon scales is constant. Therefore, we can set $\Phi' = 0$ as our initial condition. Therefore we need to check if the set of equations admits a solution with constant Φ:

$$\begin{aligned} &\delta_c' + \theta_c = 0\,, \\ &\theta_c' + \mathcal{H}\theta_c = -k^2\Phi\,, \\ &k^2\Phi = \frac{3}{2}\mathcal{H}^2\left[\delta_c + \frac{3\mathcal{H}\theta_c}{k^2}\right]\,, \end{aligned} \tag{2.15}$$

where we have used the Friedmann equation to simplify the last expression. In the matter-dominated era, $H \propto a^{-3/2}$, so that $\mathrm{d}\mathcal{H}/\mathrm{d}\tau = -\mathcal{H}^2/2$. We use the last equation above to obtain δ_c as function of Φ and θ_c and substitute in the first equation, obtaining

$$\frac{2k^2\Phi'}{3\mathcal{H}^2} + \frac{2k^2\Phi}{3\mathcal{H}} - \frac{3\mathcal{H}\theta_c'}{k^2} + \frac{3\mathcal{H}^2\theta_c}{2k^2} + \theta_c = 0\,. \tag{2.16}$$

Now we can use the equation for θ_c' to obtain a second order equation on Φ. We substitute θ_c' above obtaining

$$\frac{2k^2\Phi'}{3\mathcal{H}^2} + \left[\frac{\theta_c}{k^2} + \frac{2\Phi}{3\mathcal{H}}\right]\left(\frac{9\mathcal{H}^2}{2} + k^2\right) = 0\,. \tag{2.17}$$

One condition for constant Φ to be a solution of the system is if we obtain a second-order equation for Φ of the form $\alpha\Phi'' + \beta\Phi' = 0$. Therefore, we can test if Φ constant is a solution by deriving the expression above as function of τ and dropping terms proportional to derivatives of Φ. Using the fact that the conformal time derivative of $\mathcal{H}^{-1}$ is $1/2$ during matter domination and again the equation for θ_c', we see that the remaining terms are

$$-\left[\frac{\mathcal{H}\theta_c}{k^2} + \frac{2\Phi}{3}\right](9\mathcal{H}^2 + k^2) = 0\,, \tag{2.18}$$

where the term in square brackets can be identified with the one in the previous expression, which is proportional to Φ'. Therefore, there is no term proportional Φ and Φ =constant is a valid solution for the system in the matter-domination era. Since it comes also from an initial condition, Φ =constant is *the* solution. The other solution to the system involves a decaying solution, thus not relevant to the problem at hand.

Therefore, gravitational potentials remain constant inside of the horizon during matter-domination era. This means that the matter accretion (which makes the potential grow) and the expansion of the Universe (which dilutes the potential) exactly counteract each other. When dark energy becomes relevant, accelerating the expansion of the Universe, makes the latter dominate and potentials will decay.

In this situation, since the gravitational potential is constant and we are in matter domination and well within the horizon, we can use Eq. (2.2) to relate the potential and the matter perturbations to find that matter perturbations grow as $\propto a$.

2.2 Small scales

We have broken the study of large scales perturbations as the matter-radiation transition outside of the horizon, and the horizon crossing during matter domination. The situation for small scales is mirrored: perturbations enter the horizon during radiation domination, and they experience the transition to matter domination when they are well within the horizon.

2.2.1 Horizon crossing

During radiation domination, matter perturbations are determined by the gravitational potential, but they are not significant to influence it back, since the energy density is dominated by radiation. Therefore, the gravitational potential is influenced by radiation perturbations, and it determines the matter perturbations. The study of the dark matter perturbations in this regime requires a two step process: solve the radiation and potential perturbations, and then translate these into matter perturbations. To start we take the radiation equations in Eq. (2.3) and the second Einstein equation in Eq. (2.4) dropping the matter terms, which leaves

$$\Phi = \frac{3\mathcal{H}^2}{2k^2}\left[\mathcal{F}_{r0} + \frac{3\mathcal{H}}{k}\mathcal{F}_{r1}\right], \tag{2.19}$$

where as before we have substituted $\bar{\rho}_r$ using the Friedmann equation. Furthermore, in radiation domination, $\mathcal{H} = 1/\tau$, and substituting $\mathcal{F}_{r0}$ by Φ and $\mathcal{F}_{r1}$ using the equation above in the radiation equations we find

$$\begin{aligned} -\frac{3}{k\tau}\mathcal{F}'_{r1} + k\mathcal{F}_{r1}\left[1 + \frac{3}{k^2\tau^2}\right] &= -4\Phi'\left[1 + \frac{k^2\tau^2}{6}\right] - \frac{4k^2\tau}{3}\Phi\,, \\ \mathcal{F}'_{r1} + \frac{1}{\tau}\mathcal{F}_{r1} &= -\frac{4k}{3}\Phi\left[1 - \frac{k^2\tau^2}{6}\right]. \end{aligned} \tag{2.20}$$

As done before, we will turn these two first-order equations into a second order for Φ. We can use the second equation to express $\mathcal{F}'_{r1}$ as function of Φ and $\mathcal{F}_{r1}$, and substitute in the first equation, which is left as

$$\Phi' + \frac{1}{\tau}\Phi = -\frac{3}{2k\tau^2}\mathcal{F}_{r1}\,. \tag{2.21}$$

Now we can differentiate, and remove terms depending on $\mathcal{F}_{r1}$ and $\mathcal{F}'_{r1}$ with the expression above. We find

$$\Phi'' + \frac{4}{\tau}\Phi' + \frac{k^2}{3}\Phi = 0\,, \tag{2.22}$$

which is the wave equation in Fourier space with a damping term due to the expansion of the Universe. This implies oscillatory solutions, which must be connected to the initial condition of a constant Φ (before horizon crossing). Therefore, let us define $u \equiv \Phi\tau$, such as

$$u'' + \frac{2}{\tau}u' + \left(\frac{k^2}{3} - \frac{2}{\tau^2}\right)u = 0\,. \tag{2.23}$$

This is the Bessel equation of order 1, with solutions $j_1(k\tau/\sqrt{3})$ (the spherical Bessel function) and $n_1(k\tau/\sqrt{3})$ (the spherical Neumann function). The latter diverges as $\tau \to 0$, so that we must discard it due to the initial conditions. We can use the exact expression for $j_1(x) = (\sin x - x\cos x)/x^3$, which tends to $1/3$ as $x \to 0$. Since $\Phi(0) = 2\mathcal{R}/3$, we obtain

$$\Phi(\boldsymbol{k}, \tau) = 2\frac{j_1(k\tau/\sqrt{3})}{k\tau/\sqrt{3}}\mathcal{R}(\boldsymbol{k})\,. \tag{2.24}$$

As soon as the mode enters the horizon during radiation-dominated era, its potential starts to decay and oscillate. Effectively, the solution corresponds to a damped standing wave in Fourier space. Physically, this is because radiation pressure counteracts (and overcomes) gravity, preventing overdensities to grow. This is evident from Eq. (2.19), ignoring the dipole (which is much smaller than the monopole within the horizon): since $\mathcal{F}_{r0}$ oscillates with fixed amplitude, the potential also oscillates but proportionally to $\mathcal{H}^2 \propto \tau^{-2}$.

Neglecting the influence of dark matter induces an error in the evolution of the gravitational potential at large scales, due to its gravitational effect. The effect of free-streaming neutrinos leads to additional damping of the potential after horizon crossing.

Now we can determine the evolution of the cold dark matter perturbations, which are determined by Φ, following Eq. (2.3). Merging both equations, we find (using that $\mathcal{H} = 1/\tau$ in radiation domination)

$$\delta_c'' + \frac{1}{\tau}\delta_c' = S = -3\Phi'' + k^2\Phi - \frac{3}{\tau}\Phi'\,. \tag{2.25}$$

Two solutions to the homogeneous equation (i.e., having the source term $S = 0$) are δ_c =constant and $\delta_c = \log \tau$. Therefore, we anticipate a logarithmic growth of the matter perturbations within the horizon in the radiation-dominated epoch.

Remember that the solution to a second-order equation is the linear combination of the two homogeneous solutions and a particular solution. In this case, we do not have prior intuition about the particular solution, so we can construct it from the two homogeneous solutions (denoted by s_1 and s_2) and the source term. Such solution is the integral of the source term weighted by the Green function $[s_1(\tau)s_2(\tilde{\tau}) - s_1(\tilde{\tau})s_2(\tau)]/[s_1'(\tilde{\tau})s_2(\tilde{\tau}) - s_1(\tilde{\tau})s_2'(\tilde{\tau})]$. So here we have (adding factor of k to the arguments of the logarithms, since they will be convenient later)

$$\delta_c = C_1 + C_2 \log(k\tau) - \int_0^{\tau} \mathrm{d}\tilde{\tau} S(k, \tilde{\tau})\tilde{\tau}(\log(k\tilde{\tau}) - \log(k\tau)) \,. \tag{2.26}$$

At very early times, the integral can be neglected, and matching the initial condition ($\delta_c = \mathcal{R}$, from previous chapter), we find $C_2 = 0$ and $C_1 = \mathcal{R}$. S decays as it enters the horizon (since the potential does), hence most of the contribution to the integral comes from $k\tau \sim 1$. Therefore, the first integral will asymptote to a constant, and the second one will lead to a term proportional to $\log(k\tau)$. Therefore, after entering the horizon

$$\delta_c = A\mathcal{R} \log(Bk\tau) \,, \tag{2.27}$$

which is a constant plus a logarithmic growing mode. The constant term is C_1 plus the first integral, while the logarithmic term is the second integral:

$$\begin{aligned} A\mathcal{R} \log B &= \mathcal{R} - \int_0^{\infty} \mathrm{d}\tilde{\tau} S(k, \tilde{\tau})\tilde{\tau} \log(k\tilde{\tau}) \,, \\ A\mathcal{R} &= \int_0^{\infty} \mathrm{d}\tilde{\tau} S(k, \tilde{\tau})\tilde{\tau} \,. \end{aligned} \tag{2.28}$$

The upper limit set to infinity is allowed since the potential decays (and thus S) and the integrand vanishes at large τ. Solving this equations return $A = 6$ and $B = 0.44$. A more precise treatment, using more precise expressions for the potentials, leads to slightly different values, as found by Hu and Sugiyama in 1996.

In summary, dark matter perturbations grow even during radiation-domination era. This is in contrast of the radiation perturbations, which oscillate with constant amplitude (determining the decay of the potential) and the baryon perturbations, which are tightly coupled to photons. This is because cold dark matter does not feel any pressure that counteracts the effect of gravity, hence even if the gravitational potential decays and the Universe expands faster it keeps clustering (although not as fast as during matter domination era, where the constant potential implied $\delta_c \propto a$). As the Universe gets closer to matter domination, the expansion slows down and the perturbations start to grow faster. This growth eventually makes that the matter perturbations must be taken into account (i.e., $\bar{\rho}_c\delta_c \sim \bar{\rho}_r\mathcal{F}_{r0}$), which produces the small offset at large scales in our prediction for the gravitational potential inside the horizon during radiation domination mentioned before.

2.2.2 Sub-horizon evolution across the matter-radiation transition

As we mentioned, even during radiation domination, the growth of matter perturbations joint to the fact that the radiation perturbations oscillate at fixed amplitude eventually leads to $\bar{\rho}_c\delta_c \sim \bar{\rho}_r\mathcal{F}_{r0}$ even if $\bar{\rho}_c < \bar{\rho}_r$. Once this point is reached, the gravitational potential is determined by the matter perturbations independently of the radiation perturbations. Therefore, $\mathcal{F}_r$ can be ignored. In this subsection we will solve the evolution of perturbations in this regime and match it to the logarithmic growth from the previous subsection, which happened when the potential decays.

We start from Eq. (2.3), neglecting the role from radiation in this case, and the second Einstein equation in Eq. (2.4), and once again we want to get to a second order equation from a system of three equations. In this regime, the sub-horizon dark-matter perturbations experience the matter-radiation transition, so we will use again the variable y defined in Eq. (2.6) as the evolution variable. The three equations therefore become

$$\begin{aligned}
&\frac{\mathrm{d}\delta_c}{\mathrm{d}y} + \frac{\theta_c}{\mathcal{H}y} = -3\frac{\mathrm{d}\Phi}{\mathrm{d}y}\,, \\
&\frac{\mathrm{d}\theta_c}{\mathrm{d}y} + \frac{\theta_c}{y} = -\frac{k^2\Phi}{\mathcal{H}y}\,, \\
&k^2\Phi = \frac{3\mathcal{H}y}{2(y+1)}\delta_c\,.
\end{aligned} \tag{2.29}$$

Note that expressed in this way the gravitational potential only depends on δ_c and not on the velocity divergence because perturbations are well within the horizon and terms that are divided by $\mathcal{H}/k \ll 1$ (remember that $\theta = ikv$). Following the same routine as above, we differentiate the first equation above to get

$$\frac{\mathrm{d}^2\delta_c}{\mathrm{d}y^2} - \frac{(2+3y)\theta_c}{2\mathcal{H}y^2(1+y)} = -3\frac{\mathrm{d}^2\Phi}{\mathrm{d}y^2} + \frac{k^2\Phi}{\mathcal{H}^2y^2}\,, \tag{2.30}$$

where we have used the second equation above to substitute the derivative of θ_c, and considered that $\mathrm{d}(\mathcal{H}y)^{-1}/\mathrm{d}y = -(1+y)^{-1}(2\mathcal{H}y)^{-1}$. The first term in the right is much smaller than the second one, which has a $k^2/\mathcal{H}^2$ factor, hence we drop it, and we can substitute the second term using the Einstein equation above. Using the first equation for δ_c we can substitute the θ_c factor (neglecting the potential, which is much smaller than δ_c within the horizon, according to the Poisson equation). Thus, we have

$$\frac{\mathrm{d}^2\delta_c}{\mathrm{d}y^2}+2\frac{(2+3y)}{2\mathcal{H}y^2(1+y)}\frac{\mathrm{d}\delta_c}{\mathrm{d}y}-\frac{3}{2y(y+1)}\delta_c=0\,, \tag{2.31}$$

which is known as the Meszaros equation, and governs the evolution of sub-horizon cold dark matter perturbations after radiation perturbations have become negligible.

Now we need to find the two solutions and match the to the logarithmic evolution found above. We can use our prior knowledge about the perturbations deep in the matter era, which we have seen they grow proportionally to a. Therefore, one of the solutions must be a polynomial of y of order 1 (which would imply $\mathrm{d}^2\delta_c/\mathrm{d}y^2=0$). In this case,

$$\frac{\mathrm{d}\delta_c}{\mathrm{d}y}\frac{1}{\delta_c}=\frac{3}{2+3y}\,, \tag{2.32}$$

the solution of which is $\delta_c\propto y+2/3$, or

$$\delta_c\propto a+\frac{2a_{\mathrm{eq}}}{3}\,, \tag{2.33}$$

which approximates to the growth proportional to a for $a\gg a_{\mathrm{eq}}$.

The second solution can be found using $u\equiv\delta_c/(y+2/3)$, which satisfies

$$(1+3y/2)\frac{\mathrm{d}^2u}{\mathrm{d}y^2}+\frac{(21/4)y^2+6y+1}{y(y+1)}\frac{\mathrm{d}u}{\mathrm{d}y}=0\,, \tag{2.34}$$

and involves a first-order equation in the derivative of u. We can therefore integrate to get the solution for this derivative, and then integrate again. The first integral returns

$$\frac{\mathrm{d}u}{\mathrm{d}y}\propto(y+2/3)^{-2}y^{-1}(y+1)^{-1/2}\,, \tag{2.35}$$

and the subsequent integral leads to

$$\delta_c\propto(y+2/3)\log\left[\frac{\sqrt{1+y}+1}{\sqrt{1+y}-1}\right]-2\sqrt{1+y}\,. \tag{2.36}$$

At early times $y\ll 1$, the first solution is constant, and the second, proportional to $\log y$; at late times $y\gg 1$, the first solution scales as y and the second decays as $y^{-3/2}$. Therefore, we can denote them as the growing D_+ and decaying D_- modes, respectively.

Note that the decaying mode cannot be neglected because we need to match the solution to the logarithmic evolution from horizon crossing derived in the previous subsection, which is valid within the horizon before equality. Therefore, we can aspire to get a qualitative solution only for the modes that enter the horizon before equality.

For those modes we can match the two solutions and their first derivatives (with respect to y),

$$\begin{aligned} A\mathcal{R}\log(By_m/y_H) &= C_1 D_+(y_m) + C_2 D_-(y_m)\,, \\ \frac{A\mathcal{R}}{y_m} &= C_1 D'_+(y_m) + C_2 D'_-(y_m)\,, \end{aligned} \tag{2.37}$$

where y_m is the matching time, which must satisfy $y_H \ll y_m \ll 1$, and y_H is the horizon crossing time, which replaces $k\tau$ in the logarithm with y/y_H, valid as long y_m is deep in the radiation era. As you can see, at late times, the only term that matters is D_+, since D_- decays with time.

2.3 Transfer function

We have seen that at linear order each mode $\boldsymbol{k}$ evolves independently from the rest for all species and metric perturbations. Furthermore, while the initial conditions are random given a distribution function, the linear evolution is deterministic. Therefore, we can express any property of a field as function of its initial condition using a transfer function T_X. In the absence of anisotropic stress (e.g., as the one introduced by the massive neutrinos), the transfer function can be decomposed in the time and k dependence $T(a,k) = T(k)D(a)$, where D is known as the linear growth factor, and will be discussed later. Since the effect of neutrinos is not large, and its anisotropic stress is small, there is only a small scale-dependence on the growth factor. However, as we will discuss later, it is key to accurately describe the growth of perturbations.

Let us focus on a specific flavor of the transfer function, regarding the relation between the gravitational potential at a given time with respect to the large-scale beyond-horizon gravitational potential during matter domination (i.e., after accounting for the reduction by the 9/10 factor). Let us also consider that we can completely separate the scale and time dependence on the transfer factor into two different multiplicative factors (the transfer function and the growth factor).[7] Therefore, we can write

$$\Phi(\boldsymbol{k}, a) = \frac{3}{5}\mathcal{R}(\boldsymbol{k})T(k)\frac{D_+(a)}{a}\,, \tag{2.38}$$

where the prefactor accounts for the 9/10 factor of super-horizon scales after matter-radiation equality. The normalization of the growth factor D, although seemingly strange, is like that because it is defined in term of the matter perturbations during matter domination, rather than the gravitational potential.

[7] Massive neutrinos and, in more generality, a non-negligible anisotropic stress introduces a scale dependence in the time evolution of the matter perturbations, which breaks down this assumption.

Using the Poisson equation (in matter domination, well within the horizon), and using the matter density parameter and the definition of the critical density as we have done many times in this chapter, we have (in this limit)

$$\delta_m(\boldsymbol{k},a) = \frac{2k^2 a}{3\Omega_m H_0^2}\Phi(\boldsymbol{k},a) = \frac{2k^2}{5\Omega_m H_0^2}\mathcal{R}(\boldsymbol{k})T(k)D_+(a)\,, \tag{2.39}$$

which by definition implies that the time evolution of matter perturbation is linearly proportional to the growth factor. Note that this definition of the transfer function can be extended to any variable, especially if we define in full generality

$$\delta_x(\boldsymbol{k},a) = \mathcal{R}(\boldsymbol{k})T(k,a)\,. \tag{2.40}$$

As we saw before, modes that enter the horizon after matter-radiation equality have a constant potential. Therefore, the transfer function is very close to unity at scales beyond the size of the horizon at matter-radiation equality, those that fulfill $k \ll k_{\rm eq} = \mathcal{H}_{\rm eq}$. For the consensus cosmology, $k_{\rm eq} = 0.073\,{\rm Mpc}^{-1}\Omega_m h^2 = 0.010\,{\rm Mpc}^{-1}$.

Now, recovering the definition of the power spectrum of primordial curvature perturbations from Eq. (1.126), we find that the linear matter power spectrum is given by

$$P(k,a) = \frac{8\pi}{25}\frac{\mathcal{A}_s}{\Omega_m^2 H_0^4}T^2(k)D_+^2(a)\frac{k^{n_s}}{k_{\rm p}^{n_s-1}}\,. \tag{2.41}$$

The power spectrum is the Fourier transform of the correlation function $\langle\delta(\boldsymbol{x})\delta(\boldsymbol{x}')\rangle$, hence it must have units of volume; we can see in the expression above that this is fulfilled.

To get an analytic expression for the transfer function in this limit, we can recover the results from Eq. (2.37) and get the value for the constant multiplying the growing mode:

$$C_1 = \frac{D_-'(y_m)\log(By_m/y_H) - D_-(y_m)/y_m}{D_+(y_m)D_-'(y_m) - D_+'(y_m)D_-(y_m)}A\mathcal{R}\,. \tag{2.42}$$

The denominator is $-(4/9)y_m^{-1}(y,+1)^{-1/2} = -4/9y_m$, since $y_m \ll 1$. In that limit, $D_- \to (2/3)\log(4/y) - 2$ and $D_-' \to -2/3y$, so that

$$C_1 \to -\frac{9}{4}A\mathcal{R}\left[-\frac{2}{3}\log(By_m/y_H) - (2/3)\log(4/y_m) + 2\right]\,, \tag{2.43}$$

which happens to not depend on y_m. This returns an approximate solution at late times for the small-scale dark matter perturbations in our simplified scenario:

$$\delta_c(\boldsymbol{k},a) = \frac{3}{2}A\mathcal{R}(\boldsymbol{k})\log\left(\frac{4Be^{-3}a_{\rm eq}}{a_H}\right)D_+(a)\,, \qquad (a \gg a_{\rm eq})\,, \tag{2.44}$$

where a_H is the scale factor at which the mode k enters the horizon, $a_H H(a_H) = k$. For very small scales, the argument of the logarithm simplifies, since $a_{\rm eq}/a_H \to \sqrt{2}k/k_{\rm eq}$ (due to the time dependence of the Hubble rate during matter domination). Then, the transfer function (in this limit in which we have ignored baryons and anisotropic stress) is given by

$$T(k) = \frac{15}{4}\frac{\Omega_m H_0^2}{k^2 a_{\rm eq}} A \log\left(\frac{4Be^{-3}\sqrt{2}k}{k_{\rm eq}}\right), \qquad (k \gg k_{\rm eq}). \tag{2.45}$$

Plugging the numbers, we have

$$T(k) = 12\frac{k_{\rm eq}^2}{k^2}\log(0.12k/k_{\rm eq}), \qquad (k \gg k_{\rm eq}). \tag{2.46}$$

This approximation is valid at $k \gtrsim 1\,{\rm Mpc}^{-1}$. There have been derivations with more accurate analytic solutions, but since Boltzmann codes have become so fast and precise, they have lost most of their practical utility by now, beyond providing some qualitative understanding of the evolution of perturbations.

If there had been no logarithmic growth of the matter perturbations during radiation domination, the modes that entered the horizon before equality would have not growth until the epoch of equality, having their amplitude suppressed with respect to large-scale modes by a factor of order $(k_{\rm eq}/k)^2$ (instead having also the logarithmic factor).

We have now the tools to qualitative explain some of the features of the matter power spectrum. In the power spectrum we find a clear turnover scale at $k_{\rm eq}$. Larger scales enter the horizon after equality, hence they have had a constant potential over all their evolution (approximately). This makes that the transfer function at those scales is approximately unity, and the matter power spectrum to be $\propto k$ (accounting for the k^2 relation between δ_c and Φ and the scale dependence of the primordial power spectrum). Smaller scales, however, enter the horizon at earlier times, during the radiation-domination era, and have the potential suppressed. Although this still implies a logarithmic growth for the matter perturbations, they are suppressed by a factor $\sim (k_{\rm eq}/k)^2\log(0.12k/k{\rm eq})$, and therefore the power spectrum decreases with k.

If we keep zero curvature and h fixed, changes in Ω_m change the position and amplitude of the turnover ($k_{\rm eq} \propto \Omega_m h^2$ in physical units, $\propto \Omega_m h$ in Mpc/h units): for lower abundance of matter, equality happens later and $k_{\rm eq}$ is smaller, and viceversa.

2.3.0.1 Effect of baryons and massive neutrinos

After equality, the solution of Eq. (2.37) is not accurate due to the impact of baryons. Baryons contribute after equality to the gravitational potential, but they cluster less than dark matter due to the radiation pressure that they feel until recombination. This solution therefore overestimates the growth of matter perturbations. In a more realistic scenario, baryons suppress matter overdensities in scales below the size of the horizon at equality, given by $k_{\rm eq} \sim 0.01\,{\rm Mpc}^{-1}$ in the fiducial cosmology.

There is another big impact of baryons in the matter perturbations that we have not considered. Before decoupling, the baryon-photon fluid experiences acoustic oscillations (due to the counteracting forces of the radiation pressure and gravity). We saw similar acoustic oscillations in the potential in the radiation-dominated era. Those oscillations reflect the oscillations in the density of the baryon-photon fluid, which are known as baryon acoustic oscillations. The amplitude of the oscillations is small due to the relative abundance of baryons with respect to the total matter.

Massive neutrinos affect the expansion in the Universe (as they become non relativistic), although this does not affect the moment of equality because the non-relativistic transition happens at $z \sim 100$. However, even if non-relativistic, they do free stream, i.e., they are not cold, as dark matter and baryons. Therefore, they travel across perturbations diluting them in scales below the free-streaming scale (determined by the comoving distance a massive neutrino can travel in a Hubble time):

$$k_{\rm fs}(a) \simeq 0.063\,h\,{\rm Mpc}^{-1}\frac{m_\nu}{0.1\,{\rm eV}}\frac{a^2H(a)}{H_0}\,. \tag{2.47}$$

Therefore, the presence of massive neutrino suppresses the power spectrum at $k \gtrsim k_{\rm fs}$, in a scale-dependent time-dependent suppression, since the free-streaming scale (and the level of suppression) depends on time. The suppressing factor with respect to the massless neutrinos case at small scales asymptotes to a constant. More massive neutrinos suppress more than lighter, but at smaller scales (since their free streaming scale is smaller) and viceversa. This means that at large scales, the perturbations for more massive neutrinos may be larger than for lighter neutrinos.

After recombination, free of the radiation pressure, baryons eventually follow the dark matter distribution as they fall in its potential wells, and follow the matter equations in Eq. (2.3). Let us define the relative density perturbation and the relative velocity between baryons and dark matter:

$$\delta_{bc} = \delta_b - \delta_c\,, \qquad v_{bc} = v_b - v_c\,, \qquad \theta_{bc} = \theta_b - \theta_c\,. \tag{2.48}$$

Their evolution equations can be obtained from substracting the evolution equations of each component, yielding

$$\delta'_{bc} + \theta_{bc} = 0\,, \qquad\qquad \theta'_{bc} + \mathcal{H}\theta_{bc} = 0\,. \tag{2.49}$$

There is no impact of the gravitational potential here, because the gravitational potential cares only about the total matter. The solutions for the system above involves a solution with constant relative density perturbations and no relative velocity, and a decaying mode for the total relative velocity $\theta_{bc} \propto a^{-1}$, with $\delta_{bc} \propto \int {\rm d}\tau/a$. The latter corresponds to giving baryons an initial push such as they have a different initial condition than dark matter. This is actually the realistic case, since after recombination, baryons have a different velocity than dark matter as they fall in its potential wells. Nonetheless, this difference in the state after recombination is washed out by the gravitational pull of dark matter by the time we observe the large-scale structure.

It is relevant, nonetheless, for the early time perturbations at very small scales: after recombination, v_{bc} is supersonic, which means that baryons can travel over dark matter potential wells diluting them rather than actually falling in them. This is why for early times it is necessary to study the small-scale limit as function of a bulk relative velocity between the two species. The variance of such bulk velocity is determined by the physics of the photon-baryon plasma before recombination. The main impact is that, at early times, the supersonic bulk relative velocity suppresses the growth of structures at very small scales, with different patches of the Universe showing different levels of suppression that are correlated at large distances following the baryon acoustic oscillations pattern.

2.3.1 Growth factor

We can also discuss the time evolution of the matter perturbations, in terms of a scale-independent linear growth factor. At late times, the horizon is much larger than the scales of interest, and the only deviation that we find from the Meszaros equation is the influence of dark energy. Furthermore, at these times (after recombination), baryons do not feel any pressure and therefore behave like cold dark matter (except for the acoustic term that matters at very small scales). While dark matter and baryons start with different initial conditions after recombination, baryons fall in the dark matter potential wells and trace the dark matter perturbations faithfully. Therefore, we will use the total matter perturbations (with a energy-density weighted average).

We start from the matter equations in Eq. (2.3), multiplying the first one by a a deriving with respect to the conformal time. Neglecting the second derivative of Φ, since it is negligible within the horizon, we have

$$(a\delta_m')' = ak^2\Phi\,, \tag{2.50}$$

which we can combine with the Einstein equation of Eq. (2.4). Neglecting contributions from radiation and terms that are small when $k \gg \mathcal{H}$ and using the Friedmann equation and the density parameter, we have

$$(a\delta_m')' = \frac{3}{2}\Omega_m H_0^2 \delta_m\,. \tag{2.51}$$

To solve this equation it is better to use a as the time variable, which returns

$$\frac{\mathrm{d}^2\delta_m}{\mathrm{d}a^2} + \frac{\mathrm{d}\log(a^3H)}{\mathrm{d}a}\frac{\mathrm{d}\delta_m}{\mathrm{d}a} - \frac{3\Omega_m H_0^2}{2a^5H^2}\delta_m = 0\,, \tag{2.52}$$

which has to be solved numerically. We can use the variable $u = \delta_m H^{-1}$, that leaves the equation

$$\frac{\mathrm{d}^2 u}{\mathrm{d}a^2} + 3\left[\frac{\mathrm{d}\log H}{\mathrm{d}a} + \frac{1}{a}\right]\frac{\mathrm{d}u}{\mathrm{d}a} = 0\,. \tag{2.53}$$

The first order equation can be integrated to obtain $\mathrm{d}u/\mathrm{d}a \propto (aH)^{-3}$. If we integrate again, and remembering that the growth factor is uH, we have

$$D_+(a) \propto H(a)\int^a \frac{\mathrm{d}a'}{(a'H(a'))^3}\,. \tag{2.54}$$

Now we need to find the normalization. We can find it matching the behavior of the definition of the growth factor $D_+(a) = a$ during matter domination. Therefore, since at those times $H = H_0\sqrt{\Omega_m a^{-3}}$,

$$D_+(a) = \frac{5\Omega_m}{2}\frac{H(a)}{H_0}\int_0^a \frac{\mathrm{d}a'}{(a'H(a')/H_0)^3}\,. \tag{2.55}$$

This is only valid for matter and a cosmological constant components.

We can find the solution for the decaying mode assuming $\delta_m = H$, expressing the equation in terms of H^2 and substituting H^2 for the sum of the components and their evolution. In this case, we will find the condition that for $\delta_m = H$ to work as a solution, any component beyond matter must fulfill $p_s^2 + 2p_s = 0$, which is the same condition for the growing mode.

Finally, a relevant quantity for large-scale structure is the logarithmic derivative of the growth factor, known as the growth rate f, defined as

$$f(a) \equiv \frac{\mathrm{d}\log D_+}{\mathrm{d}\log a} \simeq (\Omega_m(a))^{0.55}\,, \tag{2.56}$$

where the last equality involves a fitting function which depends on the time-dependent matter density parameter. The growth rate reduces to $f = 1$ in the totally matter dominated Universe (i.e., $\Omega_m = 1$), and it is only when dark energy becomes relevant that the growth factor over the scale factor (D/a) (and also f) start to decay.

Before closing this chapter, let us note that there is a slightly different convention regarding the normalization of the growth factor. It can be defined in terms of early-time perturbations, as we have done so far. However, for studies of large-scale structure, it is more common to find it defined in terms of the matter power spectrum in the present day. In that case, the growth factor would fulfill

$$P_m^2(k,a) = D_{\mathrm{LSS}}^2 P_m^2(k,a_0)\,. \tag{2.57}$$

Of course, these two conventions only differ in their normalization.

2.4 Limit of linear theory

We have limited the discussion to linear perturbations so far. Non linearities, which significantly complicate the study of the growth of perturbations and large-scale structure, are bound to be relevant at small scales. There are different ways to estimate the scales at which non linearities cannot be ignored. One of them is to compute the variance of linear perturbations in a certain spatial scale. For instance, consider an spherical top-hat region in Fourier space (which corresponds to a sinc window function in configuration space, and viceversa), and the variance will be given by

$$\sigma_{\rm w}^2 = \frac{1}{2\pi^2}\int {\rm d}k k^2 W^2(k) P(k)\,, \tag{2.58}$$

where $P(k)$ is the linear power spectrum and $W(k)$ is the spherically symmetric window function in Fourier space of the region we are considering. If $\sigma^2 \gtrsim 1$, the perturbations are too large for the linear regime to accurately describe them, and non-linear growth is relevant for the study. We can therefore scan $\sigma_{\rm w}^2$ as function of radius (or scale) to find the scale $k_{\rm NL}$ at which non-linear perturbations become relevant.

Another way to estimate $k_{\rm NL}$ is to consider the variance of modes within a specific narrow logarithmic wavenumber:

$$\sigma_{\rm L}^2 = \frac{1}{\epsilon}\int_{|\log k' - \log k| < \epsilon} \frac{{\rm d}\Omega_{\rm k}{\rm d}\log k k^3}{(2\pi)^3} P(k) = \frac{k^3}{2\pi^2} P(k)\,, \tag{2.59}$$

where for the last equality we have assumed an infinitesimal wavenumber bin. Similarly, linear perturbations fulfill $\sigma_{\rm L}^2 \ll 1$, while values close tu unity indicate non-linear perturbations. Today, this corresponds to a non-linear scale of $k_{\rm NL}(a=1) \simeq 0.25\, h\,{\rm Mpc}^{-1}$, and progressively higher values as we go higher in redshift (since structure did not have time to grow so much).

CHAPTER 3

LECTURE 2: CMB PRIMARY ANISOTROPIES

In the previous chapter we studied, under some simplifying assumptions and in specific limits, how the gravitational potential evolved and how this impacted the dark matter perturbations. The matter distribution in the Universe is relevant because it is the one that determines the potential wells in which galaxies will form, and make up for the large-scale structure we observe in the Universe today. The late-times large-scale structure is probed with galaxy surveys, especially through galaxy clustering and weak lensing.[1]

However, we did not pay much attention to the photon perturbations. Given how precise the observations of the cosmic microwave background (CMB) anisotropies are, understanding photon perturbations and predicting them accurately is crucial to understand our Universe and constrain cosmological models. Furthermore, most of the cosmic tensions appear between low-redshift probes and the CMB measurements made by Planck (6). There-

[1]There are other approaches and techniques to probe the large-scale structure, including CMB secondary anisotropies like CMB lensing (4), and line-intensity mapping (5). All these approaches are complementary in their strengths and weaknesses and in the scales and times that can probe.

fore, we will dedicate some time to understand the different features in the CMB power spectrum.

As expected, the photon perturbations behave drastically different before and after recombination, which takes places around $z_* \sim 1100$.[2] Before recombination, the interactions between photons and free electrons are so frequent that photons and baryons are tightly coupled and can be described as a single fluid; after recombination, in turn, photons free stream from the last-scattering surface. Since gravitational potentials are too weak to trap photons, photon overdensities do not grow after recombination, contrary to dark matter and baryons.

As discussed earlier, we can describe photons in terms of the perturbations in their phase-space distribution. In Eq. (1.62) we defined the perturbations in the phase-space distribution as $f(x^i, P_j, \tau) = f_0(q, m)(1 + \varphi(x^i, q, \hat{\boldsymbol{q}}_j, \tau))$, which in the case of the photons can be further simplified taking the momentum-averaged perturbation $\mathcal{F}_\gamma$. In Fourier space, the momentum-averaged perturbation of the phase space distribution only depends on the Fourier mode, direction of the momentum, and conformal time, and it can be related to the photon overdensity, velocity divergence and anisotropic stress following Eq. (1.92).

Nonetheless, we cannot measure directly those photon properties. In turn, we can measure the intensity of the radiation that arrives along a given line of sight as function of frequency (and the polarization of that radiation). Therefore, rather than dealing with the photon properties, it is more convenient to work with the temperature T that determines its background phase-space distribution

$$f_0 = f_0(\epsilon) = \frac{g_*}{h_P^3} \frac{1}{\exp\{\epsilon / k_\mathrm{B} T_0\} \pm 1}\,, \tag{3.1}$$

where as in the derivation of the Boltzmann equations we use $\epsilon = aE = a\sqrt{p^2 + m^2} = \sqrt{P^2 + a^2 m^2}$ and $T_0 = aT$ as the temperature of the particles today. At linear order, perturbations in the photon distribution maintain the black-body spectrum, but change the associated temperature of the distribution. Hence, we can equally describe the perturbations in the phase space distribution with perturbations in the temperature:

$$T = \bar{T}(1 + \Theta) \implies \Theta = \frac{T - \bar{T}}{\bar{T}}\,. \tag{3.2}$$

Therefore, if we substitute this expression for the temperature in f_0, we find that $f = f_0(q/(1 + \Theta))$ in such a way that at linear order[3]

$$\Theta = -\left(\frac{\mathrm{d}\log f_0}{\mathrm{d}\log q}\right)^{-1} \varphi = \frac{1}{4}\mathcal{F}_\gamma\,. \tag{3.3}$$

[2]We will denote quantities related with recombination with a subscript ‘*’.
[3]Remember that since photons are massless, $\epsilon = q$, where $q = ap$ is the comoving momenta used in the chapter about cosmological perturbation evolution.

We cannot observe the actual state of the photons in the last-scattering surface, but how they reach us after traveling through the Universe. For instance, photons have to exit the potential they were at the last-scattering surface, which changes their energy accordingly to the sign of the potential: they lose energy if they were in an overdensity ($\Psi < 0$), and viceversa, due to the gravitational redshift. Therefore, the actual observed temperature is $\Theta_0 + \Psi_*$. Furthermore, we can only measure their properties as function of position on the sky, not in terms of any radial distance. This is why we will focus on angular summary statistics to describe the angular maps obtained from the CMB observations. In particular, we will focus on the angular power spectrum. To do that, let us define the temperature perturbation as function of a three-dimensional position,

$$\begin{aligned}\Theta(\boldsymbol{x},\hat{\boldsymbol{q}},\tau) &= \int \mathrm{d}^3k e^{i\boldsymbol{k}\boldsymbol{x}}\Theta(\boldsymbol{k},\hat{\boldsymbol{q}},\tau) = \\ &= \int \mathrm{d}^3k e^{i\boldsymbol{k}\boldsymbol{x}}\sum_{\ell=0}^{\infty}(-i)^{\ell}(2\ell+1)\Theta_\ell(\boldsymbol{k},\tau)\mathcal{P}_\ell(\mu)\,,\end{aligned}\tag{3.4}$$

where $\mu = \hat{\boldsymbol{k}}\hat{\boldsymbol{q}}$ is the cosine of the angle between $\boldsymbol{k}$ and the propagation direction of the photon. Note that the direction of the momentum of the photon must be the same as the direction in which an observer at the origin (us) looks at the sky to detect them (with a different sign). Therefore, let us change $\hat{\boldsymbol{q}}$ for the angle on the sky $\hat{\boldsymbol{n}}$. The anisotropy at the origin as function of position on the sky is

$$\begin{aligned}\Theta(\hat{\boldsymbol{n}}) &= \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell m}Y_{\ell m}(\hat{\boldsymbol{n}})\,,\\ a_{\ell m}(\boldsymbol{x}) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}e^{i\boldsymbol{k}\boldsymbol{x}}\int \mathrm{d}\Omega_{\mathrm{n}}Y^*_{\ell m}(\hat{\boldsymbol{q}})\Theta(\boldsymbol{k},\hat{\boldsymbol{q}},\tau)\,,\end{aligned}\tag{3.5}$$

where we have used the orthonormality of the spherical harmonics.

We cannot make any predicction about specific values of the perturbations in a specific point (or a specific coefficient $a_{\ell m}$ in this case); we can only predict their ensemble variance (the average is null by definition), which is measured in practice using the Ergodic hypothesis. The covariance of the expansion coefficients $a_{\ell m}$ is given by the angular power spectrum:

$$\langle a_{\ell m}a^*_{\ell' m'}\rangle = C_\ell\delta_{\ell\ell'}\delta_{mm'}\,.\tag{3.6}$$

Excluding the monopole and dipole (i.e., for $\ell \geq 2$), the power spectrum and correlation function are gauge-independent quantities.

Before recombination, there are two opposing forces influencing the photon-baryon fluid. On the one hand, there is gravity, for which the potential wells in the dark matter overdensity pull the fluid in. On the other, radiation pressure between photons and baryons grows with density, diluting the photon-baryon overdensities and therefore pushing against gravity. This situation is analog to a forced harmonic oscillator

$$\ddot{x} + \frac{K}{m}x = F\,, \tag{3.7}$$

where the driving force F is due to gravity. The total force is $mF - Kx$, where x is the position of the oscillator and K is the force constant of the oscillator. The general solution for this system has two oscillatory modes with an angular frequency $w = \sqrt{K/m}$, and a particular solution is $x = F/w^2$. Assuming that the oscillator is initially at rest the sine mode vanishes, which leaves

$$x = A\cos(wt) + \frac{F}{w^2}\,. \tag{3.8}$$

The driving forces displaces the unforced situation from zero, so that the two extreme points at each side of the oscillations are not symmetric. The shift is more dramatic for smaller frequencies. The square of the oscillator position shows that the odd and even peaks have different heights due to this shift. Therefore, a forced harmonic oscillator is determined by the external force F and the reduced spring constant K/m.

In our case, for the photon-baryon fluid, the frequency grows as we decrease the effective mass of the fluid, i.e., as we decrease Ω_b. The fewer the baryon abundance, the higher the sound speed of the fluid (and closer the peaks of the wave pattern). In turn, with more cold dark matter, the gravitational potentials are larger, which increases the driving force (and lowers the frequency), and therefore the difference in the amplitude between odd and even peaks is larger. As the fluid falls in the potential, radiation pressure increases and pushes the plasma outwards to maximum expansion, leading to an underdensity with smaller amplitude than in the absence of gravity. Then the radiation pressure reduces and the plasma clusters again, and the cycle repeats from the beginning.

On a different note, remember that even if photon and baryons are tightly coupled, their interaction rate is not infinite. This allows the photons to travel a finite distance between two scatter events. The mean free path $\lambda_{\rm MFP}$ in this case is the inverse of the derivative of the optical depth, $\lambda_{\rm MFP} = (n_e\sigma_{\rm T}a)^{-1}$ from the collision term for photons. Over a Hubble time, photons undergo $\sim n_e\sigma_{\rm T}H^{-1}$ scatter events (transforming the scattering rate to time instead of conformal time, and multiplying for the time). For a random walk like this, the total distance traveled is the mean free path times the square root of the number of steps (i.e., scatter events). Therefore, a cosmological photon moves a mean comoving distance

$$\lambda_{\rm D} \sim \lambda_{\rm MFP}\sqrt{n_e\sigma_{\rm T}H^{-1}} = \left(a\sqrt{n_e\sigma_{\rm T}H}\right)^{-1} \tag{3.9}$$

over a Hubble time. Any perturbation on scales smaller than this distance will be washed out due to all the photons diffusing over a patch of this scale, which homogenizes the photon temperature. In Fourier space, this smoothing corresponds to a damping of high-k modes. Since $\lambda_{\rm D}$ depends on the number of electrons, the diffusion scale depends on Ω_b. Larger Ω_b reduces $\lambda_{\rm D}$ which in turn reduces the damping.

We have qualitatively described what is known as the primary CMB anisotropies. However, photons do not travel completely unaffected after recombination. Instead, they are affected by evolving gravitational potential (integrated Sachs-Wolfe effect), reionization, gravitational lensing due to metric perturbations along the line of sight, and interactions with free electrons (Sunyaev-Zeldovic effect).

We will provide a more accurate qualitative understanding of the photon perturbations to understand the CMB power spectrum and how we can use it to constrain cosmological parameters. As with the case of dark matter perturbations, an almost exact treatment requires the use of numerical Boltzmann code. On what follows, we will distinguish between different regimes and stages of evolution to simplify the computations.

We will focus primarily on the CMB temperature anisotropies. However, as discussed in the previous chapters, Compton scattering generates linear polarization (in turn, cosmological perturbations do not generate circular polarization). Nonetheless, only the quadrupole of the photon perturbations generate non-zero polarization. We can also distinguish between a curl-free, scalar component of the polarization (known as E mode) and a divergence-free, pseudoscalar component (known as B mode); scalar perturbations only generate E modes, while the B modes are generated either by primordial tensor perturbations or through secondary anisotropies like lensing.

Since only the quadrupole generates polarization, $E(k) \propto \Theta_2(k)$ (in the tight-coupling approximation), and actually the monopole and quadrupole of the polarization perturbations contribute to the temperature perturbation through this same connection. Finally, note that the polarization perturbations must be significantly smaller than the temperature perturbations, since the quadrupole is suppressed in the early Universe due to Compton scattering.

3.1 Large-scale anisotropies

The large-scale limit can be treated with the same system that was discussed in Eq. (2.5). From the equation for the photon monopole, $\Theta_0' = -\Phi'$, we find that $\Theta_0 = -\Phi$ plus a constant. Similarly, from Eq. (1.134) we learn that the initial post-inflation condition is $\Theta_0 = \Phi/2$, so the constant must be $\mathcal{R} = 3\Phi_{\rm super\,hor.}/2$ (from Eq. (2.14)). The large-scale evolution of Φ is given by Eq, (2.13), but note that recombination takes places long after equality, hence $\Phi = 3\mathcal{R}/5$ in this limit. Therefore,

$$\Theta_0(\boldsymbol{k}, \tau_*) = -\Phi(\boldsymbol{k}, \tau_*) + \mathcal{R}(\boldsymbol{k}) = \frac{2}{5}\mathcal{R}(\boldsymbol{k}) = \frac{2}{3}\Phi(\boldsymbol{k}, \tau_*)\,. \tag{3.10}$$

As discussed before, the observed anisotropy is $\Theta_0 + \Psi$ (and using that $\Psi \simeq -\Phi$), so that we have

$$(\Theta_0 + \Psi)(\boldsymbol{k}, \tau_*) = -\frac{1}{5}\mathcal{R}(\boldsymbol{k}) = -\frac{1}{3}\Phi(\boldsymbol{k}, \tau_*)\,. \tag{3.11}$$

From the last two equations we see something that may be counter intuitive. On the one hand, photons are hotter ($\Theta_0 > 0$) in places where gravity is more intense ($\Phi > 0$, $\Psi < 0$). However, we do not see them actually hotter, because the energy they lose as they climb those potential wells makes them actually cooler than those coming from places where gravity is less intense. This also applies for matter over and underdensities: if we integrate the equation for δ_c in Eq. (2.3) and apply the initial condition $\delta_c = \mathcal{R}$ from the inflation chapter, we find

$$\delta_c(\boldsymbol{k}, \tau_*) = \mathcal{R}(\boldsymbol{k}) - 3\left[\Phi(\boldsymbol{k}, \tau_*) - \frac{2}{3}\mathcal{R}(\boldsymbol{k})\right] = \frac{6}{5}\mathcal{R}(\boldsymbol{k}) = 2\Phi(\boldsymbol{k}, \tau_*)\,, \tag{3.12}$$

so that the observed anisotropy in terms of the dark matter overdensity is

$$(\Theta_0 + \Psi)(\boldsymbol{k}, \tau_*) = -\frac{1}{6}\delta_c(\boldsymbol{k}, \tau_*)\,, \tag{3.13}$$

presenting a similar behavior than with respecto to the gravitational potentials. Therefore, hotter observed anisotropies corresponds to underdense regions.

3.2 Baryon acoustic oscillations

The mean-free path of photons before recombination is significantly smaller than the size of the horizon, which couples them to baryons conforming a tightly-coupled photon-baryon fluid. This condition applies when the optical depth is $\gg 1$ (i.e., $\int n_e \sigma_{\rm T} a \gg 1$). As argued before, the competing forces of radiation pressure and gravity build acoustic oscillations in the fluid.

In this limit, all moments beyond the monopole and dipole are suppressed: the photons therefore behave like a fluid and can be described by its density and velocity. We can show this starting from Eqs. (1.104), and taking the limit in which $\lambda_{\rm MFP} = (n_e \sigma_{\rm T} a)^{-1}$ is very small. For the cases in which $\ell \geq 3$, $\Theta'_\ell \sim \Theta_\ell/\tau \ll n_e \sigma_{\rm T} a \Theta_\ell$, and neglecting the coupling to the higher multipole, we have

$$\Theta_\ell \sim \frac{k}{n_e \sigma_{\rm T} a}\frac{\ell}{2\ell+1}\Theta_{\ell-1} = k\lambda_{\rm MFP}\frac{\ell}{2\ell+1}\Theta_{\ell-1}\,. \tag{3.14}$$

Therefore, for scales much larger than the mean-free path, $\Theta_\ell \ll \Theta_{\ell-1}$ (which justifies neglecting of the higher multipole above). If we neglect the contribution from the difference between the two linear polarization components given by $\mathcal{G}_\ell$, we can also neglect Θ_2.

Physically, we can understand this as follows. Consider a plane-wave perturbation: an observer at its center sees photons coming from a distance $\sim \lambda_{\rm MFP}$. Therefore, large-scale perturbations (i.e., $k\lambda_{\rm MFP} \ll 1$) do not contribute to the perturbations that the observer perceives, because they produce a constant temperature over that volume. Small scales perturbations are in turn damped by the diffusion of the photons. Therefore, considering only the first two moments:

$$\begin{aligned} \Theta_0' + k\Theta_1 &= -\Phi' \,, \\ \Theta_1' - \frac{k\Theta_0}{3} &= \frac{k\Psi}{3} - a n_e \sigma_{\rm T} \left(\Theta_1 - \frac{\theta_b}{3k} \right) \,, \end{aligned} \tag{3.15}$$

which are accompanied by the baryon equations, which we can rewrite, defining $R \equiv 3\bar{\rho}_b/4\bar{\rho}_\gamma$ and ignoring the acoustic term, as

$$\theta_b = 3k\Theta_1 - \frac{R}{n_e \sigma_{\rm T} a} \left(\mathcal{H}\theta_b - k^2\Psi + \theta_b' \right) \,. \tag{3.16}$$

The second term is much smaller due to the $R\lambda_{\rm MFP}$ factor (multiplied by $1/\tau$ and k in each case of the terms in the parenthesis). To lowest order we take $\theta_b = 3k\Theta_1$, and expand substituting this lowest-order expression in the second term, leading to

$$\theta_b \simeq 3k\Theta_1 - \frac{R}{n_e \sigma_{\rm T} a} \left(3k\mathcal{H}\Theta_1 - k^2\Psi + 3k\Theta_1' \right) \,, \tag{3.17}$$

which we can use to eliminate θ_b in the photon perturbation equations above. After rearranging a bit the terms:

$$\Theta_1' + \frac{\mathcal{H}R}{1+R}\Theta_1 - \frac{k}{3(1+R)}\Theta_0 = \frac{k}{3}\Psi \,. \tag{3.18}$$

Now we have a system of two first-order equations; as done in the previous chapter we will differentiate the equation for Θ_0, substitute the equation above, and then use the equation for Θ_0 without differentiate to substitute Θ_1, to obtain

$$\Theta_0'' + \frac{\mathcal{H}R}{1+R}\Theta_0' + k^2 c_s^2 \Theta_0 = F(k,\tau) \,, \tag{3.19}$$

where we have defined the force function

$$F(k,\tau) \equiv -\frac{k^2}{3}\Psi - \frac{\mathcal{H}R}{1+R}\Phi' - \Phi'' \,, \tag{3.20}$$

and the sound speed of the fluid as

$$c_s(\tau) \equiv \sqrt{\frac{1}{3(1+R(\tau))}} \,. \tag{3.21}$$

Note that the sound speed depends on Ω_b. If the abundance of baryons is negligible, the sound speed tends to $1/\sqrt{3}$, as for any relativistic fluid. Baryons makes the fluid *heavier*, which acts as the inverse mass in the reduced spring constant of the forced harmonic oscillator. Actually, the equation above for Θ_0 is a forced, damped harmonic oscillator. Most of the terms multiplying Φ coincide with those of Θ_0 therefore we can rewrite the equation above as

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + \frac{\mathcal{H}R}{1+R}\frac{\mathrm{d}}{\mathrm{d}\tau} + k^2 c_s^2\right\}\left[\Theta_0 + \Phi\right](\boldsymbol{k},\tau) = \frac{k^2}{3}\left[\frac{1}{1+R}\Phi - \Psi\right](\boldsymbol{k},\tau)\,. \tag{3.22}$$

We will use again the Green's method to solve the full solution, which proposes to find the particular solution starting from the two homogeneous general solutions. The drag term in the equation above goes as $R(\Theta_0 + \Phi)/\tau^2$ and the pressure ($\propto k^2 c_s^2$) is much larger for modes within the horizon or if R is small, which describes how for the scales of interest the impact of the pressure (oscillations, in this case) is much more significant than the one from the Hubble expansion. Although there is a solution including this term (the WKB solution, which assumes a solution of the $\Theta_0 = Ae^{iB}$), let us neglect the drag term, for which we have the oscillatory homogeneous solutions

$$S_1 = \sin(kr_s(\tau))\,; \qquad\qquad S_2 = \cos(kr_s(\tau))\,, \tag{3.23}$$

where the sound horizon is the comoving distance that the acoustic wave has had time to travel in time τ:

$$r_s = \int_0^\tau \mathrm{d}\tilde{\tau} c_s(\tilde{\tau})\,. \tag{3.24}$$

The total solution (including the particular solution for the driving force) can be obtained from these two solutions similarly than for Eq. (2.26) (and neglecting all instances of R outside the oscillatory homogeneous solutions):

$$\begin{aligned}\Theta_0 + \Phi =& C_1 S_1 + C_2 S_2 + \\ &+ \frac{k^2}{3}\int_0^\tau \mathrm{d}\tilde{\tau}(\Phi(\tilde{\tau}) - \Psi(\tilde{\tau}))\frac{S_1(\tilde{\tau})S_2(\tau) - S_1(\tau)S_2(\tilde{\tau})}{S_1(\tilde{\tau})S_2'(\tilde{\tau}) - S_1'(\tilde{\tau})S_2(\tilde{\tau})},.\end{aligned} \tag{3.25}$$

We can fix the integration constants to the initial condition for which both Θ_0 and Φ are constants. Therefore, the coefficient C_1 multiplying the sine must be zero, and $C_2(\boldsymbol{k}) = \Theta_0(\boldsymbol{k},0) + \Phi(\boldsymbol{k},0)$. In our limit, R is effectively very small, hence the denominator in the integral, which is $-kc_s$ can be approximated as $-k\sqrt{3}$. and the numerator can be reexpressed as the sine of the difference of the arguments, so that

$$\begin{aligned}\Theta_0 + \Phi =& (\Theta_0(0) + \Phi(0))\cos(kr_s) + \\ &+ \frac{k}{\sqrt{3}}\int_0^\tau \mathrm{d}\tilde{\tau}(\Phi(\tilde{\tau}) - \Psi(\tilde{\tau}))\sin\left[k(r_s(\tau) - r_s(\tilde{\tau})\right]\,.\end{aligned} \tag{3.26}$$

Since outside the horizon $\Theta_0 + \Phi$ is constant, only the cosine mode is excited and a clear oscillatory pattern can be appreciated in the solution. This expression can predict with accuracy the position of the acoustic peaks from a numerical solution. To get the solution we should numerically integrate the last term above, but we can simplify a bit a further. If the first term dominates, the position of the peaks is given by the extrema of $\cos(kr_s)$: $k_{\rm pk} = n\pi/r_s$, where n is a natural number, which is within 10% of the numerical solution.

Finally we can use Eq. (3.15) to relate this solution to the dipole of the photon distribution:

$$\begin{aligned} \Theta_1(\boldsymbol{k}, \tau) = & \frac{1}{\sqrt{3}}(\Theta_0(0) + \Phi(0)) \sin(kr_s) - \\ & -\frac{k}{3}\int_0^\tau \mathrm{d}\tilde{\tau}(\Phi(\tilde{\tau}) - \Psi(\tilde{\tau})) \cos\left[k(r_s(\tau) - r_s(\tilde{\tau})\right] , \end{aligned} \tag{3.27}$$

which is completely out of phase with respect to the monopole, even after accounting for the integral term.

3.3 Diffusion damping

Diffusion is characterized by a small but non-negligible quadrupole moment. Therefore, we need to recover Eq. (1.104) to account for it to obtain the equivalent of Eq. (3.15). However, we can simplify on other end: diffusion matters at very small scales, where gravitational potentials are smaller than radiation perturbations by a factor $\mathcal{H}/k^2$. Otherwise, all the considerations made in the previous section still apply, so that we can neglect all moments above the quadrupole and we have (after neglecting the effects of polarization)

$$\begin{aligned} & \Theta_0' + k\Theta_1 = 0 , \\ & \Theta_1' + \frac{k}{3}(2\Theta_2 - \Theta_0) = n_e\sigma_{\rm T}a\left(\frac{\theta_b}{3k} - \Theta_1\right) , \\ & \Theta_2' - \frac{2k}{5}\Theta_1 = -n_e\sigma_{\rm T}a\frac{9}{10}\Theta_2 , \end{aligned} \tag{3.28}$$

along with

$$3k\Theta_1 - \theta_b = \frac{R}{n_e\sigma_{\rm T}a}\left(\mathcal{H}\theta_b + \theta_b'\right) , \tag{3.29}$$

which is a small rephrase of Eq. (3.16) after dropping the potentials. We know that the time dependence of the variables involved is gonna follow sinusoidal functions, hence let us assume that already, but using the exponential form, such as $\theta \propto e^{i\int \mathrm{d}\tilde{\tau}\omega}$, where we know that $\omega \simeq kc_s$ in the tight-coupled limit. This implies that the derivative with respect to conformal time is

$$|\theta_b'| = |i\omega\theta_b| \gg \mathcal{H}|\theta_b| , \tag{3.30}$$

where we have used the approximate value of ω and that $k \gg \mathcal{H}$ at small scales. Thus, we can drop the $\mathcal{H}\theta_b$ term in the baryon equation above. Substituting the relation between θ_b and θ_b' in the equation above and expanding the denominator up to second order, we have

$$\theta_b = 3k\Theta_1 \left[1 - \frac{i\omega R}{n_e \sigma_{\rm T} a} - \left(\frac{i\omega R}{n_e \sigma_{\rm T} a}\right)^2\right] . \tag{3.31}$$

We can do the same procedure for the quadrupole. First, since we are in a regime where the mean-free path is very small, hence we can drop the Θ_2' term, which leaves

$$\Theta_2 = \frac{4k}{9 n_e \sigma_{\rm T} a} \Theta_1 \, , \tag{3.32}$$

which shows that our hierarchy closing scheme is sound: higher moments are suppressed by a $k\lambda_{\rm MFP}$ factor. Finally, the equation for the monopole is given by

$$i\omega\Theta_0 = -k\Theta_1 \, . \tag{3.33}$$

We can now insert all these expressions in the equation for the dipole, which returns the dispersion relation for ω (after collecting all the terms):

$$\omega^2(1+R) - \frac{k^2}{3} - \frac{i\omega}{n_e \sigma_{\rm T} a}\left[\omega^2 R^2 + \frac{8k^2}{27}\right] = 0 \, . \tag{3.34}$$

Note that the last term is suppressed by a mean-free path factor. If we were to neglect that term, we would recover the result of the previous section: that the frequency is kc_s.[4] Since the last term is a correction, we can write the frequency of the oscillator as the previous result plus a minor correction

$$\delta\omega = \frac{ik^2}{2(1+R) n_e \sigma_{\rm T} a}\left[c_s^2 R^2 + \frac{8}{27}\right] . \tag{3.35}$$

Therefore, the time dependence for the perturbations is given by

$$\sim \exp\left\{ik \int \mathrm{d}\tilde{\tau}\, c_s(\tilde{\tau})\right\} \exp\left\{-\frac{k^2}{k_{\rm D}^2}\right\} , \tag{3.36}$$

where we have defined the damping scale and its corresponding wavenumber as

$$k_{\rm D}^{-2} \equiv \int_0^\tau \frac{\mathrm{d}\tilde{\tau}}{6(1+R) n_e \sigma_{\rm T} a(\tilde{\tau})}\left[\frac{R^2}{1+R} + \frac{8}{9}\right] . \tag{3.37}$$

For an order-of-magnitude qualitative understanding, the above expression implies

$$\lambda_{\rm D} \sim k_{\rm D}^{-1} \sim \sqrt{\tau \lambda_{\rm MFP}} \, , \tag{3.38}$$

[4]Note that in this case we do not have any forcing term in the harmonic oscillator because we have neglected the contribution from the gravitational potential.

which matches our previous expectations (remember that $\tau \simeq \mathcal{H}^{-1}$). The diffusion scale grows with $\sim a^{1/2}$ and $\Omega_b^{-1/2}$, and damps the power spectrum at multipoles $\ell \gtrsim k_{\mathrm{D}}\tau_0 \sim 10^3$. This effect is known as the Silk damping.

3.4 Projection to anisotropies on the sky

Until now we have derived the three-dimensional perturbations in the photon-baryon fluid at recombination, but actually we are only sensitive to the projected anisotropies on the sky, once photons arrived to us. Remember that the moments were defined in terms of the angle between the direction of the propagation of the photon and $\boldsymbol{k}$, and that the direction of propagation is set by the fact that they arrive to us through a given line of sight. Therefore, we need a solution for the photon moments today in terms of the monopole and dipole at recombination.

We can use Eqs. (1.77) and (1.101), and rearrange a bit the terms to get

$$\Theta' + (ik\mu - \uptau')\Theta = \hat{S}\,, \tag{3.39}$$

where we have defined the scattering optical depth integrated backwards from today[5]

$$\uptau \equiv \int_{\tau_0}^{\tau} \mathrm{d}\tilde{\tau} n_e \sigma_{\mathrm{T}} a(\tilde{\tau})\,, \qquad\qquad \uptau' = -n_e\sigma_{\mathrm{T}} a\,, \tag{3.40}$$

and the source function

$$\hat{S} \equiv -\Phi' - ik\mu\Psi - \uptau'\left[\Theta_0 - \frac{i\theta_b}{k}\mathcal{P}_1(\mu) - \frac{1}{2}\Pi\mathcal{P}_2(\mu)\right]\,, \tag{3.41}$$

in turn using

$$\Pi \equiv \frac{1}{4}\left(\mathcal{F}_{\gamma 2} + \mathcal{G}_{\gamma 0} + \mathcal{G}_{\gamma 2}\right)\,. \tag{3.42}$$

As a side note, it is now convention to set the *moment* of recombination τ_* as the conformal time for which $\uptau = 1$, although there are also alternative conventions. We can turn the differential equation above into an integral equation. Rewrite the left-hand side of Eq. (3.39) as a factor multiplying a time derivative so that

$$\Theta' + (ik\mu - \uptau')\Theta = \Theta' + \mathcal{A}\Theta = e^{-\mathcal{A}}\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\Theta e^{\mathcal{A}}\right]\,. \tag{3.43}$$

[5]Here we face a slight conflict regarding the notation. The optical depth is usually denoted by a regular τ. Here we decide to used the variant $\uptau$ to avoid confusion with the conformal time. Other sources, especially those that do not use the synchronous gauge, solve this conflict denoting the conformal time with η. On the other hand, there are references using κ to denote the optical depth; we prefer not to use that convention to avoid confusion with the curvature.

Therefore, we can write $(\Theta e^{ik\mu\tau - \mathrm{\tau}})' = e^{ik\mu\tau-\mathrm{\tau}}\hat{S}$ and integrate over conformal time to obtain

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\Theta e^{\mathcal{A}}\right] = e^{\mathcal{A}}\hat{S} \implies \\ &\Theta(\tau_0) = \Theta(\tau_{\mathrm{init}})e^{ik\mu(\tau_{\mathrm{init}}-\tau_0)}e^{-\mathrm{\tau}(\tau_{\mathrm{init}})} + \int_{\tau_{\mathrm{init}}}^{\tau_0} \mathrm{d}\tau \hat{S}(\tau)e^{ik\mu(\tau-\tau_0)}e^{-\mathrm{\tau}}\,,\end{aligned} \tag{3.44}$$

where we have used that $\mathrm{\tau}(\tau_0) = 0$ by definition. On the other hand, $\mathrm{\tau}(\tau_{\mathrm{init}})$ blows up for early enough times, so that the exponential vanishes and we can drop the first term. Conceptually, this corresponds to the fact that Compton scattering erases effectively any initial anisotropy. For the same reason, we can move τ_{init} to 0 without any impact. Therefore, the solution for anisotropies is given by

$$\Theta(k,\mu,\tau_0) = \int_0^{\tau_0} \mathrm{d}\tau \hat{S}(k,\mu,\tau)e^{ik\mu(\tau-\tau_0)}e^{-\mathrm{\tau}}\,. \tag{3.45}$$

We need to deal now with the dependence in μ, which is inside the source function and in the exponential. In the case of the exponential is easy, because we can multiply each side of the equation by a Legendre polynomial and remember that

$$\begin{aligned}(-i)^{-\ell}A_\ell &\equiv \frac{1}{2}\int_{-1}^{1} \mathrm{d}\mu \mathcal{P}_\ell(\mu)A\,,\\ (-i)^{-\ell}j_\ell(x) &\equiv \frac{1}{2}\int_{-1}^{1} \mathrm{d}\mu \mathcal{P}_\ell(\mu)e^{ix\mu}\,,\end{aligned} \tag{3.46}$$

so that it seems we could express the multipoles of Θ as function of Bessel function integrals. Also, not that $j_\ell(x) = (-1)^\ell j_\ell(-x)$.

Dealing with the μ dependence in the source function seems more complicated. However, since it multiplies the exponential, we can repeat the trick from the previous subsection and substitute each appearance it has by a time derivative on the rest of the term:

$$\mu \to \frac{1}{ik}\frac{\mathrm{d}}{\mathrm{d}\tau}\,, \qquad (\mathrm{within}\,\hat{S})\,. \tag{3.47}$$

We can do this for the all terms in which μ appears and use integration by parts to get the desired equality. For instance, for the $-ik\mu\Psi$ term:

$$\begin{aligned}-ik\int_0^{\tau_0} \mathrm{d}\tau\mu\Psi e^{ik\mu(\tau-\tau_0)}e^{-\mathrm{\tau}} &= -\int_0^{\tau_0} \mathrm{d}\tau \Psi e^{-\mathrm{\tau}}\frac{\mathrm{d}}{\mathrm{d}\tau}\left[e^{ik\mu(\tau-\tau_0)}\right] = \\ &= \int_0^{\tau_0} \mathrm{d}\tau e^{ik\mu(\tau-\tau_0)}\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\Psi e^{-\mathrm{\tau}}\right]\,,\end{aligned} \tag{3.48}$$

where the last line is the result of the integration by parts after the surface term vanishes: the $e^{-\mathrm{\tau}(0)}$ nulls all the term $\tau = 0$, and the $\tau = \tau_0$ term does not depend on μ, hence only affects the monopole of the CMB and we cannot detect it with the anisotropies.

This procedure can be applied similarly to the other term depending on μ as well as the one depending on the Legendre quadrupole (which involves a second derivative ($\mathcal{P}_2(\mu) = (3\mu^2 - 1)/2$). Accounting for all this, the solution is

$$\Theta_\ell(k, \tau_0) = \int_0^{\tau_0} \mathrm{d}\tau S(k, \tau) j_\ell \left[k(\tau_0 - \tau)\right] \tag{3.49}$$

with a new source function defined as

$$\begin{aligned} S(k, \tau) \equiv e^{-\mathrm{\tau}} &\left[-\Phi' - \mathrm{\tau}' \left(\Theta_0 + \frac{1}{4}\Pi\right)\right] + \\ &+ \frac{\mathrm{d}}{\mathrm{d}\tau}\left[e^{-\mathrm{\tau}}\left(\Psi - \frac{\theta_b \mathrm{\tau}'}{k^2}\right)\right] - \frac{3}{4k^2}\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\left[e^{-\mathrm{\tau}}\mathrm{\tau}'\Pi\right] . \end{aligned} \tag{3.50}$$

We can see that there are many factors in the source function that depend on $\mathrm{\tau}' e^{-\mathrm{\tau}}$. Thus, let us define the visibility function as a probability density that a photon scattered for the last time at a conformal time τ, given by

$$g(\tau) \equiv -\mathrm{\tau}'(\tau) e^{-\mathrm{\tau}(\tau)} , \tag{3.51}$$

and as it is easy to understand, g decays quickly after recombination since the Universe becomes neutral (numerically, it is due to the prefactor $\mathrm{\tau}'$, the scattering rate, which gets reduced significantly as n_e decreases dramatically). Before recombination, photons scatter many times, so the visibility function is also very small. Therefore, the visibility function is a very sharp function and determines the *width* of recombination. An alternative convention to define τ_* is the time at which g peaks. For the level of precision attempted in this analytic understanding, both moments are roughly the same.

Neglecting the contribution from polarization (which is very small), the source function becomes

$$\begin{aligned} S(k, \tau) \simeq g(\tau) &\left[\Theta_0(k, \tau) + \Psi)(k, \tau)\right] + \frac{1}{k^2}\frac{\mathrm{d}}{\mathrm{d}\tau}\left[g(\tau)\theta_b(k, \tau)\right] + \\ &+ e^{-\mathrm{\tau}}\left[\Psi'(k, \tau) - \Phi'(k, \tau)\right] . \end{aligned} \tag{3.52}$$

Now in order to get an approximate analytical result, we can integrate Θ_ℓ over time, integrating the θ_b term by parts (where as above the surface term vanishes since $g(\tau) = 0$ in both ends):

$$\begin{aligned} \Theta_\ell(k, \tau_0) \simeq &\int_0^{\tau_0} \mathrm{d}\tau g(\tau)\left[\Theta_0(k, \tau) + \Psi)(k, \tau)\right] j_\ell\left[k(\tau_0 - \tau)\right] - \\ &- \frac{1}{k^2}\int_0^{\tau_0} \mathrm{d}\tau g(\tau)\theta_b(k, \tau) j'_\ell\left[k(\tau_0 - \tau)\right] + \\ &+ \int_0^{\tau_0} \mathrm{d}\tau e^{-\mathrm{\tau}}\left[\Psi'(k, \tau) - \Phi'(k, \tau)\right] j_\ell\left[k(\tau_0 - \tau)\right] . \end{aligned} \tag{3.53}$$

The first two integrals are weighted by the visibility functions and are the dominant terms; the latter integral is weighted by $e^{-\tau}$ and only contributes for $\tau \lesssim 1$, which is true after recombination. Furthermore, the gravitational potentials are constant during matter domination, as we saw in the previous chapter, hence the last line will only contribute just after recombination, where radiation still has a small influence in the evolution of the potentials, and after dark energy becomes relevant. The last line is known as the integrated Sachs-Wolf effect, and the two contributions depending on the time are known as the early ISW and the late ISW, respectively.

The fact that the visibility function is so peaked simplifies significantly the first two integrals, the rest of the integrand of which varies at much lower rate. Therefore, we can evaluate them at τ_* and remove them from the integral, which is left to be only the integral of g which is 1 by definition. For instance $\int \mathrm{d}\tau g A = A_* \int \mathrm{d}\tau g = A_*$. Using the recursion relation to express j'_ℓ as function of $j_{\ell-1}$ and j_ℓ and that at τ_* we have $\theta_b = -3\Theta_1$ (from the discussion in previous sections), we obtain

$$\begin{aligned}\Theta_\ell(k,\tau_0) \simeq & \left[\Theta_0(k,\tau_*) + \Psi)(k,\tau_*)\right] j_\ell\left[k(\tau_0-\tau_*)\right] + \\ & + 3\Theta_1(k,\tau_*)\left(j_{\ell-1}\left[k(\tau_0-\tau_*)\right] - (\ell+1)\frac{j_\ell\left[k(\tau_0-\tau_*)\right]}{k(\tau_0-\tau_*)}\right) + \\ & + \int_0^{\tau_0} \mathrm{d}\tau e^{-\tau}\left[\Psi'(k,\tau) - \Phi'(k,\tau)\right] j_\ell\left[k(\tau_0-\tau)\right] .\end{aligned} \tag{3.54}$$

Each term is usually referred to as the monopole term, the dipole or Doppler term, and the ISW, respectively.

The expression above describes the scales where diffusion is not relevant. At smaller scales, since the diffusion scale changes very quickly around recombination, diffusion cannot be included just multiplying the $\Theta_0 + \Psi$ above by the damping. In turn, including the damping in the integral of the visibility function turns out to be a much better approximation. This adds a multiplicative factor in the first line of the expression above of

$$\int \mathrm{d}\tau g(\tau) e^{-k^2/k_{\mathrm{D}}^2(\tau)} . \tag{3.55}$$

These expressions agree with numerical solutions within 10% precision. We can see that these result matches the preliminary expectations at the beginning of the chapter. The monopole depends on $\Theta_0 + \Psi$, and the Bessel functions determine how much anisotropy on a given angular scale $\sim \ell^{-1}$ is contributed by a plane wave with wave number k. On very small angular scales where we can assume plane-parallel flat sky,

$$j_\ell(x) \rightarrow^{x/\ell \to 0} \frac{1}{\ell}\left(\frac{x}{\ell}\right)^{\ell-1/2} , \tag{3.56}$$

i.e., j_ℓ is extremely small for large ℓ if $x < \ell$, or, in our case, Θ_ℓ is very close to zero if $\ell > k\tau_0$. In essence, perturbations on scales k contribute predominantly to angular scales of order $\ell \sim k\tau_0$.

3.5 CMB angular power spectrum

Θ is the perturbation of the CMB characteristic temperature, but we can only observe it today and here (note that the small variation in the position due to the location of the satellites and the time period over all observations have been made are completely negligible). Time-ordered observations are collected in a map as function of position on the sky (an angle), rather than the three-dimensional direction of the incoming photon. This is just a mere change of variables and we can use either frame indistinguishably for denoting the position on the sky.

Therefore, we can expand the temperature perturbation in spherical harmonics as discussed at the beginning of the chapter, where the harmonic indices ℓ and m are the conjugate to the angular position. Also, note that in the flat sky approximation (valid for small angular scales), the harmonic transform can be understood as a 2D Fourier transform (by turning ℓ and m into a 2D vector $\boldsymbol{\ell}$). Thus, the maximum multipole that can be measured is related with the angular resolution of a given experiment. The total number of independent bits of information is given by the number $N_{\rm pix}$ of pixels in the map, which is also equivalent to the number of independent $a_{\ell m}$ coefficients. Therefore, since each multipole ℓ involves $2\ell+1$ m values, we can estimate the maximum multipole accessible by equating $\sum^{\ell_{\rm max}}(2\ell+1) = (\ell_{\rm max}+1)^2 = N_{\rm pix}$. Note that for the CMB there is another limitation to obtain information from very high ℓ besides the angular resolution of the experiment: at some point, the diffusion damping kills any correlation at very small scales and any measured correlation is due to foregrounds and secondary anisotropies.

By definition, the mean of a given coefficient $a_{\ell m}$ vanishes, and therefore we work with their covariance, their power spectrum. Recovering Eq. (3.6):

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'} \,. \tag{3.57}$$

All the measured coefficients are in practice samples for the same underlying distribution. Moreover, for each ℓ value there are $2\ell+1$ m components, so that higher ℓ values have more statistical precision regarding the determination of their underlying distribution. The uncertainty related with the fact that we can only measure one sky and cannot access more information than the $2\ell+1$ components is called the cosmic variance, which for the angular power spectrum scales as

$$\left(\frac{\sigma(C_\ell)}{C_\ell}\right)_{\rm cosmic\,variance} = \sqrt{\frac{2}{2\ell+1}}\,, \tag{3.58}$$

although partial scale coverage adds a factor of $f_{\rm sky}^{-1/2}$ to this estimation. Furthermore, the contamination from foregrounds (which is more difficult to control at larger scales) makes very complicated to reach the cosmic variance limit at the largest scales.

From the relation between $a_{\ell m}$ and Θ_ℓ from Eq. (3.5) we can compute the power spectrum. To compute the variance of the spherical harmonic coefficients we need to compute first the variance of $\Theta(\boldsymbol{k}, \tau_0)$, where we will drop the τ_0 dependence for simplicity. There are two different sources of correlation here: the primordial perturbations (random variable) and their evolution (deterministic process). This allows us (at linear level) to separate them using the transfer function as we did in the previous chapter. In this case we define the transfer

$$\mathcal{T}(\boldsymbol{k}, \hat{\boldsymbol{q}}) \equiv \frac{\Theta(\boldsymbol{k}, \hat{\boldsymbol{q}})}{\mathcal{R}(\boldsymbol{k})}\,, \tag{3.59}$$

which by definition is deterministic and can be removed from the ensemble average. Therefore,

$$\begin{aligned}\left\langle \Theta(\boldsymbol{k}, \hat{\boldsymbol{q}})\Theta^*(\boldsymbol{k}', \hat{\boldsymbol{q}}')\right\rangle &= \langle \mathcal{R}(\boldsymbol{k})\mathcal{R}^*(\boldsymbol{k}')\rangle\, \mathcal{T}(\boldsymbol{k}, \hat{\boldsymbol{q}})\mathcal{T}^*((\boldsymbol{k}', \hat{\boldsymbol{q}}') = \\ &= (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}-\boldsymbol{k}')P_{\mathcal{R}}(k)\mathcal{T}(\boldsymbol{k}, \hat{\boldsymbol{q}})\mathcal{T}^*((\boldsymbol{k}', \hat{\boldsymbol{q}}')\,.\end{aligned} \tag{3.60}$$

We have seen that for scalar perturbations what matters, rather than $(\boldsymbol{k}, \hat{\boldsymbol{q}})$ is (k, μ), so that we find that the power spectrum is (after integrating over $\boldsymbol{k}'$)

$$C_\ell = \int \frac{\mathrm{d}^3k}{(2\pi)^3}P_{\mathcal{R}}(k)\int \mathrm{d}\Omega_{\mathrm{q}} Y^*_{\ell m}(\hat{\boldsymbol{q}})\mathcal{T}(k,\mu)\int \mathrm{d}\Omega'_{\mathrm{q}} Y_{\ell m}(\hat{\boldsymbol{q}}')\mathcal{T}^*(k,\mu')\,. \tag{3.61}$$

We can expand the transfer function as function of the Legendre polynomials as in Eq. (1.91) so that $\mathcal{T}_\ell = \Theta_\ell/\mathcal{R}$, which leaves

$$C_\ell = \frac{2}{\pi}\int \mathrm{d}k k^2 P_{\mathcal{R}}(k)|\mathcal{T}_\ell|^2\,, \tag{3.62}$$

where we have used the orthogonality of the Legendre polynomial and the normality of the spherical harmonics. For a given multipole, the power spectrum is an integral over all Fourier modes of the variance of Θ, and quantifies the variance of the distribution from which the $a_{\ell m}$ coefficients are drawn from. Let us walk over the different scale ranges in the CMB power spectrum.

Ultra-large-scale anisotropies trace perturbations that have entered our horizon only recently, providing a window to the initial conditions. In this regime we can neglect the dipole term in Θ_ℓ, which leaves $\Theta_0 + \Psi$ and the ISW. The former is known as the Sachs-Wolfe effect, and using Eq. (3.11) we have

$$C_\ell^{\mathrm{SW}} \simeq \frac{2}{25\pi}\int \mathrm{d}k k^2 P_{\mathcal{R}}(k)|j_\ell\left[k(\tau_0-\tau_*)\right]|^2\,. \tag{3.63}$$

Substituting the expression of the primordial curvature power spectrum, neglecting τ_* in favor of τ_0 in the Bessel function, and changing the variable to $k\tau_0$, there is an analytic solution to the integral in terms of Gamma functions. If we further assume $n_s = 1$, they simplify and we find that

$$\ell(\ell+1)C_\ell^{\mathrm{SW}} \simeq \frac{8}{25}\mathcal{A}_s \tag{3.64}$$

is a constant, inherited from $k^3 P_{\mathcal{R}}$ being a constant if $n_s = 1$.[6] Deviations from this constant are due to the dipole term becoming relevant at higher ℓ and the late ISW effect –relevant at $\ell \lesssim 30$, since dark energy becomes relevant at $z \lesssim 1$ – (and, to a smaller degree, the red-tilt in the primordial curvature power spectrum). Nonetheless, the amplitude of the power spectrum at these scales can roughly give an idea of the value of $\mathcal{A}_s$.

As ℓ grows the power spectrum probes scales that are within the horizon at recombination, where the acoustic oscillations form and all the terms of Θ_ℓ matter. However, note that since a given value of ℓ has support from a given k range (selected by the Bessel function), we have now a series of peaks and troughs rather than peaks and zeros in the oscillatory pattern of the power spectrum. This also produces that the peak position is slightly shifted towards lower ℓ values, roughly $\ell_{\rm pk} \simeq 0.75\pi\tau_0/r_s$. The dipole term (which, as discussed before, is smaller than the monopole and out of phase with respect to it) contributes to raise all the power spectrum amplitude, but especially the one of the troughs. Notably, the monopole and dipole terms are uncorrelated (mathematically, this is due to the properties of the Bessel functions). Finally, there is a contribution from the early ISW: if we consider that the potentials evolve at time τ_c, all sub-horizon scales $k\tau_c > 1$ will be affected, which through the Bessel function translate to scales $\ell > (\tau_0 - \tau_c)/\tau_c$. Importantly, the early ISW is coherent with the monopole of the source term (i.e., they are proportional to the same Bessel function), which magnifies its impact in the power spectrum through their cross correlation.

So far we have assumed that photons completely free stream to us from the last-scattering surface. However, after reionization, electrons are free again and photons can scatter with them. Consider an optical depth $\mathrm{\tau}_{\rm reio} \equiv \mathrm{\tau}(\tau_{\rm late})$ to a time after recombination. As photons travel through those free electrons, only a fraction $e^{-\mathrm{\tau}_{\rm reio}}$ escape and reach us, while a fraction $1 - e^{-\mathrm{\tau}_{\rm reio}}$ scatters into the beam from all directions (thus any anisotropy that they had cancels out). This involves that for photons coming with a temperature $T(1+\Theta)$, we will measure

$$T(1+\Theta)e^{-\mathrm{\tau}_{\rm reio}} + T(1 - e^{-\mathrm{\tau}_{\rm reio}}) = T(1 + \Theta e^{-\mathrm{\tau}_{\rm reio}})\,. \tag{3.65}$$

This effects only to scales within the horizon at reionization; only those with $\ell \gtrsim \tau_0/\tau_{\rm reio} \sim 100$ are affected. Reionization has a significantly larger impact in the polarization power spectrum.

[6] In the same way that $k^3 P(k)$ is the dimensionless power spectrum per logarithmic k bin, $\ell(\ell+1)C_\ell$ is the angular power spectrum per logarithmic interval in ℓ, and it is the common way to visually represent the angular power spectrum; in particular, we usually plot $\ell(2\ell+1)C_\ell/2\pi$.

CHAPTER 4

APPLIED SESSION 1 COSMOLOGICAL PARAMETERS FROM THE EARLY UNIVERSE

In this Applied Session we will discuss how cosmological parameters affect the observed CMB angular power spectrum and how they can be estimated from the position and amplitude of its peaks. We will focus on H_0 and its relation with the sound horizon, discussing how the degeneracy between them has to be taken into account when looking for solutions to the Hubble tension.

For more detail you can refer to:

Modern Cosmology. 2nd edition. Chapter 9 S. Dodelson and F. Schmidt (2020). Elsevier Press, Cambridge. DOI: 10.1016/B978-0-12-815948-4.00020-6

Hubble constant hunter's guide L. Knox and M. Millea (2020). Phys. Rev. D 101, no.4, 043533. DOI: 10.1103/PhysRevD.101.043533

Cosmic Microwave Background Anisotropies W. Hu and S. Dodelson (2002). Ann. Rev. Astron. Astrophys.
DOI: 10.1146/annurev.astro.40.060401.093926
Check also Wayne Hu website:
http://background.uchicago.edu/ whu/araa/araa.html

4.1 Summary on the physics of the CMB peaks

During the previous lectures, you discussed how acoustic oscillations in the baryon-photon fluid at recombination determine the CMB temperature fluctuations on different scales. You showed how these can be projected to angular scales and give rise to peaks at different multipoles in the observed CMB angular power spectrum. But what do we observe when we look at the sky?

The CMB is observed in the microwave part of the EM spectrum $\sim \mathcal{O}(10-1000)$ GHz. Its first measurements, back in the '60s, detected the sky-averaged signal power per unit solid angle and unit of area, which can be converted into a temperature of $\sim 2.725\,$K by using the black body equation $B_\nu(T) \propto \nu^3 \exp(h\nu/kT)$. After that moment, many observations provided maps of the CMB in different frequency ranges and with different fields of view; here, we only discuss results that relates with the three satellites that realized (almost) full-sky maps of the CMB, probing its $\mathcal{O}(10^{-5}\,\mathrm{K})$ temperature fluctuations in the different directions of the sky: COBE-FIRAS (7°, 1989-1993), WMAP (15 arcmin, 2001-2010) and Planck (5 arcmin, 2009-2013). As figure 4.1 shows, their main difference is the angular resolution, which allow us to resolve fluctuations on smaller and smaller sky-patches.

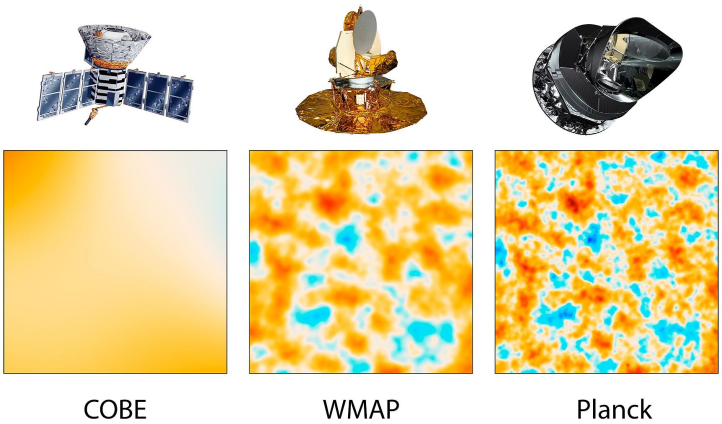

Figure 4.1 Source: NASA/JPL-Caltech/ESA

Recalling what you discussed, this means that, while COBE could neither observe the first peak, Planck gives us access to $\ell > 2000$. As figure 4.2 summarizes, each ℓ represents a certain angular scales and its power in the spectrum indicates how much the temperature fluctuates with respect to the average in patches associated with that particular angular size.

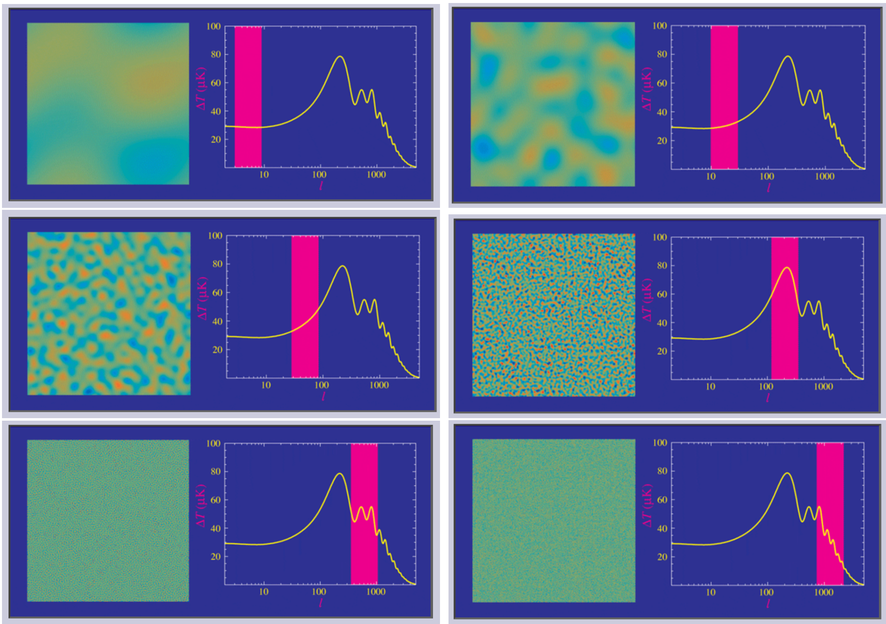

Figure 4.2 Source: Wayne Hu website.

As you saw during the lectures, this can be thought in terms of a forced harmonic oscillator; its solution gives us the first-order position of the peaks in the CMB power spectrum

$$k_{\rm pk} = \frac{n\pi}{\int_0^\tau \mathrm{d}\tilde{\tau} c_s(\tilde{\tau})} = \frac{n\pi}{r_s}\,, \tag{4.1}$$

where the speed of sound with which the wave propagates is defined as

$$c_s(\tau) \equiv \sqrt{\frac{1}{3(1+R(\tau))}} = \sqrt{\frac{1}{3(1+3\bar{\rho}_b/4\bar{\rho}_\gamma)}}\,. \tag{4.2}$$

and the comoving sound horizon (i.e., the distance travelled) through

$$r_s(z) = \int_0^\tau \mathrm{d}\tilde{\tau} c_s(\tilde{\tau}) = \int_z^\infty dz' \frac{d\tilde{\tau}}{dz'} c_s(z') = \int_z^\infty dz' \frac{c_s(z')}{H(z')}\,. \tag{4.3}$$

Think about the properties of the harmonic oscillator and how quantities that enter the standard expression affect the oscillatory behaviour. In the case of CMB, the oscillator is the photon-baryon fluid at recombination, and its fluctuations are due to the presence of sound waves driven at first by self-gravity and pressure. The photon density is directly related with the CMB intensity. How does the **baryon density** relative to the photons affect the oscillations and, therefore, the power spectral shape?

Discussion

We describe the fluctuations in the monopole of the photon temperature through the simplified equation

$$\Theta_0'' + k^2 c_s^2 \Theta_0 = F \;\leftrightarrow\; \ddot{x} + \frac{K}{m} x = F \tag{4.4}$$

The role played in the harmonic oscillator by the combination of mass and restoring force, in the CMB is due to the properties of the photon-baryon fluid. Intuitively, the more **baryons** in the fluid, the more it can be considered as "massive". This has 3 effects, summarized in figure 4.3:

- The speed of sound is smaller for larger baryon density $\bar{\rho}_b$; this implies that the time required for one oscillation is longer and the peak frequency smaller. Thus, **increasing the baryon density pushes the peaks to higher ℓ.**
- The baryon self-gravity drives the compression of the fluid, while it retains it to bounce back. Increasing $\bar{\rho}_b$ shifts the zero point of the oscillations towards the direction of the force i.e., the fluid compresses more. Since the power spectrum is related with the square of the oscillation amplitude, the overall effect is that the odd peaks, associated with contractions (which are the maxima before getting the square) gets higher than the even. Therefore, **increasing the baryon density determines a larger asymmetry between odd and even peaks**.
- The photon-baryon fluid is tightly coupled due to the scatterings between photons and free electrons, and electrons and baryons. Between two consecutive scatterings, a photon travels a mean free path of size equal to the inverse of the derivative of the optical depth $\lambda_{\text{MFP}} = (n_e \sigma_{\text{T}} a)^{-1}$, where n_e is the electron number density, σ_T the cross section and a the scale factor. In a certain time, the number of scatterings is $N \propto n_e \sigma_T$ (more electrons imply more and closer scatterings) and overall a photon travels $\lambda_D = N\lambda_{\text{mfp}}$; this quantity is the diffusion length. Fluctuations on scales smaller than the diffusion length get erased by free streaming, that restores the average temperature. The size of λ_D depends on n_e, which in turn depends on $\bar{\rho}_b$: **increasing the baryon density decreases the diffusion length and so decreases the damping of the high ℓ.**

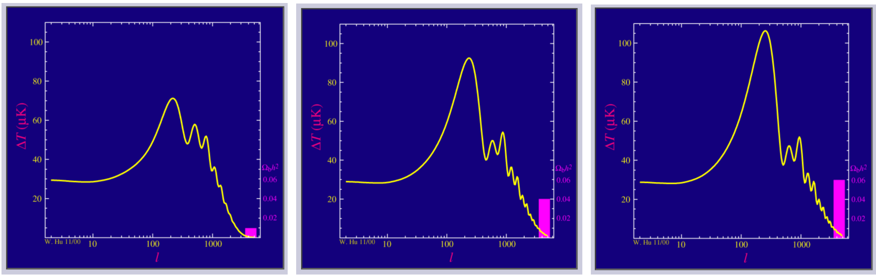

Figure 4.3 Baryon effect. Source: Wayne Hu website.

4.2 Effect of the cosmological parameters

The previous discussion already tells us that the CMB retains information on the Universe content and properties. From the theoretical point of view, its spectral shape is computed by assuming a certain set of fiducial values for the 6 parameters that describe the ΛCDM cosmological model

$$\{\Omega_b h^2, \Omega_c h^2, 100\theta_*, \tau, \ln(10^{10}A_s), n_s\} \tag{4.5}$$

where $h = H_0/100\,\mathrm{km/s/Mpc}$, $\Omega_{b,c}h^2 = \bar{\rho}_{b,c}h^2/\rho_{\mathrm{crit}}$ are the physical baryon and cold dark matter densities, θ_* is the angular scale of the sound horizon at recombination, τ is the optical depth to Thompson scattering to reionization, A_s the amplitude of the primordial power spectrum, n_s its spectral index. As figure 4.4 summarizes, we will see that the parameters affect differently different parts of the power spectrum.

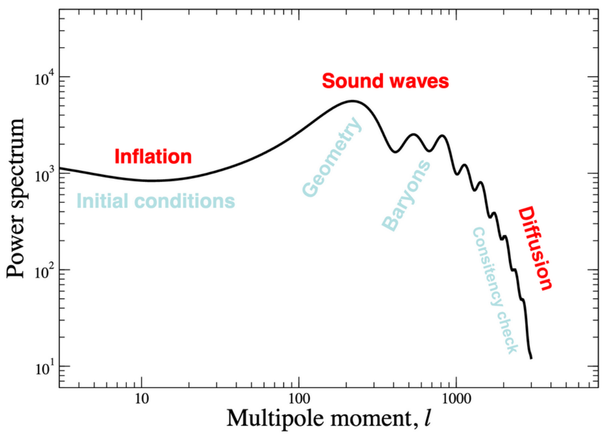

Figure 4.4 Source: `cmb.wintherscoming.no/theory_observables_content.php`

While we proceed by checking what happens to the spectrum when one or more parameters varies, remember that from an observational point of view we observe one CMB spectrum realization and we look for the best-fit values of the parameters to recover it. Table 4.5 collects the estimates of the primary and the main derived parameters obtained by Planck 2018.

$\Omega_b h^2$	0.022383	0.02237 ± 0.0
$\Omega_c h^2$	0.12011	0.1200 ± 0.0
$100\theta_{MC}$	1.040909	1.04092 ± 0.0
τ	0.0543	0.0544 ± 0.0
$\ln(10^{10} A_s)$	3.0448	3.044 ± 0.0
n_s	0.96605	0.9649 ± 0.0

$\Omega_m h^2$	0.14314	0.1430 ± 0.0011
H_0 [km s^{-1}Mpc^{-1}] . . .	67.32	67.36 ± 0.54
Ω_m	0.3158	0.3153 ± 0.0073
Age [Gyr]	13.7971	13.797 ± 0.023
σ_8	0.8120	0.8111 ± 0.0060
$S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$ . .	0.8331	0.832 ± 0.013
z_{re}	7.68	7.67 ± 0.73
$100\theta_*$	1.041085	1.04110 ± 0.00031
r_{drag} [Mpc]	147.049	147.09 ± 0.26

Figure 4.5 Main parameter estimation from Planck 2018.

Baryons: $\Omega_b h^2$

We already discussed the effects baryons have on CMB power spectrum. Since they show up in different ways, their abundance can be estimated from different properties (both from the acoustic peaks an the damping tail), thus providing many consistency checks and breaking internal degeneracies since the values of $\Omega_b h^2$ measured in different parts of the spectrum, should match. Note that the quantity the CMB measures is $\Omega_b h^2$: the presence of the h is due to the fact that the relevant parameter in determining the CMB shape is the physical density. Moreover, h collects the uncertainties on the measurement of the sound horizon, as we will see later.

Dark matter: $\Omega_c h^2$

On one side, DM creates the backbone of over- and under- densities on which the photon-baryon fluid is found: for this reason, its gravitational field can be seen as part of the external driving force F that drives the harmonic oscillator, i.e., it has similar effects to what we already discussed for baryons.

However, there is another important effect: if we keep the baryon density fixed, increasing the amount of DM increases the total matter amount of the Universe. By doing so, the epoch of matter-radiation equality gets anticipated and the capability of radiation in driving the oscillations decreases. Therefore, increasing the overall matter content decreases the amplitude of the peaks.

Primordial power spectrum: A_s, n_s

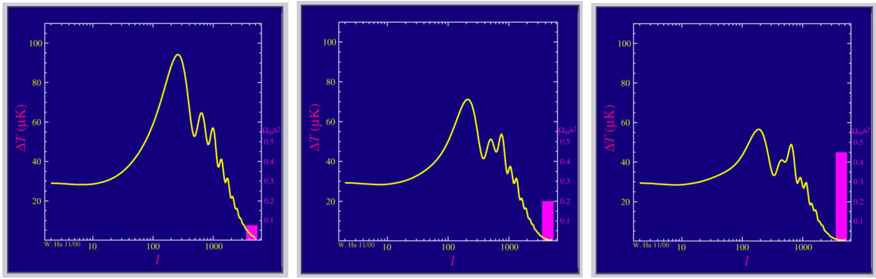

Figure 4.6 DM+baryons effect. Source: Wayne Hu website.

How do changes in the amplitude of the primordial power spectrum A_s and in its power spectral index n_s affect the CMB power spectrum?

Discussion

We write the primordial power spectrum as

$$P(k) = \frac{2\pi^2}{k^3} A_s \left(\frac{k}{k_p}\right)^{n_s - 1} \tag{4.6}$$

where $k_p = 0.05\,\mathrm{Mpc}^{-1}$ is the pivot scale, set by convention. Changing the amplitude of the primordial power spectrum re-scales the amplitude of all the multipoles by the same factor. If we change n_s to $n_s + \alpha$ with $\alpha > 0$, the power spectrum amplitude goes as $P_{\mathrm{mod}}(k) \propto P(k)k^\alpha$, therefore it increases for large k and decreases for $k < 1\,\mathrm{Mpc}^{-1}$. Then, also the CMB power spectrum gets changes by $(\ell/\ell_p)^\alpha$, where ℓ_p is the multipole over which k_p is projected. This increases the power on the small scales, while on the large scales it is compensated by the effect of the Bessel function used in the projection.

Reionization optical depth: τ

After recombination, the Universe is filled with neutral hydrogen until the first stars form. The radiation these emit reionizes the gas and increases the number of free electrons, on which CMB photons can eventually scatter. If the number of electrons is large, the number of scatterings is high and fluctuations are washed out. This only affects modes that already entered the horizon at reionization i.e., $\ell > \tau_0/\tau_{\mathrm{reion}} \sim 100$, while large scales are unaffected. The effect of τ, then, is degenerate with the combination of A_s and n_s.

The suppression can be estimated thinking that the number of scatterings is related to the optical depth as $\tau' = -(n_e \sigma_T a)$ and that only a fraction exp(−τ) of the photons that travel through the medium can escape and be observed. The same factor then rescales the peak amplitudes on small scales.

Curvature and dark energy: Ω_K, Ω_Λ

While the parameters we discussed previously affect the oscillations at recombination, the DE density Ω_Λ and the curvature $\Omega_K = 1 - \sum_i \Omega_i$ modify the way we observe the CMB power spectrum.

What is the main quantity that change when we move from a flat to a close Universe? How does this affect the peaks?

Discussion

The curvature affects the angular diameter distance:

$$D_A = \frac{1}{1+z}\int_0^z \frac{dz'}{H(z')} = \frac{\tau}{1+z}\,. \tag{4.7}$$

Since the same physical scale subtends a larger angle in a closed Universe than in a flat Universe, the inferred distance is smaller, including its value at recombination D_A^*. This mainly shows up in the position of the first peak, since in a flat Universe $\tau \propto (1+z)^{-1/2}$ and

$$\begin{aligned} &\theta_{\mathrm{pk}=1} \sim \frac{r_s^*}{D_A^*} \sim \frac{\tau_*\sqrt{3}}{\tau_0/(1+z_0)} \sim \frac{\tau_*}{\tau_0} \sim \frac{1}{\sqrt{1100}} \sim 2^\circ \\ &\to \ell_{\mathrm{pk}=1} \sim \frac{2\pi}{\theta_{\mathrm{pk}=1}} \sim 200\,. \end{aligned} \tag{4.8}$$

where we assumed $c_s \sim 1/\sqrt{3}$. In a closed Universe, instead,

$$\tau^{\mathrm{closed}} \propto \frac{\sin[\tau H_0\sqrt{|\Omega_K|}}{H_0\sqrt{|\Omega_K|}} > \tau \tag{4.9}$$

which implies that the CMB first peak should be observed at

$$\theta_{\mathrm{pk}=1}^{\mathrm{closed}} \sim \frac{\tau_*^{\mathrm{closed}}}{\tau_0^{\mathrm{closed}}} \sim \frac{\tau_*}{\tau_0\sin(\tau_0\sqrt{|\Omega_K|})} > \theta_{\mathrm{pk}=1} \tag{4.10}$$

(where we did not include the sin in the recombination term since it is at higher cosmic times). Finally we get $\ell_{\mathrm{pk}=1}^{\mathrm{closed}} < 200$ (this can be extended also to the other peaks): **in a closed Universe, the peaks shift at lower** ℓ and viceversa for the open Universe, as figure 4.7 shows. Since $\ell_{\mathrm{pk}=1} \sim 200$ is well measured, the Universe has to be nearly spatially flat.

In a flat Euclidean Universe, how does DE affect the CMB peaks?

Discussion

Similarly to the previous case, DE affects the angular diameter distance

$$D_A^* \sim \frac{1}{H_0} \int_0^{z^*} \frac{dz'}{\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda(1+z)^4}} \tag{4.11}$$

Increasing Ω_Λ, D_A^* decreases (which also implies that the age of the Universe is smaller): **with more DE, peaks shift at smaller ℓ.** Spatial curvature and DE both change the angular diameter distance to recombination and hence shift the peak angular locations, as figure 4.7 shows.

DE also modifies the large scales by modifying the way the evolution of the gravitational potential, i.e., it changes the integrated Sachs Wolfe effect.

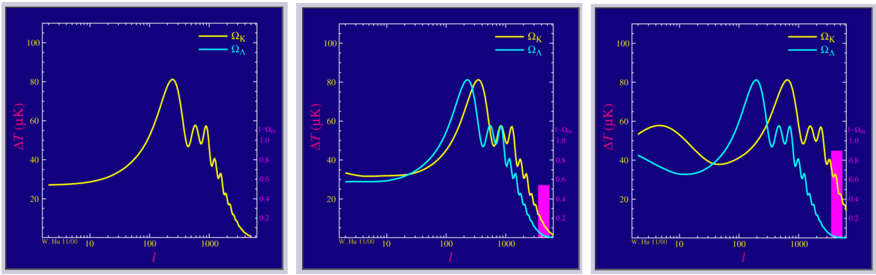

Figure 4.7 Curvature and DE effect. Source: Wayne Hu website.

H_0 and the sound horizon

The parameter set described up to now implicitly accounts for the Hubble parameter, since on one side it constrains $\Omega_{b,c}h^2$, and on the other $\Omega_K = 1 - (\Omega_b + \Omega_m + \Omega_\Lambda) \sim 1$. Up to now we did not discuss the role of the sound horizon: we will now show that it is degenerate with the choice of H_0.

Let us go back to the comoving sound horizon

$$r_s(z) = \int_0^\tau \mathrm{d}\tilde{\tau} c_s(\tilde{\tau}) = \int_z^\infty dz' \frac{d\tilde{\tau}}{dz'} c_s(z') = \int_z^\infty dz' \frac{c_s(z')}{H(z')} . \tag{4.12}$$

Depending on the value of z we choose, we can define this quantity at different "important moments". Here we are interested in the sound horizon

- at recombination/CMB last scattering surface, $r_s^* = r_s(z_* \sim 1100)$, which is defined (looking from now backwards) as the moment in which the optical depth to Thomson scattering reaches 1;

- at radiation drag, $r_s^{\rm drag}(z_{\rm drag} < z_*)$, which is defined (looking from now backwards) as the moment in which the baryon drag epoch ends, slightly later in cosmic time than recombination.

The first is relevant when discussing CMB power spectra, the second comes into play when studying the Baryon Acoustic Oscillations (BAO) i.e., the way the oscillations in the photon-baryon fluid get imprinted in the baryon power spectrum. The two quantities differ because, since the number of photons is larger, baryons take more time to fully decouple. The conversion between r_s^* and $r_s^{\rm drag}$ is straightforward ($r_s^{\rm drag} \sim r_s^*(1-2\%)$ in ΛCDM) and almost model-independent; for this reason we can safely pass from one to the other.

The sound horizon size at recombination can be estimated knowing that

$$c_s^2 = \frac{1}{1 + 3\bar{\rho}_b/4\bar{\rho}_\gamma} \tag{4.13}$$

$$H(z)^2 = \frac{8\pi G}{3}[\bar{\rho}_\gamma + \bar{\rho}_\nu + \bar{\rho}_m + \bar{\rho}_\Lambda] \tag{4.14}$$

From the second equation, we can neglect $\bar{\rho}_\Lambda$ (not effective at the time of recombination), while $\bar{\rho}_\nu$ can be estimated based on $\bar{\rho}_\gamma$. Thus, we are left with the dependence on $\bar{\rho}_\gamma, \bar{\rho}_m, \bar{\rho}_b$, which we know we can estimate from the CMB. Therefore, simply using CMB data we can compute r_s^*. But so far we know that we have to deal with angular distances and peak positions.

How can we measure the r_s^* angular size from the CMB power spectrum?

Discussion

Consider two near-by peaks ℓ_p, ℓ_{p+1} in the CMB power spectrum. We can convert them to k scales as $\ell_{p,p+1} \simeq k_{p,p+1} D_A^*$, where $D_A^* = D_A(z_*)$ is the angular diameter distance between us and recombination. From the harmonic oscillator solution, we have that $k_p = p\pi/r_s^\star$ and $k_{p+1} = (p+1)\pi/r_s^*$, so we can consider

$$\Delta\ell = \ell_{p+1} - \ell_p = k_{p+1}D_A^* - k_p D_A^* = (p+1-p)\frac{\pi D_A^*}{r_s^*} = \frac{\pi}{\theta_s^*}\,. \tag{4.15}$$

Therefore, from the peak spacing in the CMB we can estimate the angular size of the sound horizon at recombination, $\theta_s^* = \pi/\Delta\ell$.

Thus, we have a measurement of r_s^* and θ_s^* and we can combine them to estimate the distance between us and recombination, $D_A^* = r_s^*/\theta_s^*$. But we also have an analytical way to express this quantity

$$D_A^* = \frac{1}{1+z_*}\int_0^{z_*} \frac{dz'}{H(z')} \tag{4.16}$$

So, by inverting this relation we can get $H(z)$ and use it to estimate $H(z = 0) = H_0$. This is the Hubble parameter that describes the Universe expansion rate; its estimate from the CMB is indirect: for this reason, its value is model dependent and it can be degenerate with other parameters.

Provided what we said up to this point, which are the main degeneracies betweem H_0 and other parameters? Starting from them, can we modify the value of H_0 without affecting too much the CMB spectral shape?

We want to modify H_0, so we "attack" the different quantities that enter

$$D_A^* = \frac{1}{1+z_*}\int_0^{z^*} \frac{dz'}{H_0\sqrt{\Omega_\Lambda + \Omega_m(1+z)^3 + \Omega_\gamma(1+z)^4}}\,, \tag{4.17}$$

$$D_A^* = \Delta\ell \cdot r_s^*/\pi\,. \tag{4.18}$$

Since from these equations $H(z)$ and $r_s(z)$ are strongly related, we can look at the $r_s - H_0$ plane, as in figure 4.8. Instead of using r_s^*, we can also refer to $r_s^{\rm drag}$, so we are able to compare the CMB measurements with late time probes. We will discuss in detail in the next lecture that there is an inconsistency between the H_0 measurements done with CMB and with late Universe probes, the former providing a lower value of H_0 than the latter. But since $H(z)$ is degenerate with r_s, the inconsistency can be thought on the $r_s - H_0$ plane; to solve it, we would like to find a way to move Planck estimates (that are model dependent) in the intersection between SNe and BAO measurements i.e., lower r_s^* ($\sim 7\%$) and higher H_0; we now present solutions in this direction, but we keep a more detailed discussion for the final lecture of the course.

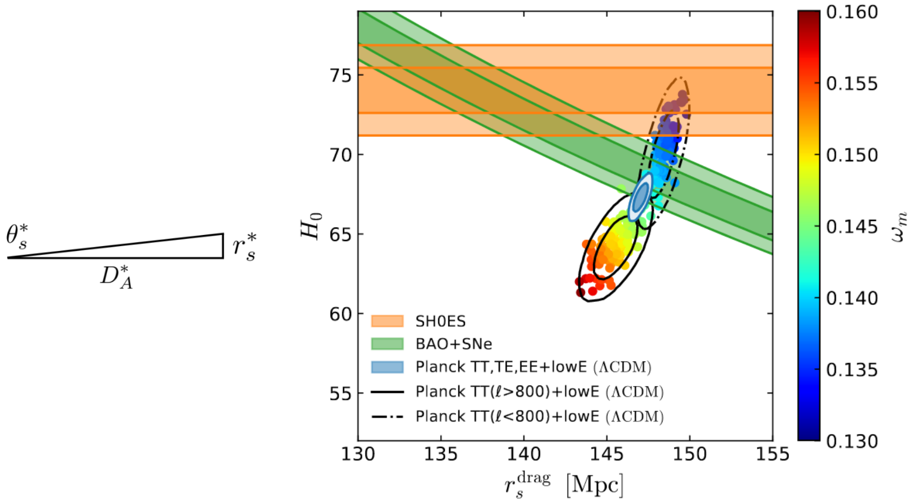

Figure 4.8 Source: Knox, Millea.

- Ω_m: if we increase its value, since $r_s^\star \propto \Omega_m^{-1/2} \to \delta r^\star / r_s^\star \propto -1/2\, \delta\Omega_m/\Omega_m$ decreases; including the effect of radiation, this soften the $1/2$ factor to $1/4$. However, we need to keep $\Delta\ell$ fix because we observe it, so we need to decrease D_A^* of the same quantity to counterbalance. But in D_A^* there is no radiation soften, so decreasing Ω_m lowers D_A^* too much. We then need to decrease Ω_Λ in the $H(z)$ computation, which however has a different z dependence than matter. The two things combined imply that $H(z)$ increases when Ω_m dominates, but it decreases when Ω_Λ takes place. Consequently, H_0 decreases and variations on Ω_m in the $r_s - H_0$ plane lead us to move orthogonally to the SNe constraints.
- D_A^*: changing the post-recombination cosmic evolution we can change the shape / value of the angular diameter distance, so to reduce r_s^* estimate.
- c_s: we can change its value by changing the pre-recombination physics: this would affect $r_s^\star$ but not H_0. The speed of sound is related to $\partial P/\partial\rho \propto \sqrt{1 + 3\bar\rho_b/4\bar\rho_\gamma}$ in the baryon-photon fluid. To lower c_s, we have to increase the inertia of the fluid without changing $\bar\rho_b$ too much, since its value is well measured on CMB. We could introduce some new non-relativistic species, tightly coupled to the photons or the baryons so that $\bar\rho_b \to (\bar\rho_b + \bar\rho')$: if we keep $\bar\rho_\gamma$ fixed, however, this new species should also affect the odd-even peak height ratio. So, if they existed they would be indistinguishable from baryons at the level of CMB and they would be already taken into account in our previous discussion. A partial coupling between baryons and DM could have similar effect; to not affect CMB temperature, polariation and lensing its amount has to be so small that its effect on c_s is negligible.
- z^*: the sound horizon computation requires to know the conformal time to the end of the baryon drag epoch, since $r_s^* = c_s \int_{z_*}^{\infty} dz/H(z)$. r_s^* decreases if z_* moves back in time, to higher photon temperatures: for this to happen, we should increase the recombination temperature e.g. with a stronger EM interaction due to a time dependency in the value of the finite structure constant α. However, variations in α or other recombination physics (e.g. faster cooling of the photons that would keep the recombination temperature at the same temperature but anticipate it back in time) affect the shape of the damping tail; therefore, we can only introduce variations $\delta\alpha/\alpha \sim 0.7 \cdot 10^{-3}$ between recombination and today, which is too tiny to lead to significant variations in the recombination temperature (atomic energies are linearly proportional to α).
- $H(z)$: finally, we can assume that the Universe expands differently from ΛCDM, for example by increasing $H(z)$ prior to recombination adding some extra component. Increasing $H(z)$ in this way decreases τ_* (we need less time to reach the same temperature) and increases n_e at the same scale factor (electrons can be found more close by), therefore recombination takes place at higher temperature. To visualize what happens in

this case, we can plot how fractional changes in the Hubble parameter reflects to the fractional change in the sound horizon r_s^* and in the photon diffusion scale λ_D, as we do in figure 4.9. The sound horizon slowly increases until recombination, i.e. it is most sensitive to the expansion rate at higher z. The damping scale instead is extremely sensitive at the time of recombination itself. Therefore, each new component that we want to include, to be effective must act prior and near recombination but has to be such that the combination of the model parameters mimics the effect of photon diffusion damping, so to compensate the changes in the CMB. Possible solutions in this sense are additional thermal relativistic species (= extra neutrinos) and early dark energy.

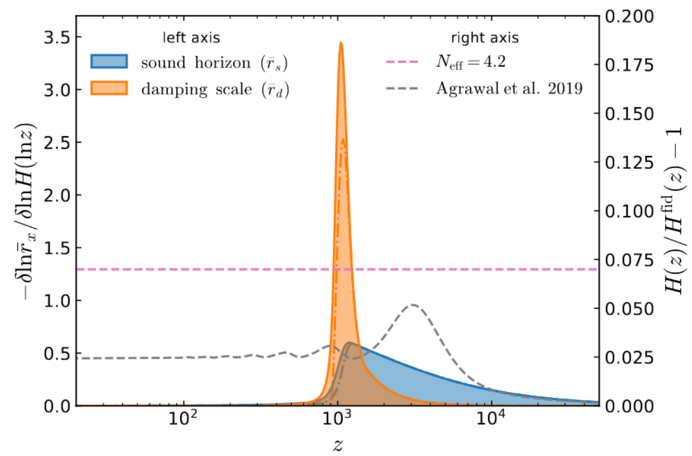

Figure 4.9 Source: Knox, Millea.

The first ones are described by the parameter $N_{\rm eff}$, i.e. the number of effective degrees of freedom, related with the energy density of relativistic particles. If $N_{\rm eff}$ increases, in the computation of $H(z)$ we have to account for changes in the relation between $\bar{\rho}_\nu$ and $\bar{\rho}_\gamma$. But then we want $\theta_s^\star$ (and $\Delta\ell$) not to change, so we also have to change D_A^* so that the overall shape and amplitude on the BAO $(D_A/r_s, H(z)r_s)$ are untouched even if H_0 is larger (this is why the $N_{\rm eff}$ line in the plot is horizontal). However, this change in $N_{\rm eff}$ alters the ratio between r_s^* and the damping scale (which is well constrained) and shifts the CMB peaks, so it is difficult to obtain.

As for early dark energy, it is usually modelled as a scalar field that close by recombination acts as a cosmological constant. It affects not only the amplitude of $H(z)$ but also its shape in z: fine tuning its parameters it is then possible to change H_0 preserving the other CMB-related quantities.

CHAPTER 5

LECTURE 3: MEASURING THE HUBBLE CONSTANT AND THE BACKGROUND EXPANSION

In the last chapter we discussed the CMB power spectrum and how it depends on the cosmological parameter and the actual cosmological model considered. This led us to understand how the value of H_0 can be inferred from the CMB anisotropies and why it is a model-dependent value. The current value reported by Planck assuming ΛCDM is $H_0 = 67.27 \pm 0.60$ km/s/Mpc.

Measurements independent from cosmology can be obtained from the recession velocity of sources in the very local Universe. That requires the calibration of a distance ladder which instead depends on astrophysics. The most precise of these measurements corresponds to the distance ladder studies using cepheids and supernovae type Ia by the SH0ES collaboration, which measured $H_0 = 73.04 \pm 1.04$ km/s/Mpc (7).

As it can be easy to appreciate, this measurement and the inferred value from *Planck* assuming ΛCDM are in $\sim 5\sigma$ tension.[1] Comprehensive studies to control potential systematic errors have returned no significant variation of

[1]This is a naive estimation of the distance in terms of the marginalized 68% confidence level uncertainties assuming Gaussian posteriors. Multidimensional parameter space and

these values, which arises the question of whether this discrepancy may imply the failure of ΛCDM and the need of extending the model to reconcile the measurements.

Finally, values from H_0 can be inferred from low-redshift measurements of the expansion rate of the Universe, which depend on cosmology but can anchor the study of CMB anisotropies by providing the expansion history of the Universe at low redshift. A compilation of measurements and inferred values of H_0 assuming ΛCDM can be found in Fig. 5.1.

This chapter discusses low-redshift measurements and inferences of H_0, for which we need to introduce first the different definitions of distances in cosmology. Measuring distances in the Universe is far from trivial, since the actual Universe is expanding. This means that the scale factor of the Universe when light leaves a given source at redshift z grows as light travels to us. Furthermore, the expansion rate also changes with time. In order to capture this, we can define the comoving distance, defined by the distance $\mathrm{d}x = \mathrm{d}t/a = \mathrm{d}\tau$ that light travels over a small time interval. This corresponds to a total comoving distance between us and a redshift z of

$$\chi = -\int_{t_0}^{t} \frac{\mathrm{d}t'}{a(t')} = \int_{a(t)}^{1} \frac{\mathrm{d}a'}{a'^2 H(a')} = \int_0^z \frac{\mathrm{d}z'}{H(z')}\,, \tag{5.1}$$

which for very small redshifts can be approximated by $\chi \approx z/H_0$. However, comoving distances are hard to measure and are more related to theory. In practice, distances are measured from the physical angle that an object or given scale subtends on the sky (standard ruler) or through the measured flux of a source of known luminosity (standard candle, and the recent analog of standard sirens since the discovery of the gravitational waves).

For small angles, a physical scale l transverse to the line of sight that subtends and angle θ on the sky will be at a distance $D_A = l/\theta$ known as the angular diameter distance. Note that the comoving size of the object is l/a', where a' is the scale factor at the time where the object is (e.g., when the light is emitted). In a flat Universe, we can relate this to θ and the comoving distance as $\theta = (l/a)/\chi$, such as the angular diameter distance is

$$D_A(a) = a\chi(a) = \frac{\chi}{1+z}\,. \tag{5.2}$$

During the last years, and especially in the context of measurements of the baryon acoustic oscillations (BAO) from galaxy surveys that we will discuss below, it is common to express this distance in terms of the comoving angular diameter distance $D_M = (1+z)D_A$. Known scales, such as the sound horizon r_d at radiation drag of the photon-baryon plasma and the signatures of which

non-Gaussian posteriors complicate the interpretation of this value and motivate the development of robust diagnosis of tension. For a discussion, see e.g., section 3 of Ref. (8).

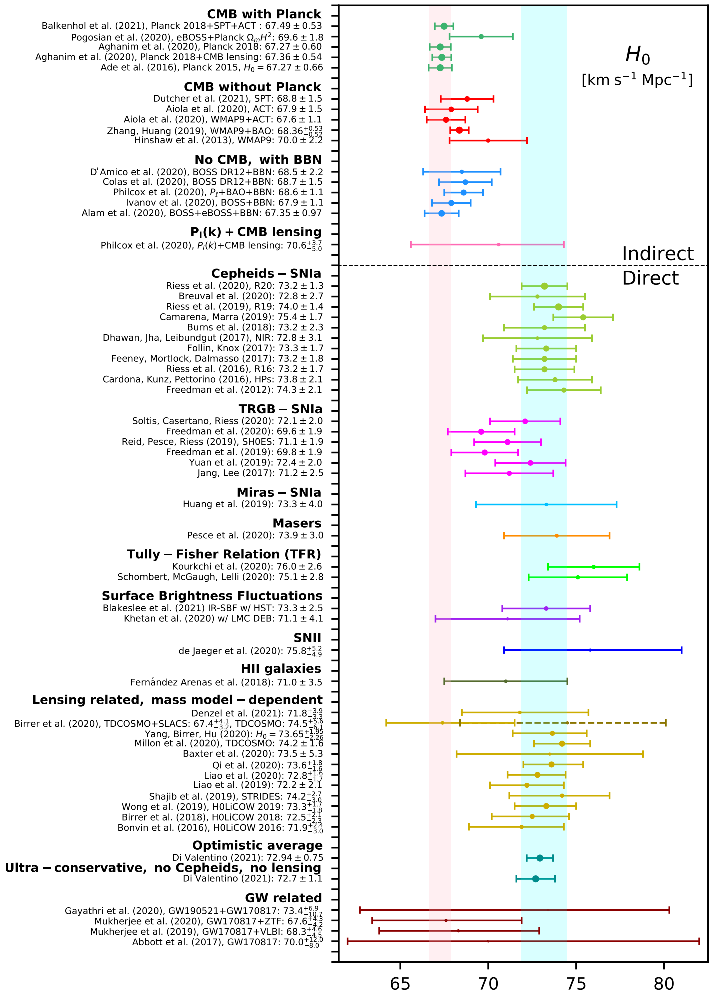

Figure 5.1 68% CL constraints of the Hubble constant H0 through direct and indirect measurements by different probes performed over the years (until 2021). The cyan vertical band corresponds to the direct measurement of H_0 from SH0ES (9) and the light pink vertical band corresponds to the inferred value from *Planck* (6) assuming ΛCDM. Figure from Ref. (10).

can be found in the current distribution of galaxies, are referred to as standard rulers.

Sources with known luminosity are called standard candles and allow us to measure luminosity distances. The measured flux of a source at z goes as $F \propto L'\chi^2(a)$, where L' is the luminosity through a spherical shell of radius

$\chi(a)$. At that distance, both the energy of the photons and a fixed time interval have redshifted, so that $L' = La^2$. Therefore, in a flat Universe, the luminosity distance (defined from the arguments above) is given by

$$D_L = \frac{\chi(a)}{a} = (1+z)\chi(z)\,. \tag{5.3}$$

For non-flat universes, the relation $D_L = (1+z)^2 D_A$ still holds.

Of course, the precision of current distance measurements to nearby galaxies has improved dramatically, what has shifted the focus on the control of systematic errors. We will not have time to discuss in detail any of the probes of H_0, but we will summarize the main measurements and comment on benefits, weaknesses and potential of each probe.

5.1 The local distance ladder

The distance ladder provides the only strictly empirical (i.e., independent of the cosmological model once the cosmological principle and general relativity have been assumed) to measure H_0 is the distance ladder. This term refers to the combination of different distance calibrators used in different steps of the *ladder* to measure the distance-redshift relation. Then, from the redshift it is possible to obtain the recession velocity of the emitter with respect to us and from there extract the H_0 value. The most used approach is to use geometric distances from parallax measurements to calibrate the luminosity to a specific source that can be treated as standard candle and can be detected at larger distances.

To obtain a precise measurement of H_0, however, it is necessary to extend the distance ladder far enough so that the measured redshift comes predominantly from cosmological redshift, largely unaffected by peculiar velocities. The higher the redshift, the highest its contribution to the total redshift. Nonetheless, if the distances are too large, there is impact from the cosmological parameters controlling the evolution of the expansion rate, e.g., Ω_m. Therefore, there is a trade-off between reducing the impact of peculiar velocities and of cosmological parameters other than H_0 itself.

The most powerful standard candle we know to date are supernovae type Ia (SNeIa), since their luminosity can be standardized in terms of the color and width of their light curve, and environmental parameters that are sometimes used (depending on the specific analysis). For each SNeIa, the light-curve fit returns the light-curve amplitude x_0 for which $m_B \equiv -2.5\log_{10} x_0$; the stretch parameter x_1 controlling the light-curve width; and the light-curve color c including intrinsic color and dust; among others. The SNeIa light curves can be standardized in terms of the distance moduli μ, defined as (see e.g. (11) and references therein)

$$\mu \equiv m_B + \alpha x_1 - \beta c - M_B - \delta_{\rm bias} - \delta_{\rm host}\,, \tag{5.4}$$

where α and β are global nuisance parameters relating stretch and color, respectively, to luminosity, M_B is the fiducial absolute magnitude of SNeIa, and δ_{bias} and δ_{host} are correction terms to account for selection biases and residual correlations between standardized brightness of a SNeIa and the host-galaxy properties, respectively. Although we do know that SNeIa have the same absolute magnitude after standardization, we do not know its actual value. Thus, SNeIa are relative distance indicators, since they depend on external calibrators.

However, although extremely bright, supernovae are rare events, and there are not enough detections in the local Universe to be calibrated directly with parallaxes. Therefore, an intermediate rung in the distance is required. The two most precise to date are variable cepheid stars and the tip of the red giant branch (TRGB). These sources can be calibrated with parallax measurements and other geometric distance calibrations. Unfortunately, they are too faint to be detected at great distances, and measurements without another rung are limited by the uncertainties introduced from peculiar velocities (see e.g., Ref. (12)). Therefore, in practice, the intermediate rung is used to reach large enough volumes to find enough SNeIa in the same hosts and calibrate their luminosity. Figure 5.2 shows an example of the complete distance ladder from the SH0ES collaboration, which uses geometric measurements to nearby cepheid stars to calibrate them, distant cepheids to calibrate nearby SNeIa, and distant SNeIa to measure the distance-redshift relation and measure H_0.

Notably, as can be seen in Fig. 5.1, the H_0 measurements using either cepheids and the tip of the red giant branch, although compatible between them, lead to very different conclusions regarding the H_0 tension. While the tension between the results from *Planck* and SH0ES is $\sim 5\sigma$, H_0 measurements from CCHP using the tip of the red giant branch are roughly consistent with *Planck*. This is unexpected, since conceptually both SH0ES and CCHP programs are very similar (they basically differ only in the SNeIa luminosity calibrator) and the samples are comparable.

Cepheids can be used as standard candle thanks to a known relation between the luminosity of these stars and their period of luminosity variation and color, although there is also a dependence on the metallicity. Cepheids are yellow supergiants that are generally found in high-surface-brightness area in star-forming regions, then susceptible to photometric errors due to crowding and blending, as well as dust extinction. However, comprehensive studies have returned a detailed systematic error budget concluding that none of these source of errors can explain the H_0 tension with Planck.

In turn, distance measurements using the tip of the red giant branch are based on the core helium-flash luminosity at the end phase of red giant branch evolution of low-mass stars, which empirically show a sharp discontinuity at a well-defined luminosity in a color-magnitude diagram. Old, blue metal-poor giant branch stars located at the tip of the red giant branch are *actual* standard candles that do not require standardization, although the evolution

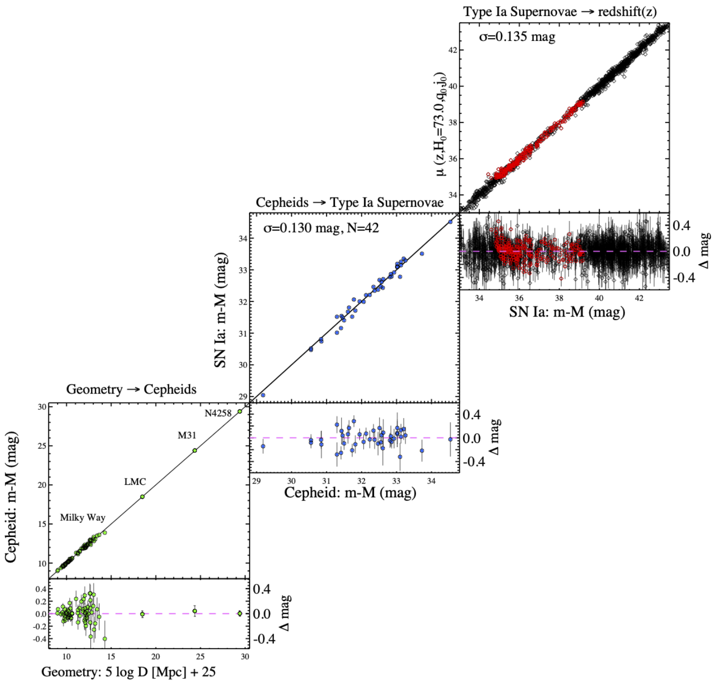

Figure 5.2 Complete distance ladder, showing the simultaneous agreement of distance pairs –geometric and cepheid-based (lower left), cepheid and SNeIa based (middle), and SNeIa and redshift based (top tight)– that provides the measurement of H_0. For each step, measurements in the x-axis serve to calibrate a relative distance indicator on the y-axis. Figure from Ref. (7).

of stars is also affected by its metallicity. In turn, since the method is best applied in the outer halos of galaxies, the effect of crowding and blending are minimal. The uncertainties in the H_0 measurements using the tip of the red giant branch is due to a more limited sample of calibrators for its luminosity. Nonetheless, nearby distance measurements using either cepheids or the tip of the red giant branch are consistent, which suggests that the zero-point calibration of the methods is not the primary reason for the different H_0 measurements.

There are ongoing efforts to improve these measurements that involve increasing the sample of high-quality calibrations of SNeIa (improving the connection between the second and third rung), increasing the number of independent geometric calibrations of cepheids and/or the tip of the red giant branch (improving the connection between the first and second rung), homogenizing the sample using the same instruments as much as possible to minimize calibration errors, and taking measurements in different wavelengths (especially the infrared) to reduce the systematics related to dust and reddening. There have been claims for the H_0 tension to be explained by non-homogeneous nearby and distant SNeIa, as well as whether potential particularities of the local volume, such as being located in an underdense region, but it has been shown that the expected impact from these sources of error is negligible.

These are the two main measurements of H_0 in the local Universe, but there are many other, if currently not competitive, alternatives. The distance ladder can be built using variable red giant stars (MIRAS) (13) to calibrate SNeIa; substituting the SNeIA by the surface brightness fluctuations method (14) or the Tully-Fisher relationship (the correlation between the rotation velocity of a galaxy and its absolute luminsoity) (15), etc. H_0 can also be measured without the need of a distance ladder directly measuring geometric distances to megamasers in the Hubble flow (16), and using standard sirens from neutron-star mergers with electromagnetic counterpart (17).

Generally, all the local measurements (with the exception of some of the measurements using the tip of the red giant branch and SNeIa) cluster around high values of H_0, those measured by the SH0ES collaboration, although most of them with larger error bars. This seems to indicate that all of them would have affected by similar systematics, although the wide range of uncertainty levels limit the application of this argument. Further references and discussions can be found in Ref. (10).

5.2 Measuring the late-time expansion history

As mentioned above, CMB constraints on low-redshift parameters, such as H_0, are necessarily model dependent. While the CMB anisotropies are very powerful to constrain physics before recombination through its effect in the perturbations in the photon-baryon fluid (see previous chapter), assuming a cosmological model is necessary to extrapolate the evolution of the Universe from the last-scattering surface to today. Once freedom beyond Ω_m and H_0 is given to that evolution, constraints from the CMB weaken significantly. This is why low-redshift probes of the expansion history of the Universe are so important. Constraining the low-redshift evolution of the Universe breaks many degeneracies in the posterior from CMB analyses.

In this section we will discuss the main probes that are employed to constrain the low-redshift expansion history of the Universe. These measurements can also be employed to infer the values of H_0. However, they extend up to $z \sim 3$, and therefore the resulting H_0 constrain depends on the cosmological model assumed. Nonetheless, agnostic parametrizations of $H(z)$ can be used to minimize the impact of the dependence on the cosmological model, marginalizing over the parametric form of $H(z)$ (or, in some cases, $w(z)$ or $\rho(z)$) to obtain a data-driven, robust constraint on H_0, if slightly weaker. One of the first attempts can be consulted in Ref. (18).

5.2.1 Cosmological supernovae type Ia

Measurements of SNeIa extend up to $z \sim 2$, redshifts significantly higher than the ones that have been traditionally used to constrain H_0 in the distance ladder (which usually are limited to $z \lesssim 0.15$). Therefore, using Eq. (5.4) we can obtain the distance moduli up to very high redshift. This sample is usually known as cosmological supernovae, and have been traditionally analyzed separately than the local supernovae (until very recently (11)). Nonetheless, given the significantly larger volume, the sample of cosmological SNeIa is significantly larger.

Without any calibrator, M_B becomes a (free) nuisance parameter to be marginalized over. Given the large sample of SNeIa, the sampling in redshift is very complete. Therefore, SNeIa are very powerful to constrain the shape of H_0, if not its amplitude (which is completely degenerate with M_B). Since the physical parameter of importance to SNeIa is the absolute magnitude M_B rather than H_0, it has been advocated to use the local determination of M_B from e.g., SH0ES as a prior to calibrate the absolute magnitude of SNeIa and provide a normalization to the constraints on $H(z)$ (19).

5.2.2 Baryon acoustic oscillations

As we discussed, baryon acoustic oscillations (BAO) appear due to the primordial sound waves propagating in the tightly coupled photon-baryon plasma in the early Universe until recombination. The acoustic waves freeze after recombination, but the density contrast that produce get imprinted in the baryon (and therefore total matter) distribution. Therefore, the oscillations that can be measured in the CMB power spectra are also imprinted in the matter (and subsequently the galaxy) distribution at low redshift, although with lower significance due to the small abundance of baryons and the fact that since recombination, baryons have fallen into dark matter potential wells. First detected in the galaxy power spectrum around fifteen years ago, BAO have been robustly measured in galaxy, quasar, and Lyman-α forest density distributions reaching percent-level precision. The BAO features are characterized by a physical scale: the sound horizon at radiation drag, $r_{\rm d}$, which is known as a standard ruler, and can therefore be used as distance calibrator.[2]

Observations measure the positions of different tracers of matter in terms of redshifts and angular positions on the sky, which must then be transformed to obtain three-dimensional clustering summary statistics (e.g., the correlation function or power spectrum) as a function of spatial distances or the corresponding Fourier mode wave numbers. Given an angular separation θ and a

[2]Radiation drag takes place slightly after recombination, but the difference in redshift is very small and, for the precision that we will consider in these discussions, we can take as the same sound horizon. Of course, Boltzmann codes and parameter inference studies do not make this approximation.

small redshift separation δz, the spatial comoving distance in the transverse direction and along the line of sight are

$$r_\perp = D_M(z)\theta, \qquad r_\parallel = \frac{\delta z}{H(z)}, \tag{5.5}$$

respectively. There are three main effects that alter these components of the observed distances: redshift-space distortions,[3] the Alcock-Paczynski effect, and the isotropic dilation. Redshift-space distortions are a physical modification to $r_\parallel$, due to the peculiar velocities of galaxies changing the redshift of observed sources along the line of sight (hence changing their position in redshift space with respect to the real space). From the cosmological principle, the clustering of biased tracers is isotropic, but these introduce anisotropies in the *observed* clustering.

Assuming a background expansion history (obtained from a *fiducial* cosmology) that differs from that of the *true* expansion rate of the Universe causes an artificial distance distortion. The fiducial cosmology is used to compute D_M and H in Eq. (5.5); therefore, the recovered $r_\perp$ and $r_\parallel$ differ from the true distances. This distortion affects $r_\perp$ and $r_\parallel$ in different ways, so it is possible to decompose it into an isotropic and an anistropic component: the isotropic dilation and the Alcock-Paczynski effect.

The Alcock-Paczynski effect and the isotropic dilation can be modeled by rescaling factors, obtained when comparing the *observed* distances, which assume the fiducial cosmology, and *true* distances: $r^{\rm true}_{\perp,\parallel} = r^{\rm obs}_{\perp,\parallel} q_{\perp,\parallel}$ (or $k^{\rm true}_{\perp,\parallel} = k^{\rm obs}_{\perp,\parallel}/q_{\perp,\parallel}$ in Fourier space). Using Eq. (5.5), the rescaling parameters are

$$q_\perp = \frac{D_M(z)}{\left(D_M(z)\right)^{\rm fid}}, \qquad q_\parallel = \frac{\left(H(z)\right)^{\rm fid}}{H(z)}\,. \tag{5.6}$$

Using $q_\perp$ and $q_\parallel$, the isotropic dilation corresponds to $\left(q_\perp^2 q_\parallel\right)^{1/3}$, and the Alcock-Paczinski effect is given by the ratio of $q_\perp$ and $q_\parallel$.

The Alcock-Paczysnki effect and the isotropic dilation are always present in the measurement of the clustering statistics as function of distance or wavenumber: it is inherent to any measurement that depends on distance scales. Nonetheless, the BAO feature is clearly distinguishable against the broadband of the summary statistic; it manifests as oscillations in Fourier space or a peak in configuration space, and large-scale clustering measurements have well-determined its location.[4]

To extract the BAO scale from the observed target summary statistic, standard BAO analyses employ a pre-computed template of the target summary

[3]We will not discuss in detail redshift space distortions in these lectures, but refer the interested reader to the excellent review of Ref. (20)

[4]The turn over of the power spectrum is the other clear feature in the matter power spectrum, but it is located at too large scales, beyond the reach of current galaxy surveys.

statistic generated assuming a given cosmology. Using the template allows for the extraction of $r_{\rm d}$, which is the only characteristic scale of matter clustering at low reshifts accessible by current experiments. The BAO scale of the template might not match the true BAO scale; therefore, a correction on $r_{\rm d}$ must be included when rescaling distances in order to fit the observed BAO feature with the template. The correction is isotropic, and the rescaling of distances becomes $r^{\rm th}_{\perp,\parallel} = r^{\rm obs}_{\perp,\parallel}\alpha_{\perp,\parallel}$ (or $k^{\rm th}_{\perp,\parallel} = k^{\rm obs}_{\perp,\parallel}/\alpha_{\perp,\parallel}$), where

$$\alpha_\perp = q_\perp \frac{(r_{\rm d})^{\rm fid}}{r_{\rm d}}, \qquad \alpha_\parallel = q_\parallel \frac{(r_{\rm d})^{\rm fid}}{r_{\rm d}} \tag{5.7}$$

provide a mapping between the observed distances (or wave numbers) and those which enter our theoretical modeling, denoted by 'th'.[5] It is important to notice that the rescaling of $r_{\rm d}$ in Eq. (5.7) is not related to the Alcock-Paczynski effect or the isotropic dilation. Hence, the rescaling between observed distances and those entering our theoretical model introduced in Eq. (5.7) is the combination of two non-physical effects: the redshift-distance transformation and the ratio between the fiducial (for the fixed template) and true $r_{\rm d}$ values. The total rescaling parameters become

$$\alpha_\perp = \frac{D_M(z)/r_{\rm d}}{(D_M(z)/r_{\rm d})^{\rm fid}}, \qquad \alpha_\parallel = \frac{(H(z)r_{\rm d})^{\rm fid}}{H(z)r_{\rm d}}, \tag{5.8}$$

where 'fid' corresponds to the fiducial cosmology that has been used to both translate redshifts into distances and compute the fixed template.

In order to avoid biasing the information obtained from the BAO feature, the shape and amplitude of the broadband are marginalized over with the introduction of nuisance parameters. After marginalization, the only remaining cosmological information in the clustering statistics is related to the BAO location and anisotropy, which is mostly encoded in the rescaling parameters. This is an entirely geometric fit to the observations; hence, it has the potential to be performed without being limited to any cosmological model without loss of generality. Specifically, the rescaling parameters are the fit parameters, and the resulting constraints are traditionally used in global analyses to infer cosmological parameters of any cosmological model.

As an example of how the BAO scale (i.e., $r_{\rm d}$) is measured from galaxy surveys, we will comment the methodology employed consider a template-based analysis of the power spectrum. The power spectrum $P(\boldsymbol{k})$ and the correlation function $\xi(\boldsymbol{s})$ are equivalent estimators for two-point clustering statistics in Fourier and configuration space, respectively, where $\boldsymbol{s}$ is the redshift space distance and $\boldsymbol{k}$ is the associated wave number. The Legendre multipoles of

[5]There are alternative parametrizations of these rescalings (or those in Eq. (5.6)), obtained through combinations of $\alpha_\perp$ and $\alpha_\parallel$. Some examples focus on the isotropic and anisotropic distortions (α, ϵ) or on the monopole and the μ^2 moment of the two-point statistics (α_0, α_2).

the power spectrum are given by

$$P_\ell(k) = \frac{2\ell+1}{2}\int_{-1}^{1} \mathrm{d}\mu P(k,\mu)\mathcal{L}_\ell(\mu), \tag{5.9}$$

where $k \equiv |\boldsymbol{k}|$ is the module of the wave number vector and μ is the cosine of the angle between the wave number vector and the line of sight. The Legendre multipoles of the correlation function, ξ_ℓ, are defined in an analogous way, and related with P_ℓ by the Fourier transform via

$$\xi_\ell(s) = i^\ell \int \frac{k^3 \mathrm{d}\log k}{2\pi^2} P_\ell(k) j_\ell(ks), \tag{5.10}$$

where $s \equiv |\boldsymbol{s}|$ is the module of the redshift space distance and j_ℓ is the ℓ-th order spherical Bessel function. Note that Eq. (5.10) equally holds for real space distances and wave numbers. In this section, we do not explicitly include the dependence on redshift, present in practically all quantities, for the sake of brevity and readability; we do, however, show the dependence on k and μ, for clarity.

The standard BAO analysis is based on fitting a template (pre-computed under a fiducial cosmology) to the observations. This template is built in such a way that the BAO feature is identifiable and isolated. In order to isolate the BAO feature, the linear matter power spectrum P_m is decomposed into a smooth component $P_\mathrm{m,sm}$ (i.e., the broadband, with no contribution from the BAO) and an oscillatory contribution O_lin. In this way, the total matter power spectrum is given by $P_\mathrm{m}(k) = P_\mathrm{m,sm}(k)O_\mathrm{lin}(k)$.

The galaxy bias b_g (that relates linear matter and galaxy perturbations) and a factor encoding the effect of redshift-space distortions F_RSD can be applied to $P_\mathrm{m,sm}$ in order to obtain the anisotropic, smoothed galaxy power spectrum in redshift space

$$P_\mathrm{sm}(k,\mu) = BF^2_\mathrm{RSD}(k,\mu)P_\mathrm{m,sm}(k), \tag{5.11}$$

where B is a constant absorbing b_g and potential variations on the amplitude of $P_\mathrm{m,sm}$, and

$$F_\mathrm{RSD}(k,\mu) = \left(1+\beta\mu^2 R\right)\frac{1}{1+0.5\left(k\mu\sigma_\mathrm{FoG}\right)^2}, \tag{5.12}$$

where $\beta = f/b_g$, and the fingers of God small-scale suppression is driven by the parameter σ_FoG, whose value is related to the halo velocity dispersion.[6]

The actual amplitude of the BAO feature is reduced with respect to the linear prediction due to non-linear collapse. In addition, non-linear clustering also introduces a sub-percent shift in the BAO scale. However, these effects

[6]The damping due to the fingers of God can also be modeled with a Gaussian function, providing similar results without losing flexibility in the fit to the observations.

can be partially reverted using density field reconstruction. The factor R in Eq. (5.12) models the partial removal of redshift-space distortions produced by the density field reconstruction and takes the following values: $R=1$ before reconstruction and $R=1-\exp\left[-\left(k\Sigma_{\rm recon}\right)^2/2\right]$ after reconstruction.[7] On the other hand, the non-linear damping of the BAO is modeled with an exponential suppression applied to $O_{\rm lin}$. The damping affects the transverse and line-of-sight directions differently; hence, we introduce two separate scales $\Sigma_\perp$ and $\Sigma_\parallel$, respectively.

The final anisotropic galaxy power spectrum, accounting for the effect of non-linearities on the BAO features and eventual density field reconstruction, can be expressed as

$$\begin{aligned} P(k,\mu) = {} & P_{\rm sm}(k,\mu)\times \\ & \times\left[1+\left(O_{\rm lin}(k)-1\right)e^{-\frac{k^2}{2}\left\{\mu^2\Sigma_\parallel^2+\left(1-\mu^2\right)\Sigma_\perp^2\right\}}\right]+ \\ & +P_{\rm shot}\,, \end{aligned} \tag{5.13}$$

where $P_{\rm shot}=n_{\rm g}^{-1}$ (where $n_{\rm g}$ is the mean comoving number density of galaxies) is a scale-independent contribution arising from the fact that we use discrete tracers of the matter density field, such as galaxies. The template for the BAO analysis is generated with Eq. (5.13).

As shown in Eq. (5.13), it is clearer to express the anisotropic power spectrum as function of k and μ, instead of $k_\perp$ and $k_\parallel$. The rescaling of distances appearing in Eq. (5.8) can be transformed to k and μ as

$$\begin{aligned} k^{\rm true} &= \frac{k^{\rm obs}}{\alpha_\perp}\left[1+\left(\mu^{\rm obs}\right)^2\left(F_{\rm AP}^{-2}-1\right)\right]^{1/2}, \\ \mu^{\rm true} &= \frac{\mu^{\rm obs}}{F_{\rm AP}}\left[1+\left(\mu^{\rm obs}\right)^2\left(F_{\rm AP}^{-2}-1\right)\right]^{-1/2}, \end{aligned} \tag{5.14}$$

where $F_{\rm AP}\equiv\alpha_\parallel/\alpha_\perp$.

Given the large scales probed, the line of sight changes with each pointing and cannot be considered parallel to any Cartesian axis. Hence, it is not possible to obtain a well-defined μ for the observations, which makes a direct measurement $P(k,\mu)$ impossible. However, one can directly measure the Legendre multipoles of the anisotropic power spectrum.[8] Then, the *observed* power spectrum multipoles are modeled as

$$\begin{aligned} P_\ell(k^{\rm obs}) = {} & \frac{2\ell+1}{2\alpha_\perp^2\alpha_\parallel}\times \\ & \times\int_{-1}^{1}{\rm d}\mu^{\rm obs}P(k^{\rm true},\mu^{\rm true})\mathcal{L}_\ell(\mu^{\rm obs})+A_\ell(k), \end{aligned} \tag{5.15}$$

[7] The specific functional form of R after reconstruction depends on the reconstruction formalism used.
[8] There are other compression options, such as the so-called angular wedges.

where $\mathcal{L}_\ell$ is the Legendre polynomial of degree ℓ, $P(k^{\rm true}, \mu^{\rm true})$ is computed using Eqs. (5.13) and (5.14), and a $\left(r_{\rm d}^{\rm fid}/r_{\rm d}\right)^3$ term has been absorbed into the constant factor B of $P_{\rm sm}$ in Eq. (5.11). Different polynomials $A_\ell(k)$ are added to each one of the power spectrum multipoles. These polynomials are added not only to marginalize over uncertainties related with non-linear clustering, but in particular to account for the possibility that the broadband of the template $P_{\rm m,sm}$ does not match the actual one. These polynomials have the form

$$A_\ell(k) = a_{\ell,1}k^{-3} + a_{\ell,2}k^{-2} + a_{\ell,3}k^{-1} + a_{\ell,4} + a_{\ell,5}k^n, \tag{5.16}$$

where $n = 1$ and $n = 2$ before and after density field reconstruction, respectively.

In summary, BAO-only analyses include the following parameters:

$$\left\{\alpha_\perp, \alpha_\parallel, B, \beta, \boldsymbol{a}_\ell, \sigma_{\rm FoG}, \Sigma_\perp, \Sigma_\parallel\right\}, \tag{5.17}$$

where $\boldsymbol{a}_\ell$ are the coefficients of A_ℓ in Eq. (5.16). All but the two first parameters $\alpha_\perp$ and $\alpha_\parallel$ are nuisance parameters.

This procedure has been proven to be extremely robust and flexible for models predicting different expansion rates at late times, and it successfully models changes in $r_{\rm d}$ due to early-time modifications of the cosmological model, as well as other contributions to the BAO feature that are not captured by these rescaling and nuisance parameters, such as phase shifts (which can be scale-dependent) or a different scale dependence of the amplitude of the oscillations.

As evident from Eq. (5.6), the only cosmological information the Alcock-Paczynski effect and the isotropic dilation are sensitive to is the late-time expansion rate. By utilizing a fixed template in the analysis, BAO measurements are also sensitive to pre-recombination physics through $r_{\rm d}$ (Eq. (5.8)) in an agnostic and independent way, incorporating information about both the expansion rate and the growth of matter perturbations. While the isotropic dilation is completely degenerate with $r_{\rm d}$, the Alcock-Paczynski effect (modeled by the ratio of $\alpha_\perp$ and $\alpha_\parallel$ when a fixed template is used) is independent of the BAO scale. Since $\alpha_\perp/\alpha_\parallel = D_M H/\left(D_M H\right)^{\rm fid}$ does not depend on H_0, the Alcock-Paczynski effect constraints the unnormalized expansion history of the Universe, independently of the BAO scale.

Contrarily to SNeIa, BAO measurements are very sparse in redshift. This causes that BAO are not very powerful to constrain the shape of $H(z)$. However, BAO provide an absolute calibration of the distance-redshift relation, in terms of the product of $r_d h$ (21). This results in very strong constraints on $r_d h$, which is common to all BAO measurements. The complementarity between SNeIa and BAO makes that the combination of only these two cosmological probes, the shape of $H(z)$ is constrained to be that of ΛCDM as inferred by *Planck* with more than 5% precision up to $z \sim 2$, as shown in

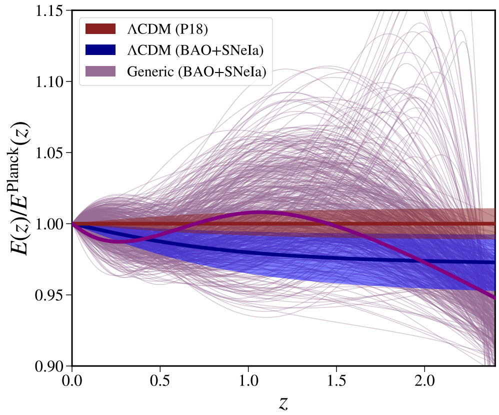

Figure 5.3 Best-fit evolution of the expansion history $E(z) \equiv H(z)/H_0$ normalized by *Planck*'s ΛCDM fit and 68% confidence level uncertainties (shaded regions, thin lines). In purple, the reconstruction from BAO+SNeIa assuming a generic expansion using flexknot splines. Figure from Ref. (22).

Fig. 5.3, with 1% constraints on $r_d h$, even marginalizing over the curvature of the Universe.

5.2.2.1 The importance of the sound horizon at radiation drag

Summarizing, the combination of BAO and SNeIa alone is powerful enough to constrain the shape of $H(z)$ and the amplitude as function of $r_\mathrm{d} h$. Then there are two potential possibilities to build the cosmic distance ladder from these measurements, depending on the calibrator, or anchor for each measurement. One possibility is to use local measurements of H_0 (or M_B) to calibrate the cosmological SNeIa and therefore normalize the $H(z)$ measurements from BAO+SNeIa. By doing that, we automatically obtain a determination of r_d. This is what is known as the *direct* distance ladder.

On the other hand, we can build the cosmic distance ladder in the opposite direction. We can impose a prior on r_d (from either CMB anisotropies or using a prior on Ω_b from standard Big Bang Nucleosynthesis and measurements of pristine gas clouds to determine r_d[9]. In both cases, this prior on r_d depends on the cosmological model assumed to model pre-recombination physics. This prior can be used to *calibrate* the BAO measurements that provide in turn a normalization to the cosmological SNeIa. This way it is possible to obtain an inference of H_0, using what is called the *inverse* distance ladder.

[9]This approach is independent on CMB anisotropies measurements. The only dependence that it has in CMB measurements comes from the determination of the redshift of reionization from FIRAS.

Of course, in a consensus cosmological model, if all experiments are consistent between them, the inverse and the direct distance ladder must return consistent results. However, this is not the case. Therefore, the H_0 tension can be reframed as a mismatch between the two anchors (r_d and H_0) of the cosmic distance ladder (23). With current measurements, $r_\mathrm{d}h \sim 100$ Mpc/h and $r_\mathrm{d} \sim 147$ Mpc as inferred from the *Planck* assuming ΛCDM (or BBN). Thus, H_0 from the inverse distance ladder would return $H_0 \sim 68$ km/s/Mpc, while the direct distance ladder returns $r_\mathrm{d} \sim 137$ Mpc.

The combination of BAO+SNeIa gives us a couple of hints about potential solutions to the H_0 tension. On the one hand, SNeIa *fix* $H(z)/H_0$ to be very similar to ΛCDM, which severely constraints modifications to the expansion history of the Universe at late times. On the other, BAO constraints on the product of $r_\mathrm{d}h$ forces to reduce the predicted value of r_d to reconcile all the measurements. This indicates that any solution to the H_0 tension must *at least* include modifications before recombination to reduce the sound horizon.

5.2.3 Strong-lensing time delays

Strong gravitational lenses can be used for cosmography using the time delay in the reception of a given signal through different paths (when there are multiple lensed signals). Here we will not discuss this method in detail, but a review can be found in Ref. (24). From Fermat's principle of least time, the light travel time through a gravitational lens is

$$t(\boldsymbol{\theta}) = D_{\Delta t}\Phi_L(\boldsymbol{\theta};\boldsymbol{\beta})\,, \qquad \Phi_L = \frac{(\boldsymbol{\theta}-\boldsymbol{\beta})^2}{2} - \phi_L(\boldsymbol{\theta})\,, \tag{5.18}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are the apparent (lensed) and true sky position of the source, Φ_L is the Fermat potential and ϕ_L is the scaled, projected gravitational potential along the lens sight-line, which receives contributions from the main lens and any matter overdensity along the line of sight. For multiple images, the time delay for each path is given by

$$\Delta t = D_{\Delta t}\Delta\Phi_L, \tag{5.19}$$

where the difference in the Fermat potential for each image can be predicted along with the deflection angle for any given mass model for the lens. Therefore, the time delay only depends on the time-delay distance $D_{\Delta t}$, defined in terms of the (physical) angular diameter distance to the deflector, to the source, and between the deflector and the source (denoted with subscripts 'd', 's', and 'ds', respectively) as

$$D_{\Delta t} \equiv (1+z_d)\frac{D_{A,d}D_{A,s}}{D_{A,ds}}\,. \tag{5.20}$$

Measuring the time delays requires dedicated monitoring of very special systems: a bright high-redshift quasar must be located behind a strong lens in a position such that multiple images are created. Practical implementation requires dedicated measurements with high cadence to detect the often small variations in the quasar brightness, used as features to determine the time delay between the images.

Note that, for a perfect mass model, strong-lensing time delays return absolute measurements of the time-delay distance, which is proportional to H_0^{-1} (and also moderately sensitive to other cosmological parameters controlling the expansion history of the Universe such as Ω_m or w). However, modeling the mass distribution of the main deflector, its neighbors, and accounting for potential small contributions along the line of sight is far from trivial. External features are used, such as Einstein ring, kinematics of the stars with spectroscopic observations, the properties of the neighbor galaxies to determine their position along the line of sight, etc. Even with this wealth of observations, uncertainties due, among other, to the mass-sheet degeneracy (i.e., the lensing system is unchanged in presence of a uniform convergence field, or a mass sheet), limits the power of this method and requires further observations to break this degeneracy.

CHAPTER 6

LECTURE 4: MEASURING THE AMPLITUDE OF CLUSTERING

The second main tension between cosmological measurements under the assumption of ΛCDM involves the amplitude of clustering on small scales. In particular, it refers to the matter clustering in the late Universe, parameterized with the combination of $S_8 \equiv \sigma_8(\Omega_m/0.3)^{1/2}$, where σ_8 is the root mean square of the matter perturbations smoothed over 8 Mpc$/h$.[1] Simply stated, the distribution of matter in hte late Universe as measured by low-redshift probes is smoother than expected from the evolution of the fluctuations observed in the CMB assuming ΛCDM. The largest tension ($\sim 3\sigma$) involves galaxy weak lensing, but similar trends can be found in measurements of the CMB lensing tomography, galaxy weak lensing, abundance of galaxy clusters, and cross-correlations of the above.

Although they are arguably more complicated, measurements based on lensing *directly* probe the matter distribution, rather than the distribution of a biased tracer. In general, dropping the complications related with the

[1] As we will see, this is the parameter combination best constrained by weak lensing measurements from galaxy surveys.

bias makes significantly easier to obtain accurate predictions. In this chapter we will discuss the basics of matter clustering measurements based on gravitational lensing. In particular, we will cover CMB lensing, CMB lensing tomography in cross-correlations with tracers of the large-scale structure, and galaxy weak lensing.

Gravitational potentials distort the paths of light. Therefore, we can investigate the matter distribution through its impact in the light that reaches us, which makes lensing a very powerful probe of the large-scale structure. Gravitational lensing is also a very powerful probe of the mass distribution of specific collapsed objects (thus a probe of dark matter) and can also be used for cosmography, as discussed in the previous chapter. These phenomena can be classified as microlensing and/or strong lensing. In turn, the most important aspect of gravitational lensing to probe the large-scale structure is weak lensing, which slightly distorts the shape of distant galaxies. Thanks to the study of the statistical distortion of distant galaxies as function of position on the sky it is possible to make mass maps, and study the clustering of matter.

6.1 Basics of weak gravitational lensing

We can describe the effect of gravitational lensing on the observed photons using the Boltzmann equation (see Sections 1.2 and 1.3.4). At low redshift, we can safely neglect any collision term for the photons, since we can ignore absorption and scatter, so that $\mathrm{d}f/\mathrm{d}t = 0$. Gravitational lensing affects both the intensity and polarization of photons (e.g., the famous conversion from E modes to B modes for the CMB polarization due to lensing along the line of sight), but we will limit our study to their intensity. Any telescope, at the end of the day, measures the integral of the specific intensity I_ν, which is the energy incident on a detector per solid angle, per unit area and time and frequency:

$$\mathrm{d}E = I_\nu \mathrm{d}\Omega \mathrm{d}A_\perp \mathrm{d}t \mathrm{d}\nu \,, \tag{6.1}$$

where $\mathrm{d}A_\perp$ is the detector area normal to the photon flux. Within a time interval $\mathrm{d}t$, the detector collects photons from a volume $\mathrm{d}^3x = \mathrm{d}A_\perp \mathrm{d}t$. On the other hand, in natural units, the photon energy and its frequency can be related through the photon momentum as $E = p = 2\pi\nu$, such as $\mathrm{d}^3\boldsymbol{p} = (2\pi)^3\nu^2\mathrm{d}\nu\mathrm{d}\Omega$. Then, since the differential photon energy that arrives to the detector is the number of photons $\mathrm{d}N = 2f\mathrm{d}^3\boldsymbol{x}\mathrm{d}^3\boldsymbol{p}/(2\pi)^3$ weighted by an energy factor $2\pi\nu$, we have

$$I_\nu(\boldsymbol{x}, \boldsymbol{p}, t) = 4\pi\nu^3 f(\boldsymbol{x}, p = 2\pi\nu, \hat{\boldsymbol{p}}, t) \,. \tag{6.2}$$

Weak lensing conserves the phase-space distribution function but changes the photon paths. Thus, the specific intensity is conserved (once the frequency variation due to redshift is taken into account, $\nu = \nu_{\rm emission}/(1+z)$), and the only change between the emitted specific intensity $I_\nu^{\rm true}$ and its observed value $I_\nu^{\rm obs}$ is related with the line of sight, namely

$$I_\nu^{\rm obs}(\boldsymbol{\theta}) = I_\nu^{\rm true}(\boldsymbol{\theta}_S)\,, \tag{6.3}$$

i.e., the observed specific intensity at position $\boldsymbol{\theta}$ on the sky is the as same as would have been observed from the direction of the true source position $\boldsymbol{\theta}_S$ in the absence of lensing.

We want to derive the relation between $\boldsymbol{\theta}_S$ and $\boldsymbol{\theta}$ as function of the lens. Since we are in the regime of weak lensing, the deflection angle is going to be small and therefore we can limit our derivation to linear order. In this limit, a source at a distance χ and position $\boldsymbol{\theta}_S$ on the sky has a position $\boldsymbol{x}_{\rm true}$ in a 3D coordinate system; in turn, the lensed image (which has shifted position $\boldsymbol{\theta}$ on the sky) is located at $\boldsymbol{x}_{\rm obs}$. These two positions are given by

$$\boldsymbol{x}_{\rm true} = (x_\perp^{\rm true}, x_\parallel^{\rm true}) = \chi(\boldsymbol{\theta}_S, 1)\,, \qquad \boldsymbol{x}_{\rm obs} = (x_\perp^{\rm obs}, x_\parallel^{\rm obs}) = \chi(\boldsymbol{\theta}, 1)\,, \tag{6.4}$$

where the radial distance between the observer and the unlensed and lensed images is the same at linear order. $x_\perp^i$ (the position in the 3D coordinate frame perpendicular along the axis i to the photon path) is the integral of

$$\frac{{\rm d}x_\perp^i}{{\rm d}\chi} = -\frac{{\rm d}x_\perp^i}{{\rm d}\tau} = -a\frac{{\rm d}x_\perp^i}{{\rm d}t} = -\hat{p}_\perp^i \tag{6.5}$$

over the comoving distance, where we have used ${\rm d}\chi = -{\rm d}\tau = -{\rm d}t/a$, i.e., that the conformal time is the same as the comoving distance but with opposite sign (going outward in distance is backwards in time). The last equality uses Eq. (1.19). Integrating the equation above we can obtain θ_S^i:

$$\theta_S^i = \frac{x_\perp^i}{\chi} = -\frac{1}{\chi}\int_0^\chi {\rm d}\tilde{\chi}\,\hat{p}_\perp^i(\tilde{\chi})\,. \tag{6.6}$$

Similarly, we can obtain the derivatives of the direction of the momentum with respect to time for a inhomogenous Universe from the geodesic equation. Using the results in Eq. (1.73) and acknowledging that ${\rm d}\hat{p}/{\rm d}t = (1/p){\rm d}p^i/{\rm d}t - (p^i/p^2){\rm d}p/{\rm d}t$, we find

$$\frac{{\rm d}\hat{p}}{{\rm d}t} = \frac{1}{a}\left[\delta^{ij} - \hat{p}^i\hat{p}^j\right]k_j(\Phi - \Psi)\,, \tag{6.7}$$

where the factor in square brackets is the projection on directions transverse to the momentum (the z axis in our small deflection approximation). Thus,

$$\frac{{\rm d}\hat{p}_\perp^i}{{\rm d}\chi} = -\frac{{\rm d}\hat{p}_\perp^i}{{\rm d}\tau} = -a\frac{{\rm d}\hat{p}_\perp^i}{{\rm d}t} = -\partial_i(\Phi - \Psi) = -2\partial_i(\Phi)\,, \tag{6.8}$$

assuming that there is no anisotropic stress for the last equality. Integrating the previous equation we find

$$\hat{p}_\perp^i(\tilde{\chi}) = -2\int_0^{\tilde{\chi}} {\rm d}\tilde{\chi}'\partial_i\Phi(\boldsymbol{x}(\boldsymbol{\theta}, \tilde{\chi}'), \tau_0 - \tilde{\chi}') + C_i\,, \tag{6.9}$$

where $\boldsymbol{x}$ is the unperturbed photon path at which the potential is evaluated. We can plug this equation in Eq. (6.6), and consider the limit of no deflection to obtain that $C_i = -\theta^i$. Therefore, we have

$$\theta^i_S = \theta^i + \frac{2}{\chi}\int_0^\chi \mathrm{d}\tilde{\chi}\int_0^{\tilde{\chi}} \mathrm{d}\tilde{\chi}'\partial_i\Phi(\boldsymbol{x}(\boldsymbol{\theta},\tilde{\chi}'),\tau_0-\tilde{\chi}')\,. \tag{6.10}$$

Therefore, if an overdensity $\Phi > 0$ is found along the direction, $\boldsymbol{x}_\perp = 0$, from the previous equation we get $\partial_i\Phi < 0$ for $x > 0$, and the bending angle for light passing through an overdensity on the positive around $x > 0$ is negative, i.e., inwards towards the overdensity, as expected.

Changing the order of the integrals (the now second integral ranging from $\tilde{\chi}'$ to χ) leaves one of them yielding $(1-\tilde{\chi}'/\chi)$. Let us rename the variables so that

$$\theta^i_S = \theta^i + \Delta\theta^i = \theta^i + 2\int_0^\chi \mathrm{d}\tilde{\chi}\partial_i\Phi(\boldsymbol{x}(\boldsymbol{\theta},\tilde{\chi}),\tau_0-\tilde{\chi})\left(1-\frac{\tilde{\chi}}{\chi}\right)\,. \tag{6.11}$$

Using $\partial_i = \partial_{\theta^i}/\tilde{\chi}$, we can write the deflection angle as the derivative of a lensing potential ϕ_L in the transverse plane on the sky:

$$\theta^i_S = \partial_{\theta^i}\phi_L = \partial_{\theta^i}\left[2\int_0^\chi \frac{\mathrm{d}\tilde{\chi}}{\tilde{\chi}}\Phi(\boldsymbol{x}(\boldsymbol{\theta},\tilde{\chi}))\left(1-\frac{\tilde{\chi}}{\chi}\right)\right]\,. \tag{6.12}$$

Summarizing, the lensing potential is a weighted integral over 2Φ along the photon path, which at linear order can be taken to be the unperturbed path. The contribution of the lenses close to the source are suppressed by the $(1-\tilde{\chi}/\chi)$ factor.

The lensing potential is the key quantity for CMB lensing. In galaxy surveys instead, the effect of weak lensing extracted from the observations is related to image distortions: since the intrinsic position of the galaxy is unknown, the overall shift of the image does not contain information; instead, since different points of the same (resolved) galaxy are subject to different deflection angles, we can statistically exploit the image distortions due to lensing. Therefore, we need to use also the first derivative of the deflection angle, or the second-derivative matrix –distortion tensor– of the lensing potential:

$$\psi_{ij} \equiv \partial_{\theta^i}\partial_{\theta^j}\phi_L(\boldsymbol{\theta})\,. \tag{6.13}$$

6.2 CMB lensing

Gravitational lensing applies both to discrete sources and diffuse fields. The latter includes, among other, the CMB and line-intensity mapping observations. Since the effects of lensing in line-intensity mapping is very small and it will be very challenging to detect, we focus on CMB lensing here. A good review for CMB lensing can be found in Ref. (4).

Due to historical conventions, we will use here the CMB temperature rather than specific intensity. In this case, up to second order in the deflection angle,

$$T_{\rm obs}(\boldsymbol{\theta}) = T(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \simeq T(\boldsymbol{\theta}) + \Delta\theta^i \partial_{\theta^i} T(\boldsymbol{\theta}) + \frac{1}{2}\Delta\theta^i \Delta\theta^j \partial^2_{\theta^i\theta^j} T(\boldsymbol{\theta}) \,, \quad (6.14)$$

where by definition $\chi = \chi_*$ is the distance to the last scattering surface and we have dropped the subscript 'true'. Since weak lensing only deflects the angle and does not change the surface brightness, it does not affect the CMB mean temperature. Therefore, the equation above also applies for the temperature perturbation Θ. We can also work in Fourier space (taking the flat sky approximation of the harmonic space), such that derivatives with respect to the angle become multiplications by $i\boldsymbol{\ell}$. Taking the 2D Fourier transform of the expression above we have

$$\begin{aligned}\Theta_{\rm obs}(\boldsymbol{\ell}) = \Theta(\boldsymbol{\ell}) - \int \frac{{\rm d}^2\boldsymbol{\ell}'}{(2\pi)^2} \boldsymbol{\ell}'\,(\boldsymbol{\ell} - \boldsymbol{\ell}')\,\phi_L(\boldsymbol{\ell} - \boldsymbol{\ell}')\Theta(\boldsymbol{\ell}') - \\ - \frac{1}{2}\int \frac{{\rm d}^2\boldsymbol{\ell}'}{(2\pi)^2} \int \frac{{\rm d}^2\boldsymbol{\ell}''}{(2\pi)^2} \boldsymbol{\ell}'(\boldsymbol{\ell}' + \boldsymbol{\ell}'' - \boldsymbol{\ell})\Theta(\boldsymbol{\ell}')\boldsymbol{\ell}'\boldsymbol{\ell}''\phi_L(\boldsymbol{\ell}'')\phi_L^*(\boldsymbol{\ell}' + \boldsymbol{\ell}'' - \boldsymbol{\ell})\,,\end{aligned} \quad (6.15)$$

where we already have integrated over the Dirac deltas coming from the exponential factors of the Fourier transform. We can use the expression above to compute the *observed* temperature power spectrum of the CMB anisotropies. To lowest order in the power spectrum of the lensing potential, defined as

$$\langle \phi_L(\boldsymbol{\ell})\phi_L^*(\boldsymbol{\ell}')\rangle = (2\pi)^2 \delta_D^{(2)}(\boldsymbol{\ell} - \boldsymbol{\ell}') C_\ell^{\phi_L\phi_L} \,, \quad (6.16)$$

we find (using $\phi_L(\boldsymbol{\ell}) = \phi_L^*(-\boldsymbol{\ell})$)

$$C_\ell^{\rm obs} = C_\ell + \int \frac{{\rm d}^2\boldsymbol{\ell}'}{(2\pi)^2} \left[\boldsymbol{\ell}'(\boldsymbol{\ell} - \boldsymbol{\ell}')\right]^2 C^{\phi_L\phi_L}_{|\boldsymbol{\ell}-\boldsymbol{\ell}'|} C_{\ell'} - C_\ell \int \frac{{\rm d}^2\boldsymbol{\ell}}{(2\pi)^2} (\boldsymbol{\ell}\boldsymbol{\ell}')^2 C^{\phi_L\phi_L}_{\ell'} \,. \quad (6.17)$$

The last term is a small damping term on the characteristic scales of the deflection squared, that encodes the smoothing of the temperature anisotropies due to the random lensing deflections. In turn, the second term (which convolutes the lensing potential power spectrum $C^{\phi_L\phi_L}_{|\boldsymbol{\ell}-\boldsymbol{\ell}'|}$ and the intrinsic temperature power spectrum $C_{\ell'}$), effectively smooths the peaks of the temperature power spectrum and adds power at scales smaller than the diffusion scale (by transferring power from large to small scales).

Although the expression above must be taken into account to accurately predict the observed CMB temperature power spectrum, it is far from exhausting the information we can obtain about the matter distribution from the effect of gravitational lensing on the CMB. More than this, gravitational lensing also introduces off-diagonal contributions in the covariance of the CMB (since lensing breaks isotropy, as can be seen through the $\boldsymbol{\ell}$ dependence in

Eq. (6.15), which collects the dependence on the two $\perp, \parallel$ directions as for $\boldsymbol{\theta}$). Therefore, they can be used, along with quadratic estimators, to reconstruct the lensing potential. Taking the off-diagonal covariance up to linear order in ϕ_L,

$$\begin{aligned}\left\langle \Theta^{\rm obs}(\boldsymbol{\ell})\Theta^{{\rm obs},*}(\boldsymbol{\ell}-\boldsymbol{L})\right\rangle &\stackrel{\boldsymbol{L}\neq 0}{=} \int \frac{{\rm d}^2\boldsymbol{\ell}'}{(2\pi)^2}\left[\boldsymbol{\ell}'(\boldsymbol{\ell}-\boldsymbol{\ell}')\phi_L(\boldsymbol{\ell}-\boldsymbol{\ell}')\langle\Theta(\boldsymbol{\ell}')\Theta^*(\boldsymbol{\ell}-\boldsymbol{L})\rangle + \right.\\ &\quad \left. + \boldsymbol{\ell}'(\boldsymbol{\ell}-\boldsymbol{L}-\boldsymbol{\ell}')\phi_L^*(\boldsymbol{\ell}-\boldsymbol{L}-\boldsymbol{\ell}')\langle\Theta(\boldsymbol{\ell})\Theta^*(\boldsymbol{\ell}')\rangle\right] = \\ &= \left[(\boldsymbol{L}-\boldsymbol{\ell})\boldsymbol{L}C_{|\boldsymbol{\ell}-\boldsymbol{L}|} + \boldsymbol{\ell}\boldsymbol{L}C_\ell\right]\phi_L(\boldsymbol{L})\,,\end{aligned} \tag{6.18}$$

where for the last equality we have used the definition of the power spectrum (assuming it diagonal, since it is the intrinsic one). The $L = 0$ mode of the lensing potential is not observable (zero gradient). Starting from the expression above we can derive an estimator for the lensing potential performing a weighted integral of the off-diagonal covariance. We define the quadratic estimator

$$\hat{\phi}_L(\boldsymbol{L}) \equiv \mathcal{N}(\boldsymbol{L})\int \frac{{\rm d}^2\boldsymbol{\ell}}{(2\pi)^2}\Theta^{\rm obs}(\boldsymbol{\ell})\Theta^{{\rm obs},*}(\boldsymbol{\ell}-\boldsymbol{L})g(\boldsymbol{\ell},\boldsymbol{L})\,, \tag{6.19}$$

where $\mathcal{N}$ is a normalization to ensure that estimator is unbiased and g is a weighting function to ensure minimum variance. Enforcing the condition of unbiased estimator $\langle\hat{\phi}_L\rangle = \phi_L$ we find a normalization

$$\mathcal{N}^{-1}(\boldsymbol{L}) = \int {\rm d}^2\boldsymbol{\ell}(2\pi)^2\left[(\boldsymbol{L}-\boldsymbol{\ell})\boldsymbol{L}C_{|\boldsymbol{\ell}-\boldsymbol{L}|} + \boldsymbol{\ell}\boldsymbol{L}C_\ell\right]g(\boldsymbol{\ell},\boldsymbol{L})\,. \tag{6.20}$$

In turn, enforcing minimum variance (at tree level and neglecting non-Gaussian contributions from the connected four-point function), we have

$$\langle\hat{\phi}_L^*(\boldsymbol{L})\hat{\phi}_L(\boldsymbol{L}')\rangle = (2\pi)^2\delta_D^{(2)}(\boldsymbol{L}-\boldsymbol{L}')2\mathcal{N}^2(\boldsymbol{L})\int\frac{{\rm d}^2\boldsymbol{\ell}}{(2\pi)^2}C_\ell^{\rm tot}C_{|\boldsymbol{\ell}-\boldsymbol{L}|}^{\rm tot}g^2(\boldsymbol{\ell},\boldsymbol{L})\,, \tag{6.21}$$

where the total power spectrum includes signal, noise and all foreground contributions; after minimization we find

$$g(\boldsymbol{\ell},\boldsymbol{L}) = \frac{(\boldsymbol{L}-\boldsymbol{\ell})\boldsymbol{L}C_{|\boldsymbol{\ell}-\boldsymbol{L}|} + \boldsymbol{\ell}\boldsymbol{L}C_\ell}{2C_\ell^{\rm tot}C_{|\boldsymbol{\ell}-\boldsymbol{L}|}^{\rm tot}}\,. \tag{6.22}$$

Summarizing, gravitational lensing smooths the CMB power spectrum peaks and boosts the small scale power spectrum by transferring power from large to small scales. More interestingly, from the off-diagonal covariance of the power spectrum we can derive an estimator of the lensing potential to the last-scattering surface. The main effect is the modulation of the anisotropy of the CMB power spectrum due to long-modes of the lensing potential. This has been exploited by CMB experiments to create convergence maps, and to measure the lensing power spectrum, which helps to constrain cosmology.

In this section we have used the flat-sky approximation to ease the derivation (and notation), but the generalization to full sky using spherical harmonics is straightforward. Finally, the angle deflection also changes rotates the linear polarization of the photons, leaking some of the power on the E-modes to the B-modes (which is a foreground to primordial B-mode searches). The quadratic estimator can be extended to also the polarization to improve its performance.

6.2.1 CMB lensing tomography

The lensing converge is given by one half the gradient of the deflecting angle, which results in the Laplacian of the lensing potential. Applying the Poisson equation (note that we need to transform the angular to spatial derivative) we can then express the convergence κ in a position on the sky as

$$\kappa(\boldsymbol{\theta}) = \frac{3\Omega_m H_0^2}{2} \int_0^{\chi} \mathrm{d}\tilde{\chi} \frac{\tilde{\chi}}{a(\tilde{\chi})} \delta_m(\boldsymbol{\theta}\tilde{\chi}, \tau_0 - \tilde{\chi}) \left(1 - \frac{\tilde{\chi}}{\chi}\right) , \tag{6.23}$$

which explicitates the relation between lensing and the matter overdensities. From this expression we can appreciate that lensing directly depends on the large-scale structure. However, the effect on the CMB (and the reconstructed convergence field) is an integrated effect.

We can *isolate* contributions from specific redshift intervals by cross correlating the CMB convergence field with low-redshift tracers of the large-scale structure. For simplicity, let us consider galaxies as the tracers. Considering only linear bias, a galaxy catalog with redshift distribution $\mathrm{d}n/\mathrm{d}z$ provides a biased projected overdensity field

$$\delta_g(\boldsymbol{\theta}) = \int_0^{\infty} \mathrm{d}z b(z) \frac{\mathrm{d}n}{\mathrm{d}z} \delta_m(\chi(z)\boldsymbol{\theta}, z) \, . \tag{6.24}$$

Since both the convergence field and the galaxy catalog trace the same large-scale structure, they are correlated. The angular power spectrum is given by

$$C_\ell^{\kappa g} = \frac{2}{\pi} \int \mathrm{d}k k^2 P_m(k, z=0) W_\ell^{\kappa}(k) W_\ell^{g}(k) \, , \tag{6.25}$$

where P_m is the matter power spectrum (for which we can include nonlinear clustering) and (in our approximation of no anisotropic stress) we have assumed $P_m(z) = D(z)^2 P_m(z=0)$. The kernels for galaxies and convergence are

$$\begin{aligned} W_\ell^g &= \int \mathrm{d}z b(z) \frac{\mathrm{d}n}{\mathrm{d}z} D(z) j_\ell(k\chi(z)) \, , \\ W_\ell^\kappa &= \frac{3\Omega_m H_0^2}{2} \int \mathrm{d}z \frac{\chi_* - \chi(z)}{\chi(z)\chi_*} D(z) j_\ell(k\chi(z)) \, . \end{aligned} \tag{6.26}$$

Physically, it is clear that the only contribution from the convergence field that is correlated with the galaxy distribution is that one that receives contributions from the redshift range that is also covered by the galaxy catalog. Mathematically, that comes from the properties of the Bessel functions, which to zero-th order can be understood as *orthogonal* on their arguments.

Therefore, as anticipated, cross-correlating the CMB-lensing convergence field with low-redshift tracers of the large-scale structure allows to probe specific redshift ranges in the lensing potential of the CMB. Combining the cross-power spectrum with the auto power spectrum of galaxies provides a better control of the galaxy bias, as well as increasing the overall signal-to-noise ratio of the measurement.

6.3 Galaxy weak lensing

We switch now to how weak gravitational lensing affects galaxies. As with any other object, lensing shifts the position of the image, but we cannot exploit this effect because we do not know the true position of the galaxy. We can however exploit the relative displacement of different parts of the galaxy image, i.e., how it is distorted due to the lensing. This is because galaxies are not points on the sky, but small, extended objects. Therefore, each of its points is perturbed by a slightly different deflection angle . The simplest case is the distortion of a circle into an ellipse. However as we will see below, this is an extremely ideal case and many complications play a role in the measurement of galaxy weak lensing.

We first need a quantitative description of the shape of the galaxy; the simplest measure are the moments of its image. For an image centered at the origin, the second moments are

$$q_{ij} \equiv \langle \theta_i \theta_j \rangle_{I_{\rm obs}} \equiv \frac{1}{F} \int \mathrm{d}^2\boldsymbol{\theta} I_{\rm obs}(\boldsymbol{\theta}) \theta_i \theta_j \,, \tag{6.27}$$

where the brackets denote the intensity-weighted average over the image and the second moments are normalized by the flux $F = \int \mathrm{d}^2\boldsymbol{\theta} I_{\rm obs}(\boldsymbol{\theta})$, which is the total angular integral of the intensity. With this choice of coordinates, the first moments of the image are null by definition. In turn, q_{ij} is 2×2 matrix that we can write as

$$q_{ij} = \frac{1}{2} q \begin{pmatrix} 1 + \epsilon_1 & \epsilon_2 \\ \epsilon_2 & 1 - \epsilon_1 \end{pmatrix} , \tag{6.28}$$

which is determined by three independent components: the trace q and the ellipcities in the two orthogonal directions ϵ_1 and ϵ_2. For circular images, ellipcities are $\epsilon_i = 0$ and $\sqrt{q}$ provides a measure of the angular size of the image. This formulation is very similar to the polarization tensor, for the intensity I and the linear polarization Stokes parameters Q and U. Therefore, we can define E and B modes in a similar way.

We therefore need to describe how lensing affects the shape tensor q_{ij}. Since we know that the effect is a deflection angle that varies across the galaxy image, we need to derive the observed position of all the galaxy points with respect to their observed angles. This corresponds to the derivative of the gradient of the lensing potential and the antisymmetric part of this transformation, which corresponds to the rotation of the image, vanishes at linear order. We will therefore consider a symmetric transformation matrix. Furthermore, the transformation that affects q (changing the size of the image) corresponds to the convergence κ defined in the previous section. Finally, the transformation affecting the diagonal and off-diagonal contribution to the ellipcity of the image is the shear γ_1 and γ_2. Therefore, we can express the transformation matrix A_{ij} in terms of the distortion tensor from Eq. (6.13):

$$A_{ij} = \delta_{ij} + \psi_{ij} \,, \qquad \psi_{ij} = \begin{pmatrix} -\kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & -\kappa + \gamma_1 \end{pmatrix} . \tag{6.29}$$

Now we need to derive how A_{ij} applies to q_{ij}. Since the observed intensity in the observed position is equal to the true intensity in the true position, we can apply Eq. (6.3) in the definition of q_{ij} and expand the deflection angle so that

$$\theta^i_S(\boldsymbol{\theta}) = \theta^i + \Delta\theta^i + \partial_{\theta^j}\Delta^i\theta^j + \cdots = A_{ij}\theta^j + \Delta\theta^i + \dots \,, \tag{6.30}$$

where $\Delta\theta^i$ and its derivative are to be evaluated at the galaxy centroid position. Since both the deflection angle and the distortion matrix are evaluated at a fixed point, we can get them outside of the integral to compute the second moment. Since we care for the shape of the galaxy, we can also drop the overall shift $\Delta\theta^i$. Therefore, we can invert the matrix to express the observed angle as function of the source angle, changing the variable of integration. This leads to

$$\begin{aligned} q_{ij} &= \frac{1}{F}\int \mathrm{d}^2\boldsymbol{\theta}_S \left|\frac{\partial\theta_k}{\partial\theta_{S,l}}\right| I_{\mathrm{true}}(\boldsymbol{\theta}_S)(A^{-1}\theta_S)_i(A^{-1}\theta_S)_j \,, \\ F &= \int \mathrm{d}^2\boldsymbol{\theta}_S \left|\frac{\partial\theta_k}{\partial\theta_{S,l}}\right| I_{\mathrm{true}}(\boldsymbol{\theta}_S) \,. \end{aligned} \tag{6.31}$$

By definition, the Jacobian is the determinant of the inverse of A, and using the properties of the determinant $|A^{-1}| = |A|^{-1}$, hence we can pull it outside of the integral, and

$$F = |A|^{-1}F_{\mathrm{true}} = \mu F_{\mathrm{true}} = \frac{F_{\mathrm{true}}}{(1-\kappa)^2\gamma_1^2 - \gamma_2^2} \,, \tag{6.32}$$

where we have defined the magnification μ of the image. Remember that lensing conserves the surface brightness, so any change in the observed flux with respect to the intrinsic one is due to a magnification of the size of the galaxy.

For the shape tensor, the Jacobian cancels with the normalization of the flux, and it becomes

$$q_{ij} = (A^{-1})_i^k (A^{-1})_j^l q_{kl}^{\rm true}\,, \tag{6.33}$$

where $q_{kl}^{\rm true}$ is the intrinsic, unlensed second-moment tensor. If we linearize the expression above we find

$$q_{ij} = q_{ij}^{\rm true} - \psi_i^k q_{kj}^{\rm true} - \psi_j^l q_{il}^{\rm true}\,, \tag{6.34}$$

which leads to the transformation of each of the independent components of the second-moment tensor:

$$\begin{aligned} q &= q_{\rm true}\left[1 + 2\kappa + 2(\epsilon_1^{\rm true}\gamma_1 + \epsilon_2^{\rm true}\gamma_2)\right]\,, \\ \epsilon_1 &= \left[1 - 2(\epsilon_1^{\rm true}\gamma_1 + \epsilon_2^{\rm true}\gamma_2)\right]\epsilon_1^{\rm true} + 2\gamma_1\,, \\ \epsilon_2 &= \left[1 - 2(\epsilon_1^{\rm true}\gamma_1 + \epsilon_2^{\rm true}\gamma_2)\right]\epsilon_2^{\rm true} + 2\gamma_2\,. \end{aligned} \tag{6.35}$$

This derivation shows how, from the measurement of galaxy shapes we can infer the matter distribution through the impact of shear, after accounting for the intrinsic shape of the galaxies themselves. Since we are in the limit of weak lensing, we can approximate the distribution of observed shapes as a proxy for the distribution of the intrinsic shapes, which is a fairly narrow distribution with a root-mean square width of $\sim \langle(\epsilon_q^{\rm true})^2 + (\epsilon_2^{\rm true})^2\rangle^{1/2}/\sqrt{2} \simeq 0.3$. The intrinsic ellipcity is random, so it is expected to cancel out after averaging the shape of many galaxies in a pixel, since the shear field is common to all of them. This is the basic concept of how we measure shear.

6.3.1 Galaxy weak-lensing statistics

The lensing potential is proportional to the gravitational potential, so it averages to zero, as the distortion tensor does. Therefore, as it is the case with all the quantities of study so far, we need to take its variance. We will consider flat sky, since most of the signal from shear comes from small scales. Taking the 2D Fourier transform of the distortion tensor, we have

$$-\psi_{ij}(\boldsymbol{\ell}) = \ell_i \ell_j \phi_L(\boldsymbol{\ell})\,. \tag{6.36}$$

Similar to the CMB polarization, the E mode corresponds to the scalar component of the distortion tensor (after removing the trace), and the B mode vanishes (since it is only generated by a curl-type deflection angle, while the deflection angle is a gradient type by definition).[2] Therefore, we have

$$E(\boldsymbol{\ell}) = \left(\frac{\ell^i\ell^j}{\ell^2} - \frac{\delta^{ij}}{2}\right)(-\psi_{ij}(\boldsymbol{\ell})) = \frac{1}{2}\ell^2\phi_L(\boldsymbol{\ell}) = \kappa(\boldsymbol{\ell})\,, \tag{6.37}$$

[2] Actually, confirming null B-mode measurements can be used to test the presence systematics in the observations.

where we have used the relation between the distortion tensor and ϕ_L above in the first equality. Therefore, the power spectrum of the E modes is proportional to the convergence and lensing potential power spectrum that we studied before, which are scalar quantities. We can compute the angular power spectrum of the E modes (or convergence field) in the similar way than we did for the angular cross-power spectrum of CMB lensing and galaxies in Eq. (6.25). In this case we have

$$C_\ell^{\kappa\kappa} = \frac{2}{\pi} \int \mathrm{d}k k^2 P_m(k, z=0) W_\ell^\kappa(k) W_\ell^\kappa(k) \,, \tag{6.38}$$

but there are a couple of particularities with respect to the CMB case. First, there is no clear source distance as before. Second, since weak lensing is a small effect, high statistics are required, thus we need to observe a large number of galaxies, possibly in a wide field. These can be obtained through photometric surveys, which however do not provide precise distances to galaxies, hence we are forced to work with statistical distance distributions, rather than precise measurements. We denote the galaxy number density distribution as a function of distance as $\mathrm{d}n/\mathrm{d}\chi$, which we normalize to unity. The lensing potential defined in Eq. (6.12) that affects the observed flux and shapes of this distribution of sources is

$$\phi_L(\boldsymbol{\theta}) = \int_0^\infty \mathrm{d}\chi \frac{\mathrm{d}n}{\mathrm{d}\chi} \phi_L^{(\chi)} \,, \tag{6.39}$$

where $\phi_L^{(\chi)}$ is the lensing potential up to a distance χ. Changing the order of integration as before we find

$$\phi_L(\boldsymbol{\theta}) = 2 \int_0^\infty \frac{\mathrm{d}\chi'}{\chi'} \Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi'), \tau_0 - \chi') \int_{\chi'}^\infty \mathrm{d}\chi \frac{\mathrm{d}n}{\mathrm{d}\chi} \left(1 - \frac{\chi'}{\chi}\right) \,. \tag{6.40}$$

By analogy with the CMB lensing case, we have now the kernel

$$W_\ell^\kappa(k) = \frac{3\Omega_m H_0^2}{2} \int_0^\infty \frac{\mathrm{d}\chi'}{\chi' a(\chi')} j_\ell(k\chi') D(z(\chi')) \int_{\chi'}^\infty \mathrm{d}\chi \frac{\mathrm{d}n}{\mathrm{d}\chi} \left(1 - \frac{\chi'}{\chi}\right) \,. \tag{6.41}$$

The shear power spectrum depends both on the amplitude of the non-linear power spectrum $P_m(k, z=0)$, but also on the abundance of matter through the presence of Ω_m in W_ℓ^κ that relates it with the gravitational potential– and through its impact on the expansion history of the Universe (which determines the geometry in the lens system). This is why weak lensing is very sensitive to the parameter combination involving σ_8 and Ω_m. Note that we use the non linear power spectrum in the equation above, while we kept the derivation of the effects from lensing to linear order. This is a good approximation because the non-linearities in the shear are much smaller than those in the three dimensional matter clustering. Therefore, nonlinear clustering and uncertainties related with the impact of the baryonic effects are usually the main limitation for theory predictions of the shear power spectrum.

There are also challenges in the observational side. The effect of lensing is very small, which requires a huge number of galaxies to recover its effect statistically. This in turn implies than (due to the shape of the luminosity function) most of the galaxies included in the analysis are very faint, which hinders the ellipticity measurements. Moreover, uncertainties in the redshift distribution of the sources and the intrinsic ellipticity of the galaxies themselves also introduce systematic uncertainties that, if larger than the statistical errors, may limit the analysis. Finally, the galaxy shapes may be intrinsically correlated, which is known as intrinsic alignments. At large scales, the main effect correlating galaxy shapes is the tidal field. This effect must be included in the analysis to avoid biased results, and can be understood as the analog for shapes of the linear bias relation for the galaxy number counts.

Since galaxy weak lensing also traces the large-scale structure, it can be cross-correlated with any other probe of the underlying matter clustering. In particular, the most common cross correlation involves galaxy clustering. The formalism to compute the cross-power spectra is similar to the one discussed in the previous section for the CMB lensing tomography. The main benefits of performing this type of cross correlations (which have been named as '3×2' analysis, also extended to $N \times 2$ if more than two probes are cross correlated, e.g., SZ effect, CMB lensing, etc) is to increase the signal-to-noise ratio, avoid potential sources of systematics, and break degeneracies with the nuisance parameters for each probe.

Of course, there are other probes of amplitude of clustering, although they do not show such large tension with respect to the predictions from *Planck* and we leave them out due to a limitation of time. Some examples include the abundance of clusters (detected using the thermal SZ effect or X-ray observations), galaxy clustering (the degeneracy with the galaxy bias can be broken including nonlinear scales), 1D Lyman-α forest power spectrum, etc.

CHAPTER 7

APPLIED SESSION 2
CLUSTERING IN THE LATE UNIVERSE

In this second Applied Session we will discuss probes of the matter distribution and clustering across cosmic time. We will initially deal with the effect of lensing, both on CMB and with respect to galaxy surveys. In this context, we will discuss how measurements from thermal Sunayev-Zeldovich clusters and weak lensing provide estimates of the S_8 parameter and we will highlight the uncertainties that arise from them. Finally, we will discuss S_8 measurements, showing how the tension seems to be related with low redshift probes.

For more detail you can refer to:

Modern Cosmology. 2nd edition. Chapter 13
S. Dodelson and F. Schmidt (2020). Elsevier Press, Cambridge.
DOI: 10.1016/B978-0-12-815948-4.00020-6

The Sigma-8 Tension is a Drag
V. Poulin, J. L. Bernal, E. Kovetz and M. Kamionkowski (2022).
arXiv:2209.06217 [astro-ph.CO].

Constraints from thermal Sunyaev-Zeldovich cluster counts and power spectrum combined with CMB

L. Salvati, M. Douspis, N. Aghanim (2018). Astron. Astrophys. 614, A13. DOI:10.1051/0004-6361/201731990.

Intrinsic and Extrinsic Galaxy Alignment
P. Catelan, M. Kamionkowski and R. D. Blandford (2001).
Mon. Not. Roy. Astron. Soc. 320, L7-L13.
DOI: 10.1046/j.1365-8711.2001.04105.x

Dark Energy Survey Year 3 results. Cosmological constraints from galaxy clustering and weak lensing
T. M. C. Abbott et al. [DES] (2022).
Phys. Rev. D 105, no.2, 023520. DOI: 10.1103/PhysRevD.105.023520

DES Y3 cosmic shear down to small scales
G. Aricò, R. E. Angulo, M. Zennaro, et al. (2023).
arXiv:2303.05537 [astro-ph.CO].

7.0.1 The S_8 parameter

To measure the amplitude of the matter clustering in the late Universe, we can use the parameter

$$S_8 = \sigma_8 \sqrt{\frac{\Omega_m}{0.3}}\,, \tag{7.1}$$

where Ω_m is the matter density parameter today and σ_8 the rms of the amplitude of the matter perturbations smoothed over a scale $R = 8\,h^{-1}$Mpc. While Ω_m controls the background amount of matter in the Universe, σ_8 describes its clustering properties: low values of σ_8 mean that the Universe is smooth. The two parameters, as figure 7.1 shows, are degenerate in determining S_8.

The more matter there is and the more it forms clustered structures, the more effectively it deviates the paths photons follow before reaching the observer. Therefore, the parameter S_8 directly relates with two effects: lensing and thermal SZ. By measuring these effects, we can constrain S_8 once we account for uncertainties and degeneracies with other parameters; these measurements can then be compared with ΛCDM predictions on the growth of structures. We will see that, while CMB lensing seems to be coherent with ΛCDM results from the CMB, weak lensing on galaxy surveys and cluster counts related with thermal SZ predict a lower value of S_8. Since these effect are due to structures at low z, as figure 7.1 shows, this range seems to be the one in which the tension is larger: this implies that the observed level of clustering is not growing as rapidly as ΛCDM predictions.

Which are, in your opinion, possible ways to explain the low value S_8 has at low z?

Discussion

The σ_8 tension is even more unclear and uncertain than the H_0 tension. On one side, the value extrapolated from CMB is related with larger scales than the ones probed e.g., by galaxy lensing: the former arrives at $\sim 0.1\,h\text{Mpc}^{-1} \sim$ size of the Local Group, the latter reaches $k \sim 1\,h\text{Mpc}^{-1} \sim$ size of medium-size dark matter halos. Therefore, small scales effects (non linearities, exotic forms dark matter or dark energy...) possibly do not affect CMB scales. On the other hand, if not properly modelled, they act as nuisance parameters in the other probes, leading to biased parameter estimations. Indeed, one of the main issues in dealing with the S_8 tension is the presence of very large uncertain in low-z measurements; we will see where do they come from.

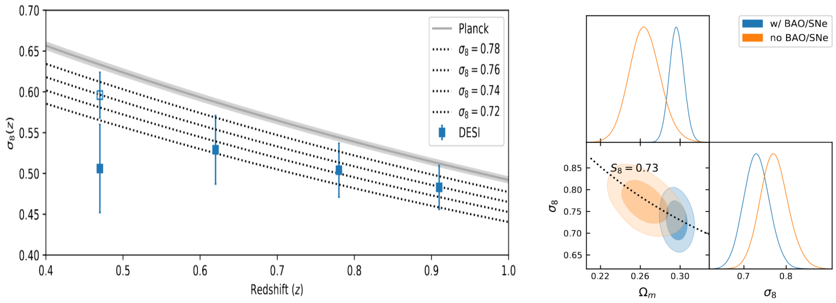

Figure 7.1 Sources: arXiv: 2111.09898.

7.0.2 CMB lensing

As you discussed in detail during the lectures, CMB lensing distorts the hot and cold spots of the temperature field around foreground masses. You computed how the lensing affects the observed CMB power spectrum

$$C_\ell^{\text{obs}} = C_\ell + \int \frac{d^2\boldsymbol{\ell}'}{(2\pi)^2} \left[\boldsymbol{\ell}'(\boldsymbol{\ell}-\boldsymbol{\ell}')\right]^2 C^{\phi_L\phi_L}_{|\boldsymbol{\ell}-\boldsymbol{\ell}'|} C_{\ell'} - C_\ell \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} (\boldsymbol{\ell}\boldsymbol{\ell}')^2 C^{\phi_L\phi_L}_{\ell'}\,, \quad (7.2)$$

ultimately through the effect of its potential, which is related to the deflection angle by

$$\theta^i_S = \partial_{\theta^i}\phi_L = \partial_{\theta^i}\left[2\int_0^\chi \frac{d\tilde{\chi}}{\tilde{\chi}} \Phi(\boldsymbol{x}(\boldsymbol{\theta},\tilde{\chi}))\left(1-\frac{\tilde{\chi}}{\chi}\right)\right]. \quad (7.3)$$

The foreground lenses deviates the photon paths, blurring the acoustic peaks and distorting the associated angular scales, as figure 7.2 shows. It is customary to describe the amplitude of CMB lensing effect through the parameter

A_L, which is defined by the relation $C_{\ell,\mathrm{obs}}^{\phi\phi} = A_L C_\ell^{\phi\phi}$: $A_L = 0$ describes the unlensed CMB, while the theoretical expectation of the lensed CMB is $A_L = 1$. From the observed CMB power spectrum, it is possible to reconstruct the lensing power spectrum: its shape is consistent with $A_L = 1$, but if only the higher multipoles are considered, a smaller preference for $A_L > 1$ seems to arise. This however can be due simply to degeneracies with other parameters or foreground cleaning.

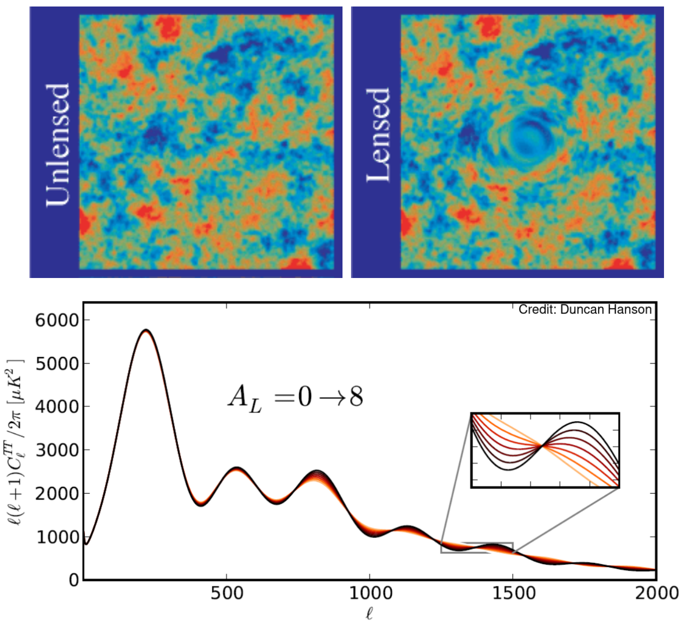

Figure 7.2 Sources: Wayne Hu website (top), `https://cosmologist.info/notes/LensedCMB-Cargese17.pdf` (bottom).

Planck 2018 results on lensing provide $S_8 = 0.832 \pm 0.013$. It is important to note that, even if the support kernel that describes the integration of the lensing potential in the $C_\ell^{\phi_L \phi_L}$ equation, spans in z between recombination and today, its peak occurs at $z \sim 2$, being therefore less sensitive to low z.

7.0.3 Thermal SZ cluster

After recombination, photons free stream in the Universe. However, the status of the Universe change as long as time passes and late time effect can affect the CMB photon path. How?

Discussion

Once stars and galaxies form, high energy radiation is produced, which in turn ionizes the gas the Universe is filled with. This happens during the reionization era around $z \sim 30$, as well as in regions of the local Universe where star formation is active, namely in the intergalactic medium inside clusters. Here, the ionizing radiation implies the presence of hot electrons, which have a non negligible Compton cross section with CMB photons. The electron are more energetic than the CMB photons, therefore the scattering these undergo result in an increased temperature for the photons themselves.

The change in the CMB photon temperature can be expressed as

$$\frac{\Delta T}{T}(\nu, \hat{\boldsymbol{n}}) = \left[\frac{h_P \nu}{k_B T} \coth \frac{h_P \nu}{2 k_B T} - 4\right] \frac{\sigma}{m_e c^2} \int n_e k_B T_e dx_{\text{phys}} \,, \tag{7.4}$$

where T, T_e respectively are the CMB and electron temperatures, ν is the photon frequency, h_P the Planck constant, σ the cross section, n_e the physical electron number density and x_{phys} the physical distance along the line of sight in the $\hat{\boldsymbol{n}}$ direction. The quantity inside the integral is the electron pressure. This effect is called thermal Sunyaev Zeldovich (tSZ).

By looking at the expression for $\Delta T/T$, infer how the CMB photon flux change when looking in the direction of a cluster filled with hot electrons. How does the CMB map in this direction change because of the tSZ effect and how does the cluster appear on the map?

Discussion

The CMB is measured in different frequency channels, which collect photons with different energies. In the angular positions associated with cluster line of sights, tSZ determines a lack of low-energy photons and an increased number of high-energy ones. Therefore, high frequency channels gain more and more photons and the observed intensity increases: in the map, this is equivalent to having hotter temperatures. Similarly, low frequency channels receive less photons and observe a lower intensity i.e., they appear colder than the average. This is evident in figure 7.3, which shows the effect of the ABELL 2319 cluster on the CMB ($\sim 2\,\text{deg}^2$ patch). The tSZ effect has a null point at frequency $\nu \sim 217\,\text{GHz}$, where

CMB appears without tSZ effect: this point corresponds to

$$\begin{aligned}
&\frac{h_P\nu}{k_BT}\coth\frac{h_P\nu}{2k_BT}-4=0\\
&\frac{\exp(h_P\nu/2k_BT)+\exp(-h_P\nu/2k_BT)}{\exp(h_P\nu/2k_BT)-\exp(-h_P\nu/2k_BT)}=4\frac{k_BT}{h_P\nu}\\
&\nu\sim 1.9\frac{2k_BT}{h_P}\sim 1.9\,\frac{2\cdot 1.4\cdot 10^{-23}\,\mathrm{J/K}\cdot 2.7\,\mathrm{K}}{6.6\cdot 10^{-34}\,\mathrm{J/Hz}}\sim 217\,\mathrm{GHz}
\end{aligned}\tag{7.5}$$

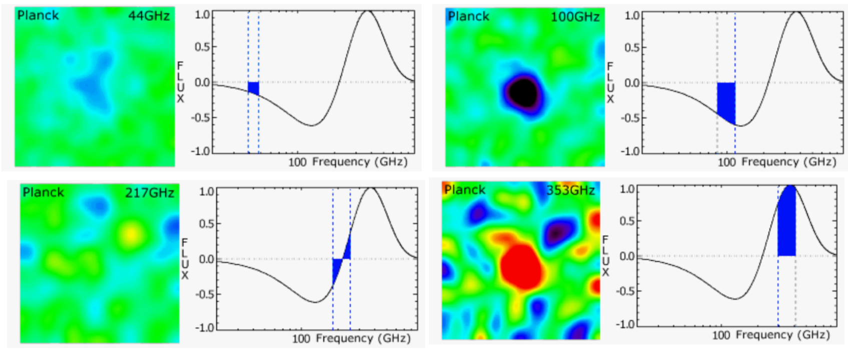

Figure 7.3 Source: ESA/Planck Collaboration

CMB maps can then be used to locate clusters across the sky: by doing so, we can count the clusters and, in analogy to what we do with any other observable in cosmology, compute the tSZ angular power spectrum.

Which ingredients do we need to model the tSZ power spectrum?

Discussion

Eq. (7.4) indicates that, to estimate the tSZ effect on CMB anisotropies, we need to know the electron number density along the line of sight. Since they are found in clusters, this information can be obtained by knowing how electrons "populate" DM halos of different masses and which is the halo mass function dn/dM, namely the number density of the halos distributed in the mass bins and integrated over the mass spectrum and the observed volume. Moreover, we need to plug an information related with the internal structure of the halos themselves.

The tSZ angular power spectrum can then be written as

$$C_\ell^{tSZ}=C_\ell^{tSZ,1h}+C_\ell^{tSZ,2h}\tag{7.6}$$

where $C_\ell^{tSZ,1h}$ represents the one halo term (due to contributions within the single halo), while $C_\ell^{tSZ,2h}$ is the two halo term (due to correlation between different halos). These are obtained as

$$C_\ell^{tSZ,1h} = \int dz \frac{d^2V}{dzd\Omega} \int dM \frac{dn}{dM}(z,M)\exp(\sigma^2_{\ln Y}/2)\cdot \\ \cdot \left[\frac{\sigma}{m_e c^2}\frac{4\pi r_p}{\ell_p^2}\int dx_r\, x_r^2 \frac{\sin(\ell x_r/\ell_p)}{\ell x_r}\ell_p n_e(z,M,x) k_B T_e(z,M,x)\right]^2 \tag{7.7}$$

$$C_\ell^{tSZ,2h} = \int dz \frac{d^2V}{dzd\Omega} P(k,z)\left[\int dM \frac{dn}{dM}(z,M) b_h(z,M)\cdot \\ \cdot \frac{\sigma}{m_e c^2}\frac{4\pi r_p}{\ell_p^2}\int dx_r\, x_r^2 \frac{\sin(\ell x_r/\ell_p)}{\ell x_r}\ell_p n_e(z,M,x_r) k_B T_e(z,M,x_r)\right]^2 \tag{7.8}$$

where r_p is the characteristic radius of the pressure profile, $x_r = r/r_p$ the dimensionless radial scale and $\ell_p = D_A(z)/r_p$. In the one halo term, $\exp(\sigma^2_{\ln Y}/2)$ describes some dispersion in the distribution, while in the two halo term b_h is the halo bias and $P(k,z)$ the matter power spectrum.

The halo mass function in the previous equations depends on the power spectrum, thus C_ℓ^{tSZ} contains information on the cosmological parameters, including σ_8. Moreover, the halo bias itself depends on this parameter since it is usually computed as $b(M,z) = 1+\delta_c/D^2(z)\sigma^2(M)$, where δ_c is the critical density required for the collapse, $D(z)$ the growth factor and $\sigma(M)$ the rms of the fluctuations on the mass scale of the halo considered.

Therefore, from tSZ we can estimate σ_8 but at the price of degeneracies with respect to the halo mass function, bias and mass calibration. The value obatined is $\sigma_8(\Omega_m/0.33)^{0.25} = 0.765 \pm 0.035$.

7.0.4 Galaxy weak lensing

Weak lensing has two effects on galaxies, which we call convergence and shear. They enter in the lensing transformation matrix as

$$A_{ij} = \delta_{ij} + \psi_{ij}\,, \qquad \psi_{ij} = \begin{pmatrix} -\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & -\kappa+\gamma_1 \end{pmatrix}. \tag{7.9}$$

The form of the lensing distortion matrix is analogous to the polarization matrix for a radiation field

$$\begin{pmatrix} -\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & -\kappa+\gamma_1 \end{pmatrix} \leftrightarrow \begin{pmatrix} I+Q & U \\ U & I-Q \end{pmatrix} = I_{ij} \tag{7.10}$$

where I is the intensity (which in our case in analogous to the convergence), while U and Q describe the polarization components in the plane orthogonal to the line of sight (the shear in our case). I, U, V are three of the Stokes parameters; the fourth is absent since we ignore circular polarization.

Use this analogy and imagine a galaxy as a circular distribution of matter. How do convergence and shear change its appearance? Then, think about a group of galaxies circularly distributed: how do they get displaced? Which uncertainty is already evident from this example?

Discussion

In general, the polarization matrix can be expressed as $I_{ij} = I\delta_{ij} + I^T_{ij}$. The first term is the trace of the matrix; in our case it due to the convergence, which magnifies the source. However, since the total flux is conserved, its effect is to change the apparent size of the source isotropically. The shear instead enters the traceless tensor I_{ij} and it introduces anisotropic stretching in the image. These effects are visualized in the top left panel of figure 7.4. The shear pattern is tangential to the mass concentration in the convergence, as the bottom panel of the figure shows.

Analogously to CMB polarization, I_{ij} can be rewritten in terms of two components that describe the behaviour under rotations. The first (E mode) is a scalar, while the second is a transverse-traceless tensor (B mode). Their effect is illustrated in the top right panel of figure 7.4.

One of the main issue with weak lensing is that we know a priori neither the intrinsic ellipticity of the galaxy nor its orientation with respect to the line of sight. These two, then, are degenerate with the distortions in the shape that the weak lensing introduces.

One interesting observable is the correlation between the observed ellipticities of galaxies, which can be used as an indicator of the shear field induced by mass inhomogeneities along the line of sight when intrinsic ellipticities are randomly distributed. In this case, B modes vanish since the shear is derived from the gradient of the lensing potential, which can not produce curl deflection angles (in directions not parallel or orthogonal to the line of sight).

In the lectures, you derived the angular power spectrum for the E-modes

$$C_\ell^{\kappa\kappa} = \frac{2}{\pi}\int dk k^2 P_m(k, z=0) W_\ell^\kappa(k) W_\ell^\kappa(k)\,, \tag{7.11}$$

where the kernel is

$$W_\ell^\kappa(k) = \frac{3\Omega_m H_0^2}{2}\int_0^\infty \frac{d\chi'}{\chi' a(\chi')} j_\ell(k\chi') D(z(\chi'))\int_{\chi'}^\infty d\chi \frac{dn}{d\chi}\left(1 - \frac{\chi'}{\chi}\right). \tag{7.12}$$

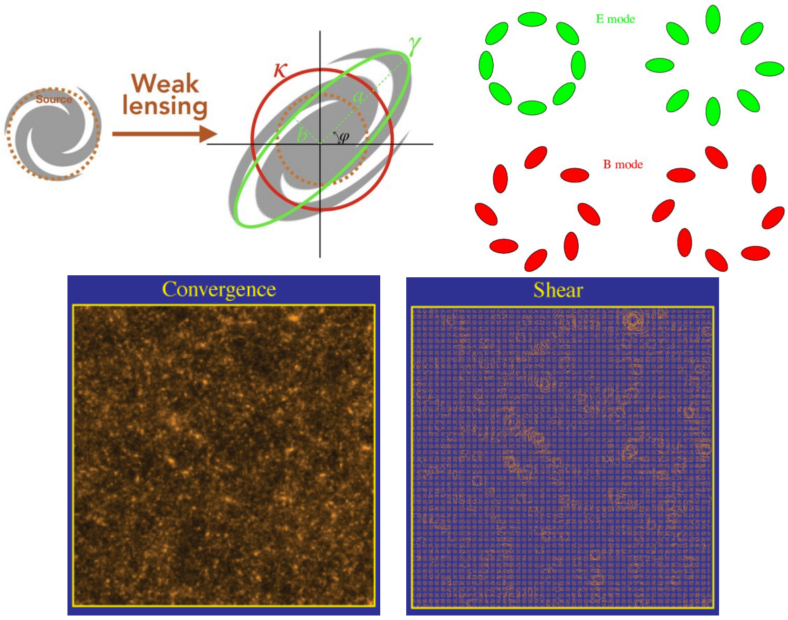

Figure 7.4 Sources: `www.researchgate.net/publication/334098793_Cosmic_Magnification_in_COSMOS` (left), arXiv:1109.1121 (right), Wayne Hu website (bottom).

We apply the flat sky Limber approximation to remove the Bessel functions:

$$\begin{aligned} \frac{2}{\pi}\int dk k^2 j_\ell(k\chi) j_\ell(k\chi') &= \frac{\delta^D(\chi-\chi')}{\chi^2} \\ \frac{2}{\pi}\int dk k^2 j_\ell(k\chi') j_\ell(k\chi') P(\mathbf{k}) &= \int \frac{d\chi'}{\chi'^2} P\left(k = \frac{\ell+1/2}{\chi'}\right) \end{aligned} \tag{7.13}$$

($k\chi \sim \sqrt{\ell(\ell+1)} \sim \ell + 1/2$ is where the Bessel functions peak).

We rewrite the previous expressions as

$$C^{ij}_{\kappa\kappa}(\ell) = \int d\chi' \frac{W^i_\kappa(\chi')W^j_\kappa(\chi')}{\chi'^2} P_{\kappa\kappa}\left(k = \frac{\ell+1/2}{\chi'}, z(\chi')\right) \tag{7.14}$$

$$W^i_\kappa(\chi') = \frac{3\Omega_m H_0^2}{2}\int_0^{\chi'_H} d\chi n^i_s(\chi) \frac{\chi'}{a(\chi')}\left(1 - \frac{\chi'}{\chi}\right) \tag{7.15}$$

We are here considering sources in two different redshift bins i, j; this is done in order to account for both auto- and cross- power spectra between redshift bins, so it is possible to perform shear tomography.

On the other hand, observations are made in real, angular space; therefore, it is easier to estimate the γ_1 and γ_2 correlation functions. Since we can arbitrary define the coordinate system and project $\gamma_{1,2}$ on it, it is customary to study the shear in terms of tangential component and cross-component, namely γ_t parallel to the line that connects two galaxies and $\gamma_\times$ oriented $\pm 45°$ with respect to it. Under this configuration, it can be shown that $\langle \gamma_t(0)\gamma_t(\theta)\rangle \pm \langle \gamma_\times(0)\gamma_\times(\theta)\rangle = \xi_\pm^{ij}$ can be re-projected in terms of the angular power spectrum (and the Bessel functions) as

$$\xi_+^{ij}(\theta) = \int \frac{\ell d\ell}{2\pi} J_0(\ell\theta) C_{\kappa\kappa}(\ell)\,, \quad \xi_-^{ij}(\theta) = \int \frac{\ell d\ell}{2\pi} J_4(\ell\theta) C_{\kappa\kappa}(\ell) \tag{7.16}$$

where θ is the angular distance between two galaxies.

The value of $\xi_\pm$ indicates how much the shear (i.e., the ellipticity of the galaxies) correlates on the different scales θ. Qualitatively, how does the $\xi_\pm(\theta)$ plots should look like? How can we extract information on S_8 from these measurements?

Discussion

The shear correlation function is larger at smaller scales: this can be computed from the shape of $J_{0,4}(\ell\theta)$, but also understood by thinking that we see correlated distortions in the shape of galaxies whose line of sights "pass" through the same matter inhomogeneity. This implies that, to be affected by the same matter "clump", galaxies (projected on the sky) can not be too far apart one from another. Therefore, $\xi_\pm(\theta)$ have to decrease going from small to large θ, with larger amplitude in the $+$ case, as you can see in figure 7.5.

The shear correlation function directly depends on the variance of the fluctuations of the matter field, whose amplitude is given by the power spectrum. To compute the variance on a certain scale (e.g., $R = 8\,h^{-1}$Mpc), we need to filter the field through a window function $\mathcal{W}(x)$ (e.g., a tophat with width R): this is done through a convolution, which in Fourier space simply becomes

$$\sigma^2_{\mathcal{W}} = \frac{1}{2\pi} \int d\ln k\, k^3 P_m(k) |\mathcal{W}(k)|^2 \tag{7.17}$$

Therefore, when performing data analysis we can look for the value of σ_8 that allows to get $P_m(k)$ and from that obtain the $C_\ell^{\kappa\kappa}$ that best fits the observations. However, from the Poisson equation we see that the amplitude of $P_m(k)$ also depends on Ω_m: for this reason, the two parameters are degenerate and it is useful to define S_8 instead.

7.0.5 Intrinsic alignment problem

Noise in the shear measurement arise if intrinsic ellipticities are already correlated, due to the formation process of galaxies in the same halo. Two are

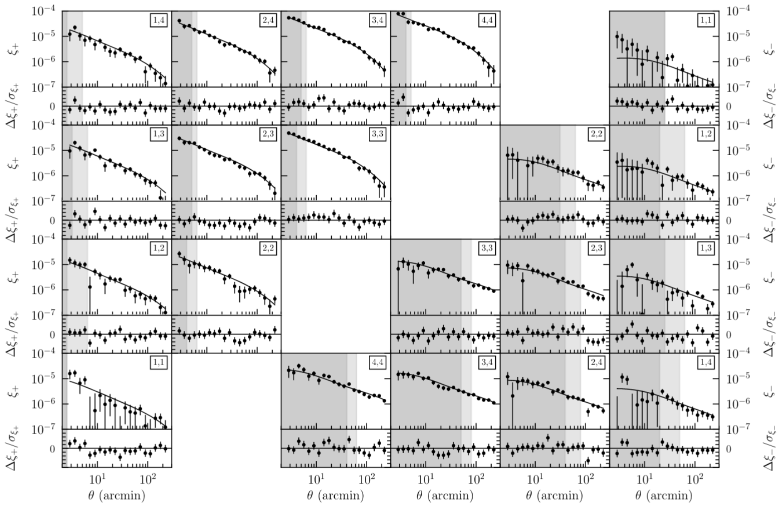

Figure 7.5 Source: Abbott et al., DES Y3.

physical quantities that can lead to intrinsic alignments: the shape of the host halo, which induces a tidal field and therefore a preferential elongation direction, or the angular momentum of the halo itself, which gets imprinted on the galaxy orientation. It can be shown that the first case is relevant for elliptical, isolated galaxies, while the second dominates the case of disk galaxies.

Let us start with the dependence on the halo shape.

Consider a galaxy that forms in the presence of a tidal gravitational potential ϕ. The galaxy is observed along the $\hat{x}$ direction and the points of its image in the sky are described by the angular position $\boldsymbol{\theta} = (\theta_y, \theta_z)$. Consider the ellipticities ϵ_+ and $\epsilon_\times$, the former describing the stretching in the $\hat{\theta}_y - \hat{\theta}_z$ direction and the latter in the direction rotated by 45°. How does ϕ affect $\epsilon_+, \epsilon_\times$? How does the galaxy shape change?

Discussion

We can consider the galaxy as a sphere of test masses that move inside the potential $\phi(\boldsymbol{x})$, which slowly varies in space. By Taylor expanding the potential around the origin, we get

$$\phi(\boldsymbol{x}) = \phi(\boldsymbol{x} = 0) + \nabla\phi|_{\boldsymbol{x}=0} + \nabla\nabla\phi|_{\boldsymbol{x}=0} \,. \tag{7.18}$$

The zeroth order is just a ground level for the potential, so it has no physical effect on the test particles motion. The linear order, which involves $g = \nabla\phi$, describes a uniform translation in the sphere that does not affect its shape. Finally, the Hessian introduces different accelerations in the different points. By looking again at the directions $\hat{\theta}_y - \hat{\theta}_z$ and rotated by 45°, we can obtain

$$\epsilon_+ \propto (\partial_y^2 - \partial_z^2)\phi \tag{7.19}$$

$$\epsilon_\times \propto 2\partial_y\partial_z\phi \tag{7.20}$$

Look at figure 7.6: if the gravitational collapse through which the galaxy forms takes place in a region with constant tidal gravitational field, then the acceleration on different sides of the galaxy differs. Thus, the collapse, instead of being spherical, is anisotropic and the galaxy acquires an intrinsic ellipticity.

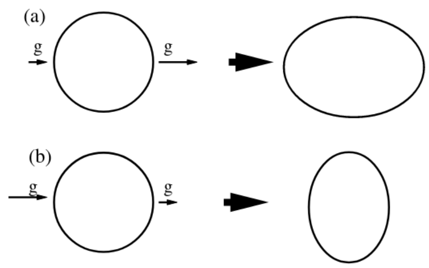

Figure 7.6 Source: Catelan et al..

In analogy to lensing shear, we can consider the contributions these ellipticities induce to the shear by integrating them along the line of sight

$$\epsilon_+ \propto \int d\chi \frac{dn}{d\chi}(\partial_y^2 - \partial_z^2)\phi\,, \quad \epsilon_\times \propto \int d\chi \frac{dn}{d\chi}(2\partial_y\partial_z^2)\phi \tag{7.21}$$

and by applying the Fourier transform (so derivatives disappear) and taking the Limber approximation, we get

$$C^{II}(\ell) \propto \left(\frac{3}{2}\Omega_m H_0^2\right)\int d\chi \left(\frac{dn}{d\chi}\right)^2 \frac{1}{\chi^2}P(\ell/\chi)\,. \tag{7.22}$$

where $dn/d\chi$ is the line of sight distribution of sources in the halo (e.g., a top hat distribution). Note that in this case the kernel that the intrinsic alignment provides to the power spectrum is similar to the one that is used for galaxy clustering. The constant of proportionality between the ellipticities and the tidal field can be estimated through the expected rms of the ellipticity of individual galaxies. Generally, it is found that the intrinsic alignments are subdominant with respect to weak lensing distortions.

We can now account for the angular momentum-induced shape. The tidal gravitational field induce a torque in the halo, producing an angular momentum $L_\alpha \propto \epsilon_{\alpha\beta\gamma}\mathcal{I}_{\beta\sigma}\partial_\gamma\partial_\sigma\phi$ (ϵ in this case is the Levi-Civita tensor and $\mathcal{I}$ the galaxy moment of inertia). If a galaxy forms in the halo, ellpiticities will be induced by $L_{x,y,z}$; even if the moment of inertia varies for different galaxies, on average they will be aligned with the major axis of the tidal gravitational field. It can be showed that in this case

$$\epsilon_+ \propto (\partial_y^2\phi)^2 - (\partial_z^2\phi)^2 \, , \, \epsilon_\times = (\partial_y\partial_z\phi)(\partial_y^2\phi + \partial_z^2\phi) \tag{7.23}$$

From this equations we can see that ellipticities show a quadratic dependence on the tidal field, which leads to curl components in the induced shear: in the case of intrinsic alignments, B modes are non vanishing.

How can we disentagle between weak lensing and intrinsic alignment?

Discussion

- Weak lensing is larger for more distance sources, since it can account for more matter along the line of sight, while intrinsic alignment is not. Moreover, the ellipticities of two galaxies nearby in the sky correlate even if they are far apart in redshift if they are due to weak lensing, while if they are intrisic they correlate only when galaxies are in the same halo, i.e. at the same redshift.
- If the same population of sources of $\sim$ known ellipticity is used, correlation between them has to be due to the intrinsic alignment and not to weak lensing.
- Shear can be cross correlated with the convergence to isolate the lensing contribution, which can be related with the density of the sources in the field (which is affected by magnification).

7.0.6 How results can be affected by uncertainties

The angular correlation of galaxy ellipticities probes the shear field and thus can be used to estimate S_8. As we saw, its effect is larger at small scales; here, however, we have less control on the modelling: non linearities emerge in the power spectrum computation; intrinsic alignment can be modelled in different ways; baryon effect can lead to relaxation processes that smooth the power spectrum. Accounting for all these effects is not easy and it leads to strong dependencies of the final results in the modelling. For instance, figure 7.7 shows how the same data from DES[1] lead to very different estimates

[1]The Dark Energy Survey results from one and three years of data have been released in 2017 and 2021. Observations were made between 2013 and 2019 using the 4m telescope

of S_8 according to the modelling adopted. Official DES results (DES TATT) account for intrinsic alignment and cut the smaller scales to avoid the effect of baryons. Including the small scales, accounting for baryons with different model prescriptions (BCM fiducial, BCM-extreme) or changing the models for intrinsic alignment (TATT to NLA) or non linearities (DES, HALOFIT, BACCOemu) lead to large variations in the final estimate: while DES official results provides $S_8 = 0.759^{+0.025}_{-0.023}$ (2.3σ tension with Planck), the revised model in Aricò et al. leads to $S_8 = 0.799^{+0.023}_{-0.015}$ (0.9σ tension with Planck).

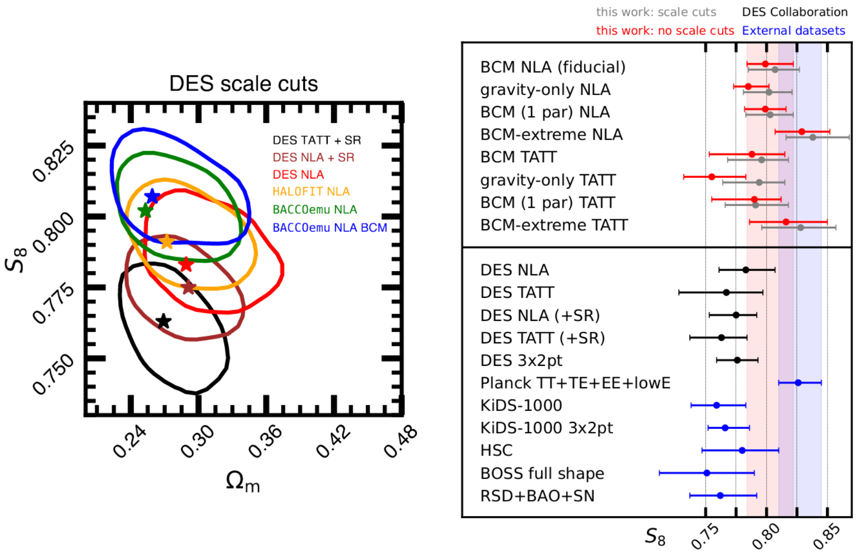

Figure 7.7 Source: Aricò et al..

at Cerro Tololo Inter-American Observatory (CTIO) in the Chilean Andes. The wide-area survey observes 5000 square degrees in the southern sky, out of the Galactic plane to avoid star and dust emission from the Milky Way. Imaging observations are made in 5 photometric bands in NUV, optical and NIR up to $z \sim 1.5$ (max nominal $z = 3$, but with very low number density in the last bin).

CHAPTER 8

LECTURE 5: COSMIC TENSIONS AND HOW TO RESOLVE THEM

During these lectures we have discuss how, from the basics of cosmological perturbation theory (which we have limited to linear order in our discussions) we can predict how the gravitational potential and the matter overdensities grow. This allows us to predict the matter power spectrum and also, together with the study of photon perturbations, to predict the angular power spectrum of CMB anisotropies. These two are the main sources of information to probe the Universe, together with standard candles, rules and sirens which allow to probe the expansion history of the Universe. Despite the success of the consensus model of cosmology, ΛCDM, which can reproduce *most* of the observations with astonishing precision, there are persisting tensions between experiments.

The largest tension involves the current expansion rate of the Universe, quantified by H_0. The inferred value by *Planck* assuming ΛCDM is $\sim 5\sigma$ smaller than the direct measurements from the distance ladder calibrated with cepheid stars and SNeIa from the SH0ES collaboration. Other local and low-redshift probes of H_0 also favor for high values of H_0 although with

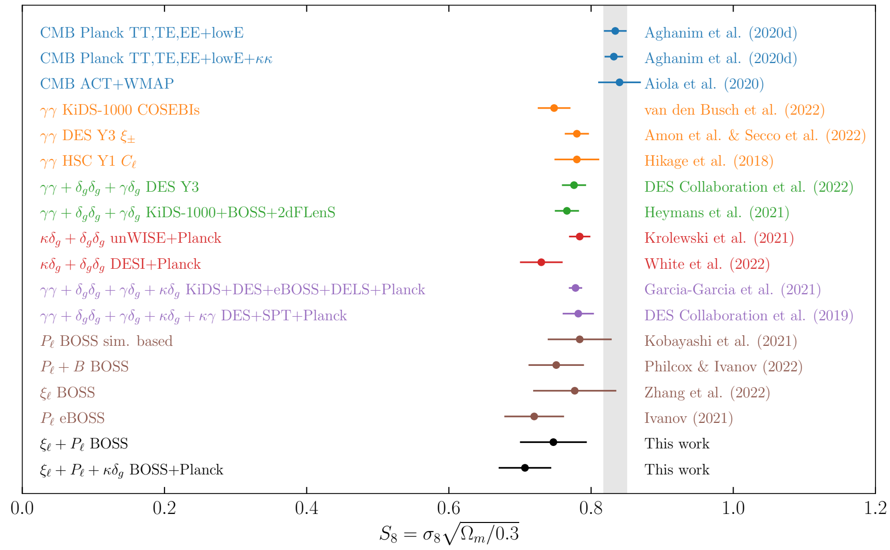

Figure 8.1 A summary of recent S_8 constraints. Different colors indicate different combinations of data that have been used for the constraints. We consider CMB measurements, marked with blue, cosmic shear (γ), projected galaxy clustering (δ_g), CMB lensing κ, redshift space clustering, marked by brown. Figure taken from Ref. (25).

higher uncertainties, which entails a smaller tension with the inferred value by *Planck*. Although non conclusive, the fact that no systematic error that could explain the tension has been found and that there is a consistent trend in low-redshift and local measurements may hint that the tension is due to actual new physics that has not been taken into account in our models. A summary of the measurements from different probes and data combinations can be found in Fig. 5.1.

Besides the H_0 tension, there is another (smaller) tension involving the clustering at small scales. In this case, probes of small-scale, low redshift clustering (in particular galaxy weak-lensing studies) measure a $\sim 2-3\sigma$ lower clustering than the prediction according to ΛCDM constrained by *Planck*. Since the largest tension involves CMB and galaxy weak lensing, it is usually quantified in the parameter combination $S_8 \equiv \sigma_8(\Omega_m/0.3)^{1/2}$ best constrained by cosmic shear.[1] Interestingly, in this case there seems to be a consistent trend between many, independent measurements and cosmological probes observing lower clustering than the prediction from *Planck* assuming ΛCDM,

[1]Remember that $\sigma_8 = \left(\int \mathrm{d}k k^2 W_8(k)^2 P_m(k)\right)^{1/2}$ is the root-mean square of the matter density perturbations today smoothed over a top-hat spherical filter of 8 Mpc/h radius, the Fourier transform of which is given by W_8.

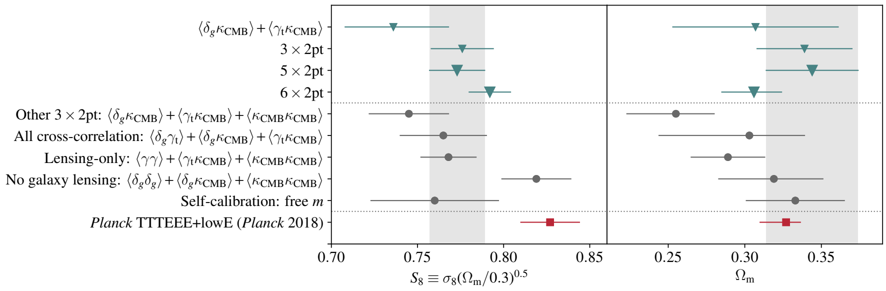

Figure 8.2 Comparison of the cosmological constraints resulting from different combination of two point function involving DES measurements of galaxy positions and shear, and SPT+*Planck* measurements of CMB lensing and primary fluctuations. 3×2 refers to galaxy clustering and shear, 5×2 adds CMB lensing only in cross correlation with galaxy clustering and shear, and 6×2 also adds the autocorrelation of CMB lensing. Figure taken from Ref. (26).

which together to the H_0 tension may indicate a high-z/low-z break down of ΛCDM. However, as we will see, it is very difficult to fix both tensions invoking a single piece of new physics. A summary of the S_8 measurements from different probes and collaborations can be found in Figs. 8.1 and 8.2.

We will dedicate this chapter to build over the concepts discussed in the previous lectures and list the conditions that, given current observations, any modification of the standard cosmological model must include to address these cosmic tensions. We will also comment some of the main proposals in the literature. A priori, we will consider the H_0 and the S_8 tensions separately (i.e., we will not attempt to solve both at the same time), but will comment potential overlaps of problems.

8.1 The H_0 tension

We have already discussed during these lectures the measurements involved in the Hubble constant tension. Furthermore, we have commented on some of the main features driving the constraining power on the cosmological observations. On one hand, we have the position of the CMB peaks, that directly depends on the sound horizon and the distance to the last scattering surface. On the other, SNeIa + BAO tightly constrain the expansion history at low redshift and the parameter combination $r_{\rm d}h$.

Here we comment on the main general features that a model must include to be on the run to solve the H_0 tension, and later mention and discuss some of them.

8.1.1 Features to solve the H_0 tension

Some of the proposals that have been ruled out during the last 5-10 years of study of the H_0 tension relied on the following deviations from ΛCDM:

- A modification of the low-redshift $H(z)$ through dark energy modifications, to keep fixed the angular diameter distance to recombination. This would allow for a higher value of H_0 with a standard $r_{\rm d}$ keeping θ_* untouched. However, this family of proposals is ruled because of the tight constraints on low-z $H(z)$ and because this potential solution does not address the *mismatch* between $r_{\rm d}$ and h regarding the BAO measurements.

- Sharp increases in $H(z)$ for $z \to 0$. This family of models was motivated by the fact that SNeIa datasets like Pantheon do not include measurements of the distance moduli at $z \lesssim 10^{-2}$. Besides the extreme fine tuning required for this possibility to apply, this proposal only works if we consider that cosmic SNeIa are normalized by H_0. However, effectively, they are normalized by the absolute magnitude as calibrated by local-distance measurements (e.g., SH0ES, CCHP, etc.). While for most cases this is equivalent, drastically changing $H(z \to 0)$ and not M_B introduces a tension between cosmological measurements and the cosmic distance ladder constrained by SNeIa.

- Violation of the distance duality relation: $D_M = D_L/(1+z)$. This relationship is usually a core assumption of most cosmological analyses. This scenario does not directly address the H_0 tension, but challenges one of the pillars it is based on. However, any consistency test has been passed without any problem, and the consistency between SNeIa and BAO does not indicate any potential problem. Furthermore, a viable particle physics model that matches all observations while affecting the distance duality relation is very challenging.

- New particle interactions (e.g., neutrino-dark matter interactions) that introduced an aparent phase-shift in the CMB oscillatory pattern. In this scenario, the location of the peaks is given by an additional phase-shift (with respect to the standard free-streaming ΛCDM neutrinos). Effectively, these models change the *interpretation* of θ_s in terms of r_s, which forces to modify the angular diameter distance to the last scattering surface by changing H_0. However, it was proved that BAO measurements are robust against early-time modifications of ΛCDM and, according to these models, H_0 would change but $r_{\rm d}$ not, which enters in tension with the strong constraints on $r_{\rm d}h$ from BAO measurements.

From this discussion and that from previous days, it seems clear a valid solution of H_0 must *at least* include modifications to ΛCDM before recombination to modify the sound horizon at radiation drag and allow for a change in H_0 keeping $r_{\rm d}h$, and the CMB peak locations, constant. However, the CMB power spectra include a great number of features that makes them very sensitive to new physics, besides the acoustic peaks.

Ignoring the suppression due to reionization and the photon diffusion, the amplitude envelope of the CMB power spectrum is controled by the matter-to-radiation ratio. Recall that the photon-baryon plasma exhibits acoustic oscillations in scales within the horizon. As a given Fourier mode crosses the horizon, the resulting gravitational potential decay provides a near-resonant driver of the oscillation. The greater the ratio of matter to radiation at horizon crossing, the less the decay, and the lower the amplitude of the resulting oscillation. The envelope grows with ℓ until it plateaus at angular scales smaller than $\theta_{\rm eq}$ ($\ell \simeq 143$), the angular scale associated to $k_{\rm eq}$. Therefore, the potential envelope depends a lot on the moment of equality and the horizon size at that moment, $k_{\rm eq} \propto \Omega_m h^2$ in ΛCDM.

The photon diffusion causes a smoothing of CMB anisotropies at very small scales, due the photon mean free path and the overall comoving distance covered by a photon in a Hubble time. Therefore, it depends on the expansion history of the Universe before recombination and in the interactions between photons and other particles. Therefore, any significant change of the expansion history must be (roughly) limited to times between matter-radiation equality and slightly earlier than recombination, such as the scale of equality and diffusion are untouched (see Fig. 4.9). This requirement, together with the effects of the phase shift and the good constraints on the peak location, rule out vanilla additional relativistic species (introduced with higher values of $N_{\rm eff}$) or some of the studied flavors of strong-interacting neutrinos.[2]

Therefore, it seems that the most promising scenarios involve the addition of components that increase the expansion rate of the Universe between those two moments. However, as we will see, the impact of these new components at linear perturbation level must be included. This, along with the creation of new degeneracies and the shift of standard ΛCDM parameters towards different values introduce additional complications.

8.1.2 Proposals beyond ΛCDM

Here we discuss *some* of the most promising alternatives to ΛCDM proposed to solve the H_0 tension. Of course, these have been proposed by different research groups at different moments in time (with different data available). Therefore, different analysis choices may bias the comparison between their performance. A fair comparison (and summary!) of promising models that can solve the H_0 tension can be found in Ref. (27). A more exhaustive (but less recent) review can be found in Ref. (10).

[2]This is why, adding relativistic species *featuring* new, specific interactions to cancel their standard effect at the perturbation level is a potential (although fine tuned) family of models that may solve the H_0 tension.

The models that show steps in the right direction to solve the H_0 constant can be classified in three main groups: addition of new relativistic species with new interactions to cancel their effect in the perturbations (sometimes labeled dark radiation); adding a new component (either dark energy or modified gravity) to boost $H(z)$ between matter-radiation equality and recombination; changing some of the core assumptions to reduce the sound horizon. We briefly discuss those that perform *relatively* well. However, it is important to note that none of them provides a satisfactory way to solve the H_0 tension without incurring in other tensions with other data sets (and none of them is currently favored over ΛCDM).

Finally, there are other potential solutions to the tension that do not directly involve cosmology. For instance, there has been proposals involving new physics on Cepheid dynamics. One example invokes a fifth force that impact the Cepheid period-luminosity relation in an environmentally-dependent manner, as in a modified gravity theory with screening mechanisms. While this possibility has not been ruled out, it would have to be adapted independently to all local measurements of H_0, which, in the case of future higher-precision measurements, may challenge this option.

From the models mentioned below, those involving varying electron mass, a majoron or a flavor of early dark energy are the most promising so far to find a solution to the H_0 tension, even if they are not completely successful or have each one their caveats.

8.1.2.1 Solutions including dark radiation

Additional number of relativistic species that mostly interact with the other components of the Universe through gravity are, effectively, dark radiation, which can also include exotic interactions with other particles. In case the additional components have some mass, they usually receive the name 'non-cold dark matter'. Additional dark radiation enhances the radiation density at early times, which increases the expansion rate and therefore introduces a degeneracy with H_0. If we capture the additional number of species in the effective neutrino number $N_{\rm eff}$, the fractional density of radiation is

$$\Omega_r = 4.18 \times 10^{-5} h^{-2} \left(\frac{T_0}{2.7255\,\mathrm{K}} \right)^4 \left(\frac{1 + \frac{7}{8}\left(\frac{4}{11}\right)^{4/3} N_{\rm eff}}{1 + \frac{7}{8}\left(\frac{4}{11}\right)^{4/3} 3.044} \right), \tag{8.1}$$

where 3.044 is the standard number for 3 neutrinos. From the expression above and the dependence of the sound horizon on Ω_r through R, we can see the effect that a higher dark radiation has on $r_{\rm s}$ and the Hubble constant. However, this perfect degeneracy is broken at the perturbation level, where dark radiation introduces a shift in the peak position and amplitude by the neutrino drag effect and modify the Silk damping. This is why models including non-free-streaming dark radiation have been explored.

Interactions between dark radiation and baryons or photons would have too large consequences on the CMB, the first possibility to try to cancel the effects of the dark radiation at the perturbation level is to include self interactions. One option is to consider a strongly self coupled relativistic fluid. The self interaction increases the clustering of the dark radiation, which reduces the Silk damping and neutrino drag and allows for larger $N_{\rm eff}$.

Nonetheless note that Big Bang nucleosynthesis, together with studies of pristine gas to infer primordial atom abundances, impose strong constraints on standard $\Delta N_{\rm eff}$ (see e.g., Ref. (28) and references therein). Therefore, any sizable $\Delta N_{\rm eff}$ large enough to solve the H_0 tension must be generated *after* Big Bang nucleosynthesis. This behavior is looked for in models involving a $\sim$ eV-scale majoron (a pseudo-Goldstone boson arising from the spontaneous symmetry breaking of a global $U(1)$ lepton number symmetry) that is produced in many neutrino models. For sufficiently large coupling between the majoron and the neutrinos, they will thermalize at a temperature dependent on the majoron mass. After thermalization, the neutrino free-streaming is effectively damped during a specific time interval until the majoron completely decays into neutrinos (which increases $N_{\rm eff}$). Together with additional dark radiation (which can come from a higher-energy-scale majoron that primordialy decayed into dark radiation), this model is well poised to solve the H_0 tension.

There are other models combining free-streaming and self interacting dark radiation (looking for higher Ω_r), or self-interacting dark radiation scattering on dark matter (which enhances and suppresses small-scale perturbations of dark radiations and dark matter, respectively, counteracting the effects of high H_0 and $N_{\rm eff}$ on the high-ℓ CMB power spectrum).

8.1.2.2 Solutions involving variations of early dark energy

Broadly speaking, early dark energy refers to a wide family of models that add a component to the Universe that drives a boost in the expansion history of the Universe, usually between matter-radiation equality and recombination. Usually it takes the form of a scalar field initially frozen in its potential by Hubble friction that, after becoming dynamical, quickly dilutes with respect to other components of the Universe.

For most early dark energy models the dynamics can be summarized as follows. First, the field is froze in its potential, such that the background energy density is constant. In this situation, the fractional contribution $f_{\rm EDE}$ to the total energy density, $f_{\rm EDE}(z) \equiv \bar{\rho}_{\rm EDE}(z)/\bar{\rho}_{\rm tot}(z)$, grows with time, until some mechanism releases the scalar, the field becomes dynamical and the background energy density dilutes faster than matter. Thus, the contribution of early dark energy to the Hubble expansion is localized in redshift and effectively reduces $r_{\rm s}$. $f_{\rm EDE} \sim 10\%$ at the peak of the contribution, at $z \sim 10^3 - 10^4$ should return a H_0 value consistent with SH0ES.

There are many models or flavors of early dark energy, including early modified gravity. In general, they can be classified in terms of the shape of the scalar field potential, the mechanism through which they become dynamical,

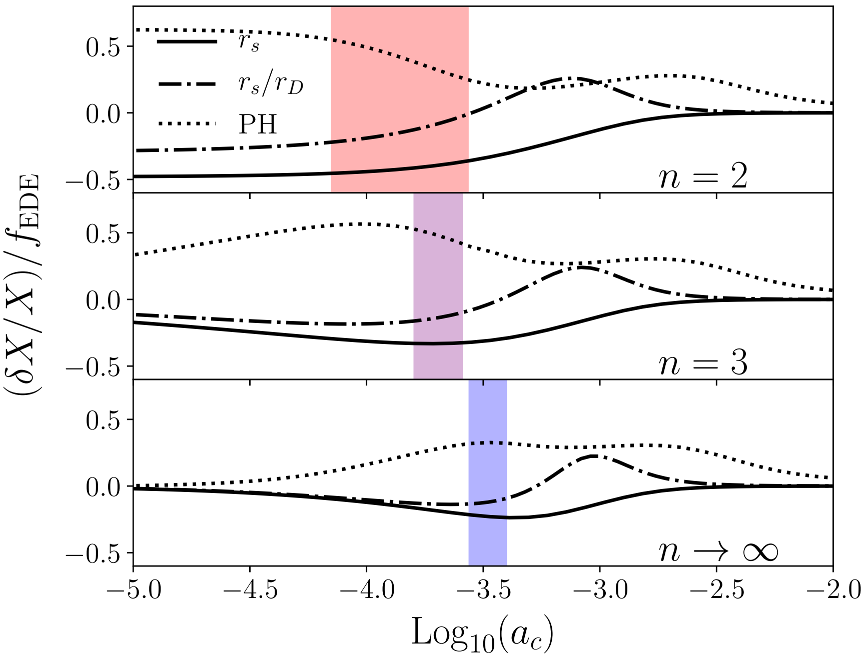

Figure 8.3 The variation of the scales that are 'fixed' by the CMB data with respect to the fraction of EDE at the maximum of its contribution as function of the moment of such maximum. All other cosmological parameters are fixed at their *Planck* best-fit values. The colored bands indicate the marginalized 1σ range of the moment of the maximum for each EDE model considered here. Figure taken from Ref. (29).

and whether or not the scalar field is minimally coupled. A recent review on the topic, with a comprehensive discussion on the phenomenology, can be found in Ref. (30). Most of these models perform similarly in light of current data; according to forecasts, future CMB experiments and galaxy surveys will provide data precise enough to discriminate between the different flavors of early dark energy.

Nonetheless, some of them have been developed to address some of the caveats of this family of models, especially a 'second coincidence problem' (why the field is relevant exactly between matter-radiation equality and recombination), or the fine tuning issues of some of the additional parameters that are hard to fit in UV-complete theories.

In any case, possibly the biggest issue that early dark energy faces is sourced from the degeneracies between the early dark energy parameters and the standard ΛCDM. In particular, early dark energy models require a higher value of n_s to fit *Planck* data, which results in an enhance power spectrum at small scales with respect to the ΛCDM prediction. This exacerbates the S_8 tension between CMB experiments and low-redshift probes. Although arguably a solution for H_0 does not need to address the S_8 tension, probes of small-scale clustering such as the 1D Lyman-α forest power spectrum could be used to test this model. First studies significantly disfavor sizable contributions from early dark energy, but further systematic checks and studies are required for a definitive answer. This caveat can be circumvented extended early dark energy models with small-scale clustering suppression.

8.1.2.3 Solutions involving shifted recombination

Rather than in the case of the dark radiation or the changes in the $H(z)$ close to recombination through some flavor of early dark energy, here we consider the family of models in which the change in the sound horizon is produced by changes in the recombination history.

Primordial magnetic fields could generate small-scale quasilinear inhomogeneities in the baryon density around recombination. In these scales, much smaller than the photon mean free path, the effective sound speed is much lower than that of a relativistic plasma, which facilitate the clustering of baryons. This inhomogeneities can change the ionization history of the Universe, which does affect the CMB anisotropies at much larger scales: this clumpier plasma recombines ealier, which reduces the sound horizon. The corresponding shift in the CMB power spectra can be counterbalanced by an increase in H_0. While promising, our ignorance about primordial magnetic fields and its impact in the components of the Universe hinder the exploration and development of these models. Three-zone models have been used to explore this possibility in an agnostic way. Interestingly, the strength of the magnetic field required to solve the tension is of the right order of magnitude to explain the existence of large-scale magnetic fields.

A similar effect can be achieved by a varying effective electron mass (or similarly the fine structure constant). Shifting the energy gap between successive excitation levels the temperature at which photon-dissociation of the hydrogen an helium becomes inefficient changes. This introduces a strong degeneracy between the redshift of recombination and these properties (broken by secondary effects like radiative transfer at recombination, two-photon decay rate, photo-ionization, recombination coefficients, Thomson scattering etc). Interestingly, varying the electron mass does not affect the Silk damping since the parameter dependence cancel. Spacetime variation in fundamental parameters are expected in theories of modified gravity or extra dimentions, but can be phenomenologically parameterized with specific relations, with a huge model space to be covered. Uniform, time-independent variations of the electron mass (in flat or curved universes) has been shown to help in increasing the H_0 value inferred from *Planck* measurements.

8.1.2.4 Last remarks

There is a remarkable effect in the analysis of most of the models mentioned above. Many of them involve the addition of new species or new effects that are canonically parameterized with a parameter that controls the amplitude or magnitude (e.g.,, $f_{\rm EDE}$) of the new effect, with the ΛCDM limit set when that parameter adopts a null value. If there are more additional parameters controlling the properties of the new species, the prior volume diverges: the new additional parameters can take *any* value if the *fraction* parameter is null. According to the Bayes theorem, the obtained posterior will therefore favor that region of the parameter space, even if the likelihood does allow higher contributions from the new physics. This effect may hinder the interpretation of the results, and have motivated frequentist approaches like the profile likelihood to complement Bayesian analyses.

Note that early-Universe modifications of ΛCDM are focused on changing $r_{\rm s}$ to keep the $r_{\rm d}H_0$ product fixed for a larger value of H_0. However, they do not modify the shape of the expansion history of the Universe at low redshift.observed It is this regime the one that dominates in the cosmic time integral

$$t(z) = \frac{977.8}{H_0}\int_0^z \frac{{\rm d}z'}{(1+z')E(z')}\,{\rm Gyr}, \tag{8.2}$$

with $H(z)$ in ${\rm km\,s^{-1}Mpc^{-1}}$. Following Eq. (8.2), the age of the Universe is $t_{\rm U} \equiv t(\infty)$. In the scenario in which only the early Universe is modified, the only impact in t_U is through H_0 (and small changes in other parameters due to the new parameter degeneracies). Therefore, independent measurements of the age of the Universe can add an important constraint to potential solutions of the H_0 tension. The oldest globular clusters can be used to infer a cosmology-independent value for the age of the Universe (31; 32): current results, limited by systematic uncertainties, agree with the inferred value of t_U from *Planck* assuming ΛCDM. If better precision can be achieved and the current result is maintained, this could hint the need to add late-time new physics to the early-time solutions of H_0 in order to be consistent with the estimated age of the Universe (22).

Finally, it is remarkable that not only the early dark energy models may have a problem with the S_8 tension. The other models do not increase the S_8 tension, but they also do not reduce it. There is no model explored so far that has been able to address both tensions at the same time with a *single* modification or new addition to the standard cosmological model. Nonetheless, although Occam's razor would favor such possibility, there is nothing preventing the combination of models that separately address each tension.

8.2 The S_8 tension

Although with lower significance, low-redshift probes of the small-scale clustering (especially the ones related with weak lensing), show a tension with respect to the prediction from the *Planck* results if ΛCDM is assumed. This tension is usually framed in the parameter combination best constrained by weak lensing, S_8, but the tension is extended to other probes which directly

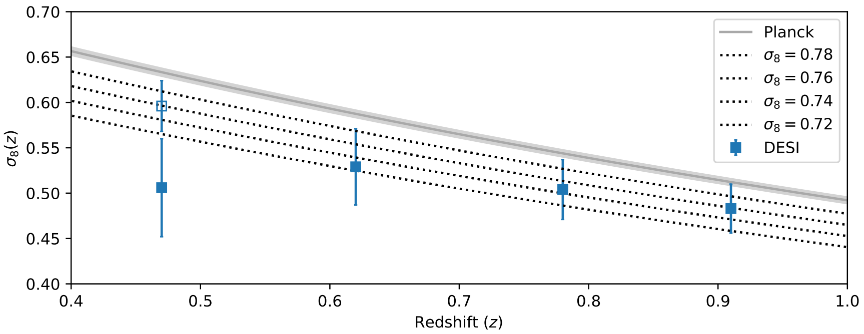

Figure 8.4 Constraints on $\sigma_8(z)$ from the cross correlation between the DESI Legacy Survey catalogs (separated in 4 redshift bins) and *Planck*. In each case the uncertainty in the ΛCDM parameters has been propagated in the prediction for $\sigma_8(z)$, showing the mean and the 68% confidence level uncertainty for *Planck* in grey. The dotted lines show the ΛCDM prediction for $\Omega_m = 0.3$ and varying values of σ_8. Figure taken from Ref. (33).

constrain σ_8. In general, the low-redshift probes manifesting (different levels of) the S_8 tension include galaxy clustering (with cosmological parameters beyond BAO and $f\sigma_8$ using full-shape analyses), galaxy clusters using the thermal SZ effect, galaxy weak lensing, CMB lensing tomography, and the cross correlations between them. A summary can be seen in Figs. 8.1 and 8.2. Interestingly, the deviation of σ_8 as measured from CMB lensing tomography from the ΛCDM prediction grows as the redshift decreases (see Fig. 8.4).

8.2.1 Features to solve the small-scale clustering tension

The situation among the different cosmological probes involved in the measurements (of inferred values) of small-scale clustering is significantly less clear than for the H_0 tension. On the one hand we have the extrapolation to low redshift from the CMB temperature and polarization power spectra by *Planck*, which lies above the measurements from cosmic shear by galaxy surveys. The highest tension on S_8 is with respect to KiDS ($\sim 3\sigma$), but there is also tension with DES; the uncertainties of HSC are large enough for the tension to be small. It is important to note, however, that cosmic shear probes from galaxy surveys depend on significantly smaller scales than CMB power spectra measurements (see Fig. 8.5). This indicates that, if the tension is due to unaccounted-for new physics, the deviations from ΛCDM must be either at redshifts after recombination or affecting only the growth of perturbations at small scales, beyond what can be probed with CMB anisotropies.

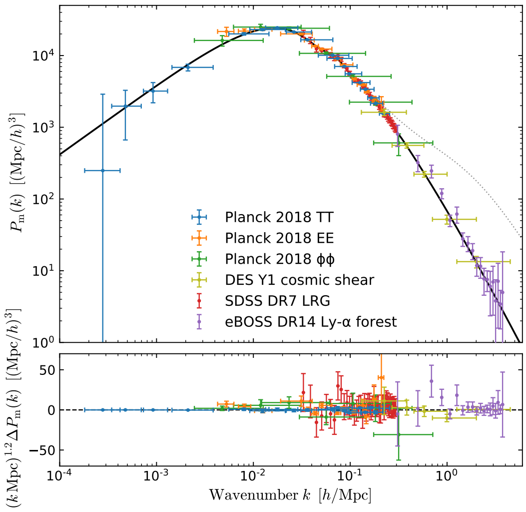

Figure 8.5 Compilation of constraints on the 3D linear matter power spectrum at $z = 0$ from *Planck* CMB power spectra on the largest scales, SDSS galaxy clustering on intermediate scales, and DES cosmic shear and SDSS Lyman-α forest clustering on the smallest scales. The solid black line in the theoretical prediction for the ΛCDM best-fit parameters of *Planck*, while the dotted line shows, for reference, the theoretical prediction for the non-linear effects. The bottom panel shows the residuals between the theoretical prediction and the constraints. Figure taken from Ref. (34).

On the other hand, CMB lensing depends on the projected gravitational potential along the line of sight (i.e., the lensing potential), as we studied in the previous chapter. Perhaps surprisingly, the CMB lensing power spectrum returns a constraint on S_8 consistent with the inferred value from the CMB temperature and polarization power spectra. Nonetheless, note that the kernel for the lensing potential for the CMB, even if extending from $z = 0$ to z_* peaks at $z \sim 2$ and has little support at lower redshifts. Furthermore, from Fig. 8.5, the CMB lensing power spectrum cannot probe scales that are accessible to cosmic shear. Therefore, although CMB lensing does not restrict small-scale deviations from ΛCDM it does push the potential deviations affecting all scales to redshifts $z \lesssim 1$.

Finally, similar to the discussion for the H_0 tension, we have seen that SNeIa+BAO strongly constrain the expansion history of the Universe at low redshift.

Therefore, from this discussion we can conclude that a potential solution to the S_8 tension must fulfill the following condition to succeed:

- Leave the expansion history of the Universe (*at least* at low redshift) untouched. In general, solutions must affect only the evolution of perturbations without modifying the background densities or the expansion rate. This already remove dark-energy related solutions from the table.

- If an early-time deviation from ΛCDM, it must only affect the growth of perturbations at small scales, beyond the access of the CMB power spectrum. Otherwise, models are ruled by *Planck* temperature and polarization anisotropies.
- If affecting a wide range of scales (i.e., overlapping with scales that are accessed by *Planck*), the deviation from ΛCDM must arise at $z \lesssim 1$. Otherwise, *Planck* lensing power spectrum rules the model out.

There are other probes that constrain σ_8 (see Fig. 8.1), but they are either at low redshifts without probing the small scales that cosmic shear can probe, such a galaxy clustering, or only probe those very small scales without a large redshift coverage, like the SZ cluster abundance.[3] In general, there is a consistent trend (although with lower significance) for these measurements to favor a lower value of σ_8 and/or S_8 with respect to the predictions from *Planck* assuming ΛCDM. Finally, 1D power spectrum from Lyman-α forest will be able to probe the small-scale power spectrum at high redshift, discriminating between the two potential ways that deviations from ΛCDM may solve this tension.

There is an alternative approach to address this tension. Rather than assuming directly a cosmological model, it is possible to consider the growth rate f (the logarithmic derivative of the linear growth factor), which can be robustly and accurately approximated as

$$f(a) = \Omega_m^\gamma(a)\,, \tag{8.3}$$

where γ is the growth index, which in flat ΛCDM with standard general relativity is predicted to be $\simeq 0.55$. See e.g., Ref (35) for a recent study using this approach. Thus, a measured deviation from this value would suggest an inconsistency between the model and observations. This approach is very similar than the agnostic parameterization of $H(z)$ to look for deviations from ΛCDM in the background expansion history.

The combination of measurements of $f\sigma_8$ and BAO from galaxy clustering, *Planck* and DES favors $\gamma > 0.55$ at $\sim 3.7\sigma$, which corresponds to a strong suppression of the perturbations at low (i.e., those for which $\Omega_m < 1$) with respect to the standard flat ΛCDM with standard general relativity. As shown in Fig. 8.6, a higher γ corresponds to a higher S_8 inferred from large-scale structure probes, and a lower value from *Planck*. This effect shows that, for a free growth index, which implies a deviation from general relativity and ΛCDM, the S_8 can be solved.

[3]Note that the tension between *Planck* power spectrum and the SZ cluster abundance depends on the prior for the cluster mass used, and the tension is only significant when a prior from gravitational lensing is used.

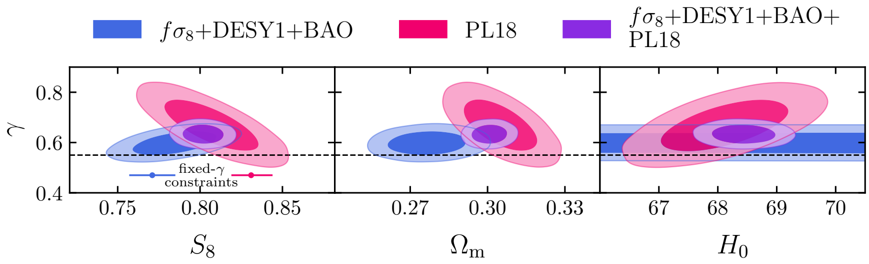

Figure 8.6 68% and 95% confidence level marginalized constraints allowing for a free growth index γ and the concordance model $\gamma = 0.55$ (marked with a dashed line). The horizontal bars in the left-most panel indicate the 68% confidence level marginalized constraints on S_8 for a fixed $\gamma = 0.55$. Figure taken from Ref. (35).

8.2.2 Beyond ΛCDM potential solutions

Given that the tension on S_8 is more recent and less significant than in the case of H_0, there has been less development at the level of model building or, at least, the exploration of the model space, for solutions of the tension. The first steps involved the exploration of massive neutrinos and energy transference in the dark sector. However, none of this models perform well because the former introduces a time-dependent suppression of the power spectrum at *all times*, while the latter modifies the background energy densities at low redshift, which changes the background expansion history. Therefore, they were disfavored by CMB lensing (and even by cosmic shear itself) and BAO+SNeIa, respectively. A review of explored models can be found in Ref. (8).

Deviations from ΛCDM that can fulfill the requirements listed in the previous section are usually limited to the dark sector. This is because if baryons are affected at early times, they will most likely significantly alter the CMB power spectrum, and if only affected at late times they have a small impact in the overall power spectrum. The exception is probably the elastic scattering between dark matter and baryons. This scattering transfer momentum between dark matter and baryons and suppresses the small-scale power spectrum, with the time evolution of the suppression controlled by the specific relative-velocity dependence in the cross section. Preliminary studies considering a relative-velocity-independent cross section, which corresponds to efficient interactions only at early times, shows indications of good performance.

A similar effect can be obtained including ultra-light axions. These particles, if light enough, transition from behaving like dark energy to behave like dark matter around or even after matter-radiation equality. In this case, they leave strong signatures in the relative peak heights (if transition occurs before matter radiation equality) or in the integrated Sachs-Wolfe effect, difussion damping and sound horizon. Axions with masses $\gtrsim 10^{-25}$ eV do not leave noticeable signatures on the CMB. In any case, all of them also suppress the power spectrum at small scales, due to an effective pressure caused by quantum effects (i.e., suppression takes place at scales of the order of the Compton wavelength for the axion). Therefore, depending on the abundance of axion and their mass, it is possible to find different kind of suppression of the power spectrum (from strong cut offs at small scales from asymptotically constant suppression at all scales).

Other models open parameter degeneracies with the standard parameters in such a way that allow for larger uncertainties in parameter constraints. Therefore, even if the mean or the best fit of the analysis does not significantly shift, the larger uncertainties can accommodate the prior on S_8 coming from cosmic shear measurements. The list of models showing these results include dark matter decaying into dark radiation and warm dark matter (which, phenomenologically, impacts cosmological observables as massive neutrinos with a time-dependent mass), cannibal dark matter (in which dark matter particles undergo a $3 \rightarrow 2$ process), or models including friction between dark matter and dark energy (coincidentally, the efficient rate for this reaction is such that coincides with the matter-dark energy equality).

Nonetheless, if we want to match the requirements for deviations from ΛCDM listed in the previous subsection with the conclusions that can be extracted from the free-γ analysis of the growth rate, we may have more information to inform the exploration of models. Since f is the logarithmic derivative of the linear growth factor, its modification affects all linear scales, a priori. Therefore, this analysis may hint a preference for models suppressing perturbations at late times, rather than only at the small scales at early times. Models that would directly modify γ are usually related with modified-gravity models (in particular within a sub-class of Horndeski models) and improvements in future observations will allow to probe them and discriminate between them.

8.2.3 Last remarks

It is important to note that cosmic shear involves nonlinear matter clustering. Even if circumventing all the complications related with nonlinear bias and redshift space distortions, pushing to small scales to obtain more information from cosmic shear observations imply a very challenging theoretical modeling for the non linearities, in particular due to the effects that baryons have in small-scale clustering: astrophysical feedback from e.g., supermassive black hole accretion or SNe explosions prevents small-scale clustering pushing the gas outside dark matter halos.

This introduces two main complications for the study of cosmic shear and the development and tests of models attempting to solve the S_8 tension. On the first hand, it limits the amount of information we can obtain, and on the other, it forces to model builders to develop the nonlinear predictions for their models if they want to use the cosmic shear likelihoods beyond a prior on S_8.

Very recently, a reanalysis of DES cosmic shear measurements included improvement in hte theoretical modeling of the data, in particular related with the impact of baryonic physics in weak lensing, which allowed to include smaller scales than ever before in the analysis (36). Standard analyses remove the scales affected by baryions (informed by hydrodynamical simulations). Instead, this analysis accounts for the effects of galaxy formation and gas physics using a baryonification algorithm on top of N-body simulations. This is meant to explicitly include the effects of baryons in a flexible way and marginalize over them. The model displaces particles from a gravity only simulation according to analytic corrections, using 7 parameters to describe the halo mass in which half of the cosmic gas graction is expelled from the halo by astrophysical processes (the parameter the observations are most sensitive to), the density profile of the gas, the galaxy-halo mass ratio, the AGN feedback range and the gas fraction-halo mass slope. Interestingly, this analysis returns a S_8 value that, while it has only slightly smaller error bars than in the standard DES analysis (they explicitly check that using the same range of scales their results only improve by $\sim 10\%$) is closer to *Planck* predictions assuming ΛCDM. Although by default DES measurements were closer to *Planck*, these results may hint that the S_8 tension, at the very least regarding cosmic shear, may be related to baryonic effects (although the impact of photo-z errors and intrinsic alignments at these scales must be studied more in detail). This result, however, would not explain the (lower significant) tensions with other cosmological probes.

Bibliography

[1] S. Dodelson and F. Schmidt, Modern Cosmology. 2020.

[2] D. Blas, J. Lesgourgues, and T. Tram, "The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes," **2011** no. 7, (July, 2011) 034, `arXiv:1104.2933 [astro-ph.CO]`.

[3] A. Lewis, A. Challinor, and A. Lasenby, "Efficient computation of CMB anisotropies in closed FRW models," The Astrophysical Journal **538** (2000) 473–476, `arXiv:astro-ph/9911177 [astro-ph]`.

[4] A. Lewis and A. Challinor, "Weak gravitational lensing of the CMB," Phys. Rept. **429** (2006) 1–65, `arXiv:astro-ph/0601594 [astro-ph]`.

[5] J. L. Bernal and E. D. Kovetz, "Line-intensity mapping: theory review with a focus on star-formation lines," Astron. Astrophys. Rev. **30** no. 1, (2022) 5, `arXiv:2206.15377 [astro-ph.CO]`.

[6] **Planck** Collaboration, N. Aghanim et al., "Planck 2018 results. VI. Cosmological parameters," Astron. Astrophys. **641** (2020) A6, `arXiv:1807.06209 [astro-ph.CO]`. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[7] A. G. Riess et al., “A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^1 Mpc^1 Uncertainty from the Hubble Space Telescope and the SH0ES Team,” Astrophys. J. Lett. **934** no. 1, (2022) L7, `arXiv:2112.04510 [astro-ph.CO]`.

[8] E. Abdalla et al., “Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies,” JHEAp **34** (2022) 49–211, `arXiv:2203.06142 [astro-ph.CO]`.

[9] A. G. Riess, S. Casertano, W. Yuan, J. B. Bowers, L. Macri, J. C. Zinn, and D. Scolnic, “Cosmic Distances Calibrated to 1% Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with ΛCDM,” Astrophys. J. Lett. **908** no. 1, (2021) L6, `arXiv:2012.08534 [astro-ph.CO]`.

[10] E. Di Valentino, O. Mena, S. Pan, L. Visinelli, W. Yang, A. Melchiorri, et al., “In the realm of the Hubble tension—a review of solutions,” Class. Quant. Grav. **38** no. 15, (2021) 153001, `arXiv:2103.01183 [astro-ph.CO]`.

[11] D. Brout et al., “The Pantheon+ Analysis: Cosmological Constraints,” Astrophys. J. **938** no. 2, (2022) 110, `arXiv:2202.04077 [astro-ph.CO]`.

[12] W. D. Kenworthy, A. G. Riess, D. Scolnic, W. Yuan, J. L. Bernal, D. Brout, et al., “Measurements of the Hubble Constant with a Two-rung Distance Ladder: Two Out of Three Ain’t Bad,” Astrophys. J. **935** no. 2, (2022) 83, `arXiv:2204.10866 [astro-ph.CO]`.

[13] C. D. Huang, A. G. Riess, W. Yuan, L. M. Macri, N. L. Zakamska, S. Casertano, et al., “Hubble Space Telescope Observations of Mira Variables in the Type Ia Supernova Host NGC 1559: An Alternative Candle to Measure the Hubble Constant,” `arXiv:1908.10883 [astro-ph.CO]`.

[14] J. P. Blakeslee, J. B. Jensen, C.-P. Ma, P. A. Milne, and J. E. Greene, “The Hubble Constant from Infrared Surface Brightness Fluctuation Distances,” Astrophys. J. **911** no. 1, (2021) 65, `arXiv:2101.02221 [astro-ph.CO]`.

[15] R. B. Tully et al., “Cosmicflows-4,” Astrophys. J. **944** no. 1, (2023) 94, `arXiv:2209.11238 [astro-ph.CO]`.

[16] D. W. Pesce et al., “The Megamaser Cosmology Project. XIII. Combined Hubble constant constraints,” Astrophys. J. Lett. **891** no. 1, (2020) L1, `arXiv:2001.09213 [astro-ph.CO]`.

[17] **LIGO Scientific, Virgo, 1M2H, Dark Energy Camera GW-E, DES, DLT40, Las Cumbres Observatory, VINROUGE, MASTER** Collaboration, B. P. Abbott et al., “A gravitational-wave standard siren measurement of the Hubble constant,” Nature **551** no. 7678, (2017) 85–88, `arXiv:1710.05835 [astro-ph.CO]`.

[18] A. Heavens, R. Jimenez, and L. Verde, “Standard rulers, candles, and clocks from the low-redshift Universe,” Phys. Rev. Lett. **113** no. 24, (2014) 241302, `arXiv:1409.6217 [astro-ph.CO]`.

[19] G. Efstathiou, “To H0 or not to H0?,” Mon. Not. Roy. Astron. Soc. **505** no. 3, (2021) 3866–3872, `arXiv:2103.08723 [astro-ph.CO]`.

[20] A. J. S. Hamilton, “Linear Redshift Distortions: a Review,” The Evolving Universe **231** (Jan, 1998) 185, `arXiv:astro-ph/9708102 [astro-ph]`.

[21] L. Verde, J. L. Bernal, A. F. Heavens, and R. Jimenez, “The length of the low-redshift standard ruler,” Mon. Not. Roy. Astron. Soc. **467** no. 1, (2017) 731–736, `arXiv:1607.05297 [astro-ph.CO]`.

[22] J. L. Bernal, L. Verde, R. Jimenez, M. Kamionkowski, D. Valcin, and B. D. Wandelt, “The trouble beyond H_0 and the new cosmic triangles,” Phys. Rev. D **103** no. 10, (2021) 103533, `arXiv:2102.05066 [astro-ph.CO]`.

[23] J. L. Bernal, L. Verde, and A. G. Riess, “The trouble with H_0,” JCAP **10** (2016) 019, `arXiv:1607.05617 [astro-ph.CO]`.

[24] T. Treu, S. H. Suyu, and P. J. Marshall, “Strong lensing time-delay cosmography in the 2020s,” Astron. Astrophys. Rev. **30** no. 1, (2022) 8, `arXiv:2210.15794 [astro-ph.CO]`.

[25] S.-F. Chen, M. White, J. DeRose, and N. Kokron, “Cosmological analysis of three-dimensional BOSS galaxy clustering and Planck CMB lensing cross correlations via Lagrangian perturbation theory,” JCAP **07** no. 07, (2022) 041, `arXiv:2204.10392 [astro-ph.CO]`.

[26] **DES, SPT** Collaboration, T. M. C. Abbott et al., “Joint analysis of Dark Energy Survey Year 3 data and CMB lensing from SPT and Planck. III. Combined cosmological constraints,” Phys. Rev. D **107** no. 2, (2023) 023531, `arXiv:2206.10824 [astro-ph.CO]`.

[27] N. Schöneberg, G. Franco Abellán, A. Pérez Sánchez, S. J. Witte, V. Poulin, and J. Lesgourgues, “The H0 Olympics: A fair ranking of proposed models,” Phys. Rept. **984** (2022) 1–55, `arXiv:2107.10291 [astro-ph.CO]`.

[28] N. Schöneberg, J. Lesgourgues, and D. C. Hooper, "The BAO+BBN take on the Hubble tension," JCAP **10** (2019) 029, arXiv:1907.11594 [astro-ph.CO].

[29] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, "Early Dark Energy Can Resolve The Hubble Tension," Phys. Rev. Lett. **122** no. 22, (2019) 221301, arXiv:1811.04083 [astro-ph.CO].

[30] V. Poulin, T. L. Smith, and T. Karwal, "The Ups and Downs of Early Dark Energy solutions to the Hubble tension: a review of models, hints and constraints circa 2023," arXiv:2302.09032 [astro-ph.CO].

[31] D. Valcin, J. L. Bernal, R. Jimenez, L. Verde, and B. D. Wandelt, "Inferring the Age of the Universe with Globular Clusters," JCAP **12** (2020) 002, arXiv:2007.06594 [astro-ph.CO].

[32] D. Valcin, R. Jimenez, L. Verde, J. L. Bernal, and B. D. Wandelt, "The age of the Universe with globular clusters: reducing systematic uncertainties," JCAP **08** (2021) 017, arXiv:2102.04486 [astro-ph.GA].

[33] M. White et al., "Cosmological constraints from the tomographic cross-correlation of DESI Luminous Red Galaxies and Planck CMB lensing," JCAP **02** no. 02, (2022) 007, arXiv:2111.09898 [astro-ph.CO].

[34] S. Chabanier, M. Millea, and N. Palanque-Delabrouille, "Matter power spectrum: from Lyα forest to CMB scales," Mon. Not. Roy. Astron. Soc. **489** no. 2, (2019) 2247–2253, arXiv:1905.08103 [astro-ph.CO].

[35] N.-M. Nguyen, D. Huterer, and Y. Wen, "Evidence for suppression of structure growth in the concordance cosmological model," arXiv:2302.01331 [astro-ph.CO].

[36] G. Aricò, R. E. Angulo, M. Zennaro, S. Contreras, A. Chen, and C. Hernández-Monteagudo, "DES Y3 cosmic shear down to small scales: constraints on cosmology and baryons," arXiv:2303.05537 [astro-ph.CO].