SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Physics Area – PhD course in

Astroparticle Physics

Testing deviations from the ΛCDM model with electromagnetic and gravitational waves

Candidate:

Dimitar Ivanov

Supervisors:

Stefano Liberati

Matteo Viel

Academic Year 2019-20

Abstract

The ΛCDM model has been extensively tested over the past decades and has been established as the standard model of cosmology. Despite its huge successes, it faces some serious theoretical problems, especially related to the nature of dark matter and dark energy. Different possible modifications and extensions have been considered in the past in order to solve these problems, and a question of central importance is how these modifications can be tested and experimentally distinguished from the standard case. The aim of this thesis is twofold. First, it presents two different modifications of the dark sector. A model in which dark matter forms a Bose-Einstein condensate in high density regions and in the process forms a non-minimal coupling to the metric is considered as a possible deviation from the cold dark matter scenario, while the possibility that dark energy is inhomogeneous in space is discussed as a possible deviation from a pure cosmological constant scenario. Second, it examines ways to test these possibilities via two main observational channels - gravitational waves and electromagnetic waves. The gravitational wave event GW170817 is used to test the dark matter model and to put constraints on the mass of the dark matter field and the strength of the non-minimal coupling. The luminosity distance and the redshift of light are highlighted as important observables for dark energy, and generalised theoretical formulae for these observables are derived for conformally FLRW and perturbed FLRW spacetimes. The luminosity distance and the redshift are finally used to test for possible anisotropies of the accelerated expansion of the universe.

Basic conventions

Throughout the thesis the signature of the metric is $\mathrm{diag}(-1, 1, 1, 1)$. We work in natural units $c = \hbar = 1$ unless stated otherwise. Also the following abbreviations are frequently used:

- DM – Dark Matter
- DE – Dark Energy
- BEC – Bose-Einstein Condensate
- ΛCDM - Λ Cold Dark Matter
- FLRW – Friedmann-Lemaitre-Robertson-Walker
- GR – General Relativity
- EFE – Einstein's Field Equation
- GW – Gravitational Wave
- GRB – Gamma-Ray Burst

Publications

This thesis is based on the following written work:

1. "Non-perturbative results for the luminosity and area distances"
   Dimitar Ivanov, Stefano Liberati, Matteo Viel, Matt Visser
   JCAP06(2018)040
   https://doi.org/10.1088/1475-7516/2018/06/040
   arXiv: 1802.06550

2. "Perturbative treatment of the luminosity distance"
   Dimitar Ivanov, Stefano Liberati, Matteo Viel, Matt Visser
   Physical Review D 98, 063505
   https://doi.org/10.1103/PhysRevD.98.063505
   arXiv: 1803.08766

3. "Testing Non-minimally Coupled BEC Dark Matter with Gravitational Waves"
   Dimitar Ivanov, Stefano Liberati
   JCAP07(2020)065
   https://doi.org/10.1088/1475-7516/2020/07/065
   arXiv: 1909.02368

4. "Testing the isotropy of the universe with future mock data"
   Dimitar Ivanov, Stefano Liberati, Matteo Viel
   In preparation

Acknowledgements

There are many people that I would like to thank. I am extremely grateful to my supervisors Stefano Liberati and Matteo Viel for the useful advice, frequent meetings and constant support during my PhD. I am indebted to my collaborators Matt Visser and Takeshi Kobayashi for their many good ideas and valuable contribution towards our projects. I would like to thank my colleagues Francesco di Filippo, Giovanni Tricella and Andrea Oddo for their help with the Italian language and for their patience while sharing an office with me, and in particular to Francesco for the many chess games that we played together. Special thanks to Raul Carballo-Rubio for his frequent help and advice, and for the many insightful discussions on various topics. I would like to thank the whole SISSA community - both academic and administrative staff, for creating an intellectually stimulating and friendly environment in which to develop and prosper. Finally, a huge 'thank you' to my parents and my brother for their moral support and their persistent faith in me.

The Earth is a very small stage in a vast cosmic arena.
Carl Sagan

Contents

1 Introduction

The Universe as we know it is about 13.7 billions of years old. Yet we have been doing science in the modern sense of the word only for the last 400 years. This is about 10^{-8} of the age of the Universe. For most of its history the Universe remained unknown to itself. Yet through its natural evolution following impersonal dynamical laws, in some regions of space the right kind of conditions were created which allowed for complexity to evolve. A small planet orbiting around a medium-sized star in the outskirts of a spiral galaxy, saw the formation of life which through billions of years of evolution through natural selection led to an astonishing variety of species. And one of these species developed the right kind of neural and cognitive system that allowed it not only to survive better and dominate over the other species, but also to explore the world, to ask deep questions about the nature of reality and ultimately to do science. The period of time for which we have been doing science is like an instant compared to the whole cosmic history. Yet during that instant we have uncovered many of the principles guiding the Universe, have established well tested mathematical theories and models, and have thus allowed the Universe to understand itself.

During that cosmic instant we have learned remarkable truths, for example that the world is made of atoms, that spacetime is curved, that all matter has wave-like properties, that the Universe used to be in a very hot dense state 13.7 billion years ago and has been expanding ever since, that we are the product of the same natural laws that guide everything that happens around us.

Yet why do we care? Why do we explore the Universe and build theories about phenomena that are completely removed from our daily lives? One possible answer is that we are simply curious - it is built into our nature to ask questions and make discoveries. But there are also deeper reasons for exploring the Universe. The Universe is our home - only through studying it carefully we can understand our place in it and how our lives fit in the grand scale of everything. Continual progress and development is what distinguishes our species from the rest of the biological world, and progress is only possible when we push hard against the boundaries of our knowledge. Furthermore, progress in our understanding goes hand in hand with technological progress, and therefore we have

to find out as much as we can about the Universe and its guiding principles, if we want to have some control over our future as a species.

Yet there seem to be limits to what we can know. We are located inside the Universe and are subject to the same laws that govern it and which, therefore, limit our freedom. We observe the Universe from one particular point of spacetime, and try to make inferences about the whole Universe just from our limited local observations. Furthermore, our direct experience can probe only a very narrow range of spatial and time scales. Fortunately, we have devised various instruments and techniques which allow us to probe a much wider range of scales - from the Planck scale to cosmological scales. Finally, our neural systems, while astonishingly complex and sophisticated, are not explicitly designed for understanding the Universe and therefore may have internal limitations.

Despite all those hindrances, our progress is astonishing. We have by now established theories and standard models in many areas of science including Biology, Chemistry, Geology, Particle Theory and Cosmology. Over the past decades our knowledge of gravity, fundamental physics and the nature and distribution of matter in the Universe has been put together in the standard model of Cosmology - the ΛCDM model. The ΛCDM model is both very simple, built on strong pillars and in remarkable agreement with the data coming from very different sources. It tells a very convincing story about the Universe from its earliest times until its distant future and fits perfectly with theories from other areas of science. Yet despite its successes it is unlikely to be the final word as far as Cosmology is concerned. It still has many unresolved issues such as the nature of the two main constituents of the Universe - dark matter and dark energy, and some other problems that call for different modifications and extensions. In the spirit of the progress made so far, we need to push the boundaries of the ΛCDM model and try to extend it until all of its problems have been solved and the agreement with the data is complete.

In this thesis we will explore some of the possible modifications of the ΛCDM and how they can be tested. We will lay the foundation of this exploration in this chapter by looking in more detail at the model, the evidence for it, its problems, possible alternatives and extensions, and how they can be tested. The upcoming chapters will look at more specific issues related to dark matter and dark energy.

1.1 The ΛCDM model

Constructing a cosmological model The ΛCDM model is one of infinitely many possible cosmological models that one can in principle construct. The construction of any cosmological model requires four main assumptions [1]:

1. A theory of gravity
2. Matter content
3. Symmetries
4. Topology

These assumptions should be based both on theoretical developments and experimental evidence. Now we will give a brief overview of what these assumptions are in the case of ΛCDM.

Theory of Gravity The gravitational dynamics in the Solar System is governed by Einstein's General Theory of Relativity. Spacetime is dynamical and the curvature of spacetime is determined by the matter content in the Universe according to Einstein's Field Equations (EFE) [2]

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.1}$$

where $g_{\mu\nu}$ is the metric, $R_{\mu\nu}$ is the Ricci tensor, $T_{\mu\nu}$ is the stress-energy tensor and Λ is the cosmological constant (more about it later). EFE can be derived by varying with respect to the metric the Einstein-Hilbert action[1]

$$S_{EH} = \frac{1}{16\pi G}\int (R - 2\Lambda)\mathrm{d}^4x. \tag{1.2}$$

The stress-energy tensor $T_{\mu\nu}$ is covariantly conserved (this is a consequence of diffeomorphism-invariance and the fact that there are no direct couplings between matter and curvature in the Einstein-Hilbert action)

$$\nabla_\mu T^{\mu\nu} = 0. \tag{1.3}$$

Both (1.1) and (1.3) need to be solved simultaneously in order to find a solution in General Relativity. In practice this is extremely difficult due to the non-linearity of the system of equations.

The ΛCDM model assumes that the EFEs (1.1) hold not only in the Solar System but on cosmological (arbitrarily large) scales. Despite the successes of General Relativity so far, this is far from obvious. General Relativity is certainly the simplest gravitational theory that we can consider. However, there exist a multitude of alternative theories which modify General Relativity on cosmological scales but reduce to General Relativity on solar-system scales. Examples of such theories are Scalar-tensor, Vector-tensor and Bimetric theories, Horava gravity, Einstein-Aether, and so on. These alternative theories are more complicated than GR in the sense that they add extra ingredients, for example extra gravitational degrees of freedom, extra terms in the Einstein-Hilbert action or extra dimensions. While GR is a useful starting assumption for cosmological model-building it need not be the final theory of gravity on large scales, and as we will see later there are benefits from exploring other possibilities in this regime.

Assumptions of symmetries and topology The observable Universe appears to be the same in all spatial directions. CMB observations and galaxy number counts bring strong support to the idea of isotropy. On the other hand, there is no reason to think

[1]One would have to add a Gibbons-Hawking-York boundary term in the action, or else perform Palatini variation.

that we occupy a special position in the Universe. Ever since Copernicus removed the Earth as the centre of the Universe, there is an increasing tendency to decrease the specialness of our place in space and time. This justifies the Copernican principle - our position in space is typical; we don't occupy a special place in the cosmos. The assumption of isotropy around the Earth, together with the Copernican principle leads to the Cosmological principle - the Universe is homogeneous and isotropic on large scales. What the Cosmological principle tells us is that among the infinite possible geometries that could describe the large-scale structure of space-time, only those for which spacetime can be sliced into maximally-symmetric patches of space with constant curvature should be taken into account. This leaves us with the FLRW metric:

$$ds^2 = -dt^2 + a^2(t)\Big[\frac{dr^2}{\sqrt{1-kr^2}} + r^2 d\Omega^2\Big], \tag{1.4}$$

where $k = \{-1, 0, 1\}$ corresponds to hyperbolic, flat or spherical spatial patches respectively.

The Cosmological Principle is a very strong assumption. It not only allows us to solve exactly the EFEs, but also allows us to make observations and draw inferences about the Universe prior to the assumption of any theory of gravity. For example, in the so-called cosmographic approach, the scale factor is expanded as a Taylor series in time [3]:

$$\begin{aligned} a(t) = a_0\{1 + H_0(t-t_0) - \frac{1}{2}q_0 H_0^2(t-t_0)^2 \\ + \frac{1}{3!}j_0 H_0^3(t-t_0)^3 + \frac{1}{4!}s_0 H_0^4(t-t_0)^4 + o[(t-t_0)^5]\} \end{aligned} \tag{1.5}$$

where the cosmographic parameters are defined as:

$$H = \frac{\dot{a}}{a}; \qquad q = -\frac{1}{H^2}\frac{\ddot{a}}{a}; \qquad j = \frac{1}{H^3}\frac{\dddot{a}}{a}; \qquad s = \frac{1}{H^4}\frac{\ddddot{a}}{a}. \tag{1.6}$$

This allows us, just on the basis of homogeneity and isotropy, to find a theoretical relation between two observable quantities - the luminosity distance d_L (measured for a special class of objects with a fixed intrinsic luminosity, the so-called standard candles) and the redshift z (measured from the absorption line spectrum of any distant object):

$$\begin{aligned} d_L(z) = \frac{z}{H_0}\Big\{1 + \frac{1}{2}\Big[1 - q_0\Big]z - \frac{1}{6}\Big[1 - q_0 - 3q_0^2 + j_0 + \frac{k}{H_0^2 a_0^2}\Big]z^2 \\ + \frac{1}{24}\Big[2 - 2q_0 - 15q_0^2 - 15q_0^3 + 5j_0 + 10q_0 j_0 + s_0 + \frac{2k(1+3q_0)}{H_0^2 a_0^2}\Big]z^3 + O(z^4)\Big\}. \end{aligned} \tag{1.7}$$

Fitting this theoretical formula with observations of standard candles allows us to constrain the parameters in (1.5) and hence to constrain the evolutionary history of the Universe or, adopting a 4-dimensional perspective, to constrain the geometry of the Block Universe.

In addition to symmetries, we need to make an assumption about the global structure of spacetime, in particular the topology. The global topology of spacetime is not fixed by neither the choice of metric, nor by EFE. Fixing the value of k in (1.4) does not fix the topology of the Universe. More precisely, if $k = 1$ the Universe is forced to be closed and finite, whereas if $k = 0$ or $k = -1$, the Universe can be closed and finite (non-simply connected) or open and infinite (simply connected). Current observations favour $k = 0$, while the issue about the topology remains unsettled.

Matter content The requirements of homogeneity and isotropy force the stress-energy tensor of all matter in the Universe to be that of a perfect fluid:

$$T_{\mu\nu} = (\rho + P)U_\mu U_\nu + P g_{\mu\nu}. \tag{1.8}$$

The energy density ρ and pressure P are functions of the cosmological time t and are related by the equation of state

$$P = w\rho \tag{1.9}$$

where w is a constant which varies from one cosmological species to another.

The covariant conservation of the stress-energy tensor (1.3) leads to the following evolution equation for the energy density:

$$\dot{\rho} + 3H(\rho + P) = 0. \tag{1.10}$$

This implies that the density of the different species scales differently with the scale factor:

$$\rho_I \propto a^{-3(1+w_I)}. \tag{1.11}$$

There are three different species of matter that enter into the ΛCDM model:

1. Dust ($w_m = 0$). This comprises all particles that move at low (non-relativistic speeds) more notably ordinary (baryonic) matter and dark matter. It leads to the following equation of state and evolution of the energy density:

$$P_m = 0, \qquad \rho_m \propto a^{-3}. \tag{1.12}$$

2. Radiation ($w_r = \frac{1}{3}$). This comprises relativistic particles such as photons, gravitons and neutrinos. It leads to

$$P = \frac{1}{3}\rho \qquad \rho \propto a^{-4}. \tag{1.13}$$

3. Dark Energy: that is any fluid that leads to an accelerated expansion $q < 0$. The most popular candidate is the cosmological constant which appears in (1.1). We can move that term to the RHS and treat it as a matter source with:

$$T^{\Lambda}_{\mu\nu} = -\Lambda g_{\mu\nu}. \tag{1.14}$$

It is easy to verify by going to Riemann normal coordinates that this is the stress-energy tensor of a perfect fluid with $w_{de} = -1$. Therefore for a cosmological constant,

$$P = -\rho \qquad \rho \propto a^0. \tag{1.15}$$

Background evolution The assumptions of homogeneity and isotropy imply that the metric is FLRW (1.4) where the only degrees of freedom are the scale factor $a(t)$ and the sign of k, while the stress-energy-tensor is that of a perfect fluid (1.8) where the only degrees of freedom are P and ρ. The EFE reduce to the Friedmann equations:

$$\Big(\frac{\dot{a}}{a}\Big)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \tag{1.16}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P). \tag{1.17}$$

The first is a constraint equation, while the second is a dynamical equation for the scale factor.

Looking carefully at the second of these equations, we notice that it predicts a decelarating Universe ($\ddot{a} < 0$) as long as the Strong Energy Condition (SEC) is satisfied $\rho + 3P > 0$. However, observations of $d_L(z)$ of standard candles indicate that the Universe is actually accelerating its expansion. That is why we need to introduce Dark Energy, an additional component which can be either a cosmological constant or something else, in order to explain this apparent acceleration. Dark Energy is any fluid which has an equation of state $w_{de} < -\frac{1}{3}$, thus violating the SEC and allowing us to fit the observations.

Evolution of perturbations Cosmological perturbations evolve according to the perturbed Einstein's equations. Their fate depends on their type (matter or radiation), their scale (sub-horizon or super-horizon) and the initial conditions. For example, matter perturbations evolve as [4]

$$\ddot{\Delta} + \mathcal{H}\dot{\Delta} - \frac{3}{2}\mathcal{H}^2\Delta = 0, \tag{1.18}$$

where Δ is the gauge-invariant density contrast (equivalent to the Bardeen potentials but for perturbations of $T_{\mu\nu}$ - see Appendix A). There are two solutions - a growing mode $\Delta_G \propto a$ and a decaying mode $\Delta_G \propto a^{-3/2}$. The growing mode leads to the amplifications of perturbations in the linear regime as structure formation begins.

There are 6 free parameters that enter into the ΛCDM model [5]. The first two are the fractional densities of baryons and cold dark matter, Ω_b and Ω_{cdm}. It is assumed that the curvature perturbations are adiabatic and the initial power spectrum can be parametrised as

$$P_{\mathcal{R}}(k) = A_s\Big(\frac{k}{k_0}\Big)^{n_s - 1}, \tag{1.19}$$

where $k_0 = 0.05 Mpc^{-1}$. The initial amplitude A_s and the primordial spectral index n_s are other two free parameters. The last two parameters are the reionisation optical depth τ and the observed angular size of the sound horizon at recombination θ_{MC}. The ΛCDM model is completely defined by these 6 parameters (most updated constraints can be found in [5]). Beside that, there are other derived parameters, and also parameters whose values are fixed within ΛCDM but can vary in alternative models (an example is w in the equation of state for dark energy).

1.2 Evidence for the ΛCDM model

There are various cosmological probes which have established the ΛCDM model as the standard cosmological model. Among these are the Cosmic Microwave Background (CMB), Baryonic Acoustic Oscillations (BAO), galaxy number counts, Big Bang Nucleosynthesis (BBN), standard candles and gravitational waves (GW). All of the ingredients of a cosmological model - matter content, theory of gravity, symmetries and topology - have been tested independently and so far all the evidence suggests that despite a few problems that are to be discussed in the next section, the ΛCDM is the simplest model that fits all the data with just 6 free parameters.

Starting with the matter content, the evidence for a cosmological constant comes from observations of standard candles such as supernovae type Ia and from the CMB [5, 6]. The Planck data also demonstrates convincingly that beside the baryonic contribution towards the matter in the universe, there is an additional non-baryonic contribution which we call dark matter. Other evidence for the existence of dark matter comes from the rotation curves of spiral galaxies, the dispersion velocities of elliptical galaxies and gravitational lensing of clusters such as the Bullet Cluster and the Musket Cluster [7, 8]. In addition, dark matter is a vital part of our best theories of structure formation. The baryonic matter of the Universe consists mostly of hydrogen and helium whose abundances are accurately predicted by Big Bang Nucleosynthesis, while heavier elements are explained by theories of stellar evolution and structure formation.

General Relativity has been tested beyond reasonable doubt on the scale of the Solar System. Experimental tests such as the deflection of light during a total eclipse, the precession of perihelia, gravitational redshift and gravitational (Shapiro) time delay have established its validity on these scales [2]. In addition, observations beyond the Solar System such as the energy loss through gravitational-wave emission from binary pulsars, the time delay between gravitational waves and gamma-ray bursts emitted by binary neutron star systems and the physics of the BBN are completely compatible with General relativity and put strong constraints on alternative theories of gravity, as we will discuss latter.

Evidence for cosmic isotropy come from the CMB (the temperature of the CMB is uniform in all directions within one part in 10^5) and from galaxy number counts. Furthermore, our best theories of structure formation assume that the early Universe is described by a FLRW background with linear perturbation on top of it, and these perturbations are enough to produce the variety of structures around us. The topology of the universe can also be tested by looking for multiple images of distant sources of light - so far there is no evidence that the universe might be non-simply connected. All the data is consistent with a flat open universe.

1.3 Problems with the ΛCDM model

Despite its successes, the standard cosmological model has some theoretical problems and some observational tensions which point to new physics beyond the ΛCDM. The

theoretical problems concern mostly the nature of dark matter and dark energy, the observational tensions concern the estimation of cosmological parameters by different cosmological probes.

Problems with Λ The simplest explanation of dark energy is that it is the cosmological constant appearing in (1.1) which, as we saw, behaves as a perfect fluid with a negative pressure. However, that explanation has two big problems that need to be resolved - the cosmological constant problem and the coincidence problem [9].

There are two aspects to the cosmological constant problem. First, if we start with the equation for the deceleration parameter in (1.6) and substitute the second Friedmann equation (1.17), we get an expression for q in terms of Ω_m and Ω_Λ:

$$\begin{aligned} q &= -\frac{1}{H^2}\frac{\ddot{a}}{a} \\ &= \frac{1}{H^2}\frac{4\pi G}{3}(\rho + 3P) \\ &= \frac{1}{H^2}\frac{4\pi G}{3}(\rho_m - 2\rho_\Lambda) \\ &= \frac{1}{2}\Omega_m - \Omega_\Lambda. \end{aligned} \tag{1.20}$$

By using that in the current age the contributions from radiation and curvature are negligible: $\Omega_{m,0} + \Omega_{\Lambda,0} = 1$, we obtain expressions for $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ in terms of q_0:

$$\Omega_{m,0} = \frac{2}{3}(1 + q_0), \qquad \Omega_{\Lambda,0} = \frac{1}{3}(1 - 2q_0). \tag{1.21}$$

The best current value of the deceleration parameter is $q_0 = -0.55$ which leads to $\Omega_{m,0} \approx 0.3$, $\Omega_{\Lambda,0} \approx 0.7$. Translating the cosmological constant density in natural units where $c = \hbar = 1$, we get

$$\rho_{\Lambda,obs} \sim (10^{-3} eV)^4. \tag{1.22}$$

The total cosmological constant in (1.1) can be split into bare cosmological constant which is simply a free parameter that needs to be determined, and vacuum energy:

$$\Lambda = \Lambda_b + \Lambda_{vac}. \tag{1.23}$$

For reasons of naturalness Λ_b is usually set to 0 or at least is assumed to be negligible compared to Λ_{vac}, while Λ_{vac} can be estimated within the Standard Model of Particle Physics. The main contribution to the vacuum energy comes from the zero-point fluctuations of all the quantum fields. This contribution diverges due to arbitrary high frequency modes, but the theory can be regularised by imposing a cut-off. A natural choice for a cut-off is the Planck scale $M_p = \frac{1}{\sqrt{8\pi G}} \approx 10^{18}$ GeV which leads to an expected value for the vacuum energy of:

$$\rho_{\Lambda,theor} \sim (10^{27}\text{eV})^4. \tag{1.24}$$

We see that

$$\frac{\rho_{\Lambda,theor}}{\rho_{\Lambda,obs}} \sim 10^{120}, \tag{1.25}$$

which is a huge discrepancy between theory and observation. This is the first aspect of the cosmological constant problem. There must be something that cancels this huge vacuum energy predicted by the Standard Model, or else we are making a mistake in the way we estimate it.

The second aspect of the cosmological constant problem is what really causes the observed accelerated expansion. It might be a residual of the vacuum energy which is imperfectly cancelled, or it might be something different such as a dynamical quintessence field or a modification of General Relativity.

The second problem with the cosmological constant is called the "coincidence problem". The relative density of the cosmological constant compared to matter evolves as

$$\frac{\Omega_\Lambda}{\Omega_m} = \frac{\rho_\Lambda}{\rho_m} \propto a^3. \tag{1.26}$$

At early times this fraction is extremely small, while at late times it is extremely large. The fact that we live in a special period of the history of the Universe where the dark energy density is comparable to the matter density is termed the "coincidence problem".

Both the cosmological constant problem and the coincidence problem urge us to rethink the nature of dark energy and to explore candidates other than the cosmological constant. (However, see [10] for an alternative point of view.)

Problems with CDM While CDM is very successful at fitting observations on large (cosmological) scales it experiences some serious problems on small (galactic) scales [7, 11]. Most of the problems arise as mismatches between CDM predictions from numerical N-body simulations and what is actually observed.

An example is the missing-satellites problem. N-body simulations of CDM predict that massive halos such as the Milky Way halo will have a large number of subhalos which host substructures such as dwarf galaxies (these substructures are referred to as satellites). However, the ones that we have detected so far are not enough to validate the CDM prediction. It might be that they are simply not bright enough to be detected, or it might be that we are missing some new physics associated with dark matter.

A closely related issue is the "too big to fail" problem. CDM predicts that the satellites of the Milky Way and other galaxies will be large and dense - too large and too dense to evade detection so far ("too big to fail" being detected).

Another difficulty with CDM is that it cannot reliably reproduce the slope (on a logarithmic plot) and the scatter of the baryonic Tully-Fisher relation:

$$M_b \propto v_c^4, \tag{1.27}$$

which is a scaling relation between the mass of baryonic mass in a galaxy M_b and the circular velocity of the galaxy v_c.

Finally, the most serious problem on small scales is the core-cusp problem. CDM simulations predict that galaxy halos have a cusp at the centre, i.e. the matter density rises sharply for decreasing radius, while observations of the velocity curves of spiral galaxies indicate that there is a core - the density profile flattens and becomes smooth. We will discuss the core-cusp problem in more detail later.

All these problems ask us to search for modifications of CDM on galactic scales.

The H_0-tension Another big problem faced by the current ΛCDM paradigm is the famous H_0-tension [12]. The problem is that the value of the Hubble constant measured from local standard candles is quite different from the value measured from the CMB, with non-overlapping error bars.

The local value of the Hubble constant is measured from variable stars with a fixed intrinsic luminosity known as Cepheids. The current best estimate is

$$H_0 = 73.24 \pm 1.74\,\mathrm{km/s/Mpc}. \tag{1.28}$$

On the other hand, the CMB can be used to fix the free parameters of the ΛCDM and hence predict from the conditions in the early Universe the value of H_0 in the late Universe. The value inferred in this way is

$$H_0 = 67.4 \pm 0.5\,\mathrm{km/s/Mpc}. \tag{1.29}$$

There is a tension of 3.6σ, which is quite significant. Possible explanations for the tension are a systematic error in at least one of the measurements (which becomes increasingly unlikely as both teams improve their techniques and seek for independent confirmation), or new physics beyond the ΛCDM.

1.4 Possible deviations from the ΛCDM model

Numerous extensions and modifications of the standard cosmological model have been considered in response to the problems presented in the previous section. They concern both components of the dark sector and generally present deviations from the idealised case of a cosmological constant and cold dark matter.

Dynamical dark energy If dark energy is truly constant, it will be very difficult to explain why it is comparable to the matter density at the current epoch. On the other hand, if it is a dynamical field, this fact can be explained more easily. The idea that dark energy is a dynamical scalar field, and therefore changes with time and possibly space is known as quintessence dark energy [13]. The basic idea is that a scalar field ϕ would have the following density and pressure:

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi), \tag{1.30}$$

$$P_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi), \tag{1.31}$$

where $V(\phi)$ is the potential of the scalar field. Therefore, the equation of state at any time is given by

$$w_\phi = \frac{P_\phi}{\rho_\phi} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}. \tag{1.32}$$

When $V(\phi) \gg \frac{1}{2}\dot{\phi}^2$, $w_\phi \approx -1$, and the scalar field behaves approximately as a cosmological constant. For other times, depending on the evolution of ϕ, there will be a deviation from the cosmological constant behaviour, but the field by itself will lead to an accelerated expansion as long as $w_\phi < -\frac{1}{3}$.

Modified equation of state of dark energy In more general terms, dark energy might have an equation of state with $w < -\frac{1}{3}$ but $w \neq -1$, or where w is a function of the scale factor and hence the redshift [14]

$$w_{de}(a) = w_0 + w_a(1-a), \qquad w_{de}(z) = w_0 + \frac{z}{1+z} w_a. \tag{1.33}$$

Such a modification would change the background evolution of the universe at late times when dark energy becomes important, and could therefore improve the H_0-tension. The reason for that is that in the CMB measurement, the value of H_0 in the late universe is reconstructed from the state of the early universe by assuming the ΛCDM model. There, the evolution of the Hubble parameter in terms of redshift is given by

$$H(z) = H_0(1+z)^{3/2}\sqrt{\Omega_m + \Omega_\Lambda(1+z)^{-3}}. \tag{1.34}$$

However, in a model where w is also a function of redshift, this evolution is modified [12]:

$$H(z) = H_0(1+z)^{3/2}\sqrt{\Omega_m + \Omega_\Lambda \exp\Big(3\int_0^z \frac{w(z')}{1+z'}dz'\Big)}. \tag{1.35}$$

This would, therefore, change the value of H_0 inferred from early universe physics. Demanding that the two values of H_0 inferred from local measurements and from CMB are in closer agreement would put constraints on the evolution of w.

Another way to improve the H_0-tension is to make modifications in the physics of the early universe. There are two possible types of modifications - changes in the details of recombination, and changes in the early-time expansion history. The former can be achieved by modifying Y_p^{BBN} - the primordial He mass fraction, while the latter can be achieved by increasing N_{eff} - the effective number of relativistic species in the early universe. See [12] for more details on these two approaches.

Acceleration from modified gravity It is well-known that several modified theories of gravity have the potential to produce an accelerated expansion without the need for dark energy [15]. Among these, most prominent are scalar-tensor theories where an extra scalar gravitational degree of freedom is added in addition to the metric. The most general scalar-tensor theory which leads to second-order equations of motion and

is free of any Laplacian instabilities and ghosts is Horndeski's theory. The Horndeski Lagrangian is a sum of four contributions:

$$\mathcal{L} = \sum_i \mathcal{L}_i, \tag{1.36}$$

where each of the four Lagrangians are given by

$$\mathcal{L}_2 = G_2(\phi, X), \tag{1.37}$$

$$\mathcal{L}_3 = G_3(\phi, X)\Box\phi, \tag{1.38}$$

$$\mathcal{L}_4 = G_4(\phi, X)R - G_{4,X}(\phi, X)[(\Box\phi)^2 - \nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi], \tag{1.39}$$

$$\begin{aligned}\mathcal{L}_5 = {} & G_5(\phi, X)G_{\mu\nu}\nabla^\mu\nabla^\nu\phi + \frac{G_{5,X}}{6}[(\Box\phi)^3 \\ & - 3(\Box\phi)(\nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi) + 2(\nabla_\mu\nabla_\nu\phi\nabla^\nu\nabla^\lambda\phi\nabla_\lambda\nabla^\mu\phi)],\end{aligned} \tag{1.40}$$

and where X is the kinetic term

$$X \equiv \frac{1}{2}\nabla_\mu\phi\nabla^\mu\phi. \tag{1.41}$$

Suppose that we choose the functions G_i to be constants and to be determined by a single mass scale μ so that all terms are dimensionally consistent. Varying this Lagrangian and expanding the resulting field equations in a FLRW background leads to Friedmann-like equations where the energy density and pressure of matter and radiation are given by the standard formulae, while the contribution of the scalar field can be collected into effective "dark energy" density and pressure that are related by an equation of state of a perfect fluid where w is negative and varies with time [16]. Thus the solution describes a Universe with accelerating expansion even though there is no cosmological constant in the theory and the strong energy condition is satisfied. In this sense, modified gravity is an alternative explanation for the acceleration that we observe, which frees us from the obligation to postulate the existence of an unknown cosmological species and gives us a better handle over the cosmological constant problem.

Fuzzy dark matter One of the most popular alternatives to CDM is a light scalar field, also known as fuzzy dark matter [17, 18]. The action takes the following form:

$$S[g] = S_{EH}[g] + S_{DM}[g, \Phi]. \tag{1.42}$$

In principle there is also a baryonic contribution, which we ignore for now. The action for the dark matter field is a sum of a kinetic term and a potential term

$$S_{DM}[g, \phi] = -\int \mathrm{d}^4x\sqrt{-g}\Big(\frac{1}{2}g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi + V(\phi)\Big), \tag{1.43}$$

where $V(\phi)$ is most often taken to be the sinusoidal potential for axions

$$\begin{aligned} V(\phi) &= m^2(1-\cos\phi) \\ &\approx \frac{1}{2}m^2\phi^2 + o(\phi^4). \end{aligned} \tag{1.44}$$

In order to be consistent with current constraints, the mass has to be of the order $m \sim 10^{-22}$eV. Varying the action leads to the Klein-Gordon equation and the EFE:

$$\nabla_\mu\nabla^\mu\phi = m^2\phi, \qquad G_{\mu\nu} = 8\pi G T^\phi_{\mu\nu}, \tag{1.45}$$

where $T^\phi_{\mu\nu}$ is the stress-energy tensor for the scalar field:

$$T^\phi_{\mu\nu} = g_{\mu\nu}\Big(-\frac{1}{2}\nabla_\rho\phi\nabla^\rho\phi - V(\phi)\Big) + \nabla_\mu\phi\nabla_\nu\phi. \tag{1.46}$$

Expanding these equations in a FLRW metric, gives the equation for a damped harmonic oscillator

$$\ddot{\phi} + 3H^2\dot{\phi} + m^2\phi = 0, \tag{1.47}$$

and the Friedmann equations with energy density and pressure given by

$$\rho_\phi = \frac{1}{2}(\dot{\phi}^2 + m^2\phi), \tag{1.48}$$

$$P_\phi = \frac{1}{2}(\dot{\phi}^2 - m^2\phi). \tag{1.49}$$

On large time- and spatial scales, one can take an oscillation average, in which case the density and pressure become

$$\rho_\phi \approx m^2\phi, \qquad P_\phi \approx 0, \tag{1.50}$$

and the field behaves as CDM. However, on time- and length scales comparable to and smaller than the period of oscillations $T = m^{-1}$ and the Compton wavelength, the pressure is non-zero and there is a modification of the CDM-like behaviour. In order to see this, it is useful to take the non-relativistic limit and to make a fluid approximation by defining a complex scalar field φ by

$$\phi = \frac{1}{\sqrt{2m}}(\varphi e^{-imt} + \varphi^* e^{imt}). \tag{1.51}$$

The Klein-Gordon equation implies that

$$\dot{\varphi} + \frac{3}{2}H\varphi = \frac{i}{2a^2m}\nabla^2\varphi. \tag{1.52}$$

Writing φ in polar form

$$\varphi = |\varphi|e^{i\theta}, \tag{1.53}$$

and defining the fluid density and velocity by

$$\rho \equiv m\varphi\varphi^*, \tag{1.54}$$

$$\vec{v} \equiv \frac{1}{am}\vec{\nabla}\theta, \tag{1.55}$$

allows to recast the dynamical equations in a fluid representation:

$$\dot{\rho} + 3H\rho + \frac{1}{a}\vec{\nabla}.(\rho\vec{v}) = 0, \tag{1.56}$$

$$\dot{\vec{v}} + H\vec{v} + \frac{1}{a}(\vec{v}.\vec{\nabla})\vec{v} = \frac{1}{a}\vec{\nabla}V_q. \tag{1.57}$$

Eqs. (1.56) and (1.57) are the continuity and Euler equation. We see that in the Euler equation there is an extra term compared to the standard perfect fluid case. The quantum potential

$$V_q = \frac{1}{2a^2m^2}\Big(\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}}\Big) \tag{1.58}$$

is a consequence of the small scale oscillations, and its effect is that it prevents the building-up of large gradients of the density thus offering a potential solution to the core-cusp problem.[2]

In addition, it provides a potential solution to the missing satellites problem. The oscillations of the scalar field cause the matter power spectrum to be suppressed on small scales compared to the standard case, thus decreasing the number of satellites that get formed and bringing us closer to the observed number. This behaviour is demonstrated in Fig. 1.1.

Dark matter as a Bose-Einstein condensate Another promising dark matter candidate that reduces to CDM on large scales, but shows deviations from the CDM-like behaviour on small scales, is a Bose-Einstein condensate [20, 21]. Consider a gas of bosons trapped in a Newtonian potential. Under certain conditions it forms a BEC and this modifies the small-scale behaviour in a way similar to fuzzy dark matter.

Working in the Heisenberg representation where the operators depend on time while the state vector is time-independent, the Hamiltonian of the system of interacting bosons is given by

$$\begin{aligned}\hat{H}(t) = &\int d^3r\hat{\Psi}^\dagger(\vec{r},t)\Big[-\frac{\hbar^2}{2m}\nabla^2 + V_g(r)\Big]\hat{\Psi}(\vec{r},t) \\ &+ \frac{1}{2}\int d^3r d^3r'\hat{\Psi}^\dagger(\vec{r},t)\hat{\Psi}^\dagger(\vec{r}',t)V(\vec{r}-\vec{r}')\hat{\Psi}(\vec{r},t)\hat{\Psi}(\vec{r},t),\end{aligned} \tag{1.59}$$

[2]Note that there are some claims in the literature that fuzzy dark matter might actually be incompatible with observations of galactic rotation curves [19].

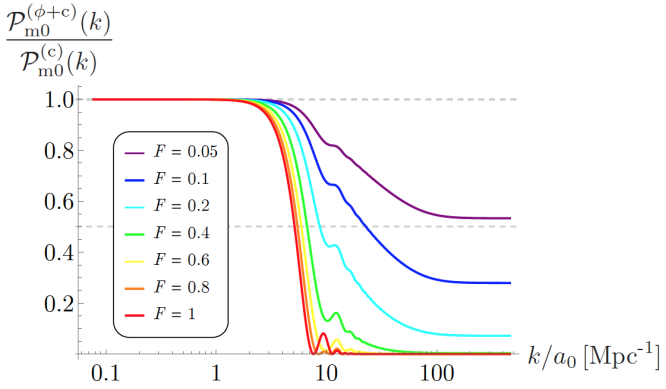

Figure 1.1: The linear matter power spectrum for a light scalar dark matter field of mass $m \sim 10^{-22}$eV, for various values of $F \equiv \frac{\rho_\phi}{\rho_{dm}}$ - the ratio of the scalar dark matter density to the total dark matter density. As F increases, the power spectrum gets more and more suppressed at small scales (large wavenumbers), thus offering a potential solution to the missing satellites problem. Taken from [18].

where $\hat{\Psi}^\dagger(\vec{r},t)$ is the boson creation operator, $\hat{\Psi}(\vec{r},t)$ is the boson annihilation operator, $V_g(r)$ is the gravitational potential, and $V(\vec{r}-\vec{r}')$ is the inter-bosonic potential. The field operator $\hat{\Psi}$ can be split into a condensate wavefunction plus quantum excitations:

$$\hat{\Psi}(\vec{r},t) = \psi(\vec{r},t)\hat{\rm I} + \hat{\varphi}(\vec{r},t). \tag{1.60}$$

The number density is defined in terms of the condensate wavefunction and it is normalised in such a way so that it integrates to the total number of particles in the condensate phase:

$$\rho(\vec{r},t) = |\psi(\vec{r},t)|^2, \qquad \int \rho d^3r = N. \tag{1.61}$$

The dynamics of the Bose field is given by the Heisenberg equation of motion

$$i\hbar\frac{\partial}{\partial t}\hat{\Psi} = [\hat{\Psi}, \hat{H}], \tag{1.62}$$

and for a rare, low-density gas of bosons the inter-bosonic interactions can be approximated by point interactions:

$$V(\vec{r}-\vec{r}') = \lambda\delta(\vec{r}-\vec{r}'), \tag{1.63}$$

where the coupling constant λ is given by

$$\lambda = \frac{4\pi\hbar^2 a}{m}, \tag{1.64}$$

where a is the scattering length. Substituting (1.60) into (1.62), we obtain the Gross-Pitaevski equation for the evolution of the condensate wavefunction:

$$i\hbar\frac{\partial\psi}{\partial t} = \Big[-\frac{\hbar^2}{2m}\nabla^2 + V_g(\vec{r}) + \lambda|\psi|^2\Big]\psi. \tag{1.65}$$

The gravitational dynamics is given by the Poisson equation

$$\nabla^2 V(r) = 4\pi G \rho_m, \tag{1.66}$$

where the mass density is defined by

$$\rho_m = \rho m. \tag{1.67}$$

As before, we split the condensate wavefunction in amplitude and phase, which allows us to define the fluid velocity

$$\psi = \sqrt{\rho} e^{\frac{i}{\hbar} S}, \qquad \vec{v} = \frac{1}{m} \vec{\nabla} S. \tag{1.68}$$

Substituting that into (1.65) gives the continuity equation, and the Euler equation

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v}.\vec{\nabla})\vec{v} = -\frac{\lambda}{m} \vec{\nabla} \rho - \frac{1}{m} \vec{\nabla} V_g - \frac{1}{m} \vec{\nabla} V_q. \tag{1.69}$$

Comparing that to the identical equation for fuzzy dark matter, we see that they are very similar, except that this equation is valid in a flat spacetime, has a Newtonian gravitational potential included, and also an extra term on the RHS proportional to the gradient of ρ. This extra term comes from the gradient of the pressure which arises due to the interactions of the bosons:

$$P(\rho) = \frac{1}{2} \lambda^2 \rho^2. \tag{1.70}$$

We are interested in a static fluid for which $\vec{v} = 0$. In that case, the continuity equation implies that ρ is independent of time, while the Euler equation gives

$$\lambda \vec{\nabla} \rho = -\vec{\nabla} V_g - \vec{\nabla} V_q. \tag{1.71}$$

Using the Poisson equation and rewriting everything in terms of the mass density ρ_m gives

$$\frac{\lambda}{m} \nabla^2 \rho_m + \nabla^2 V_q = -4\pi G \rho_m. \tag{1.72}$$

This equation should be solved in order to find the equilibrium density profile. The solution is

$$\rho_m(r) = \rho_c \frac{\sin(kr)}{kr}, \tag{1.73}$$

where ρ_c is the density at the centre of the halo, and $k = \sqrt{\frac{Gm^3}{\hbar^2 a}}$. We see that the cusp of the standard CDM scenario is removed, and instead there is a smooth core. The radius of the halo can be defined in terms of the first zero of the profile (1.73) when $kr = \pi$, which gives

$$R = \pi \sqrt{\frac{\hbar^2 a}{Gm^3}}. \tag{1.74}$$

Dark matter and modified gravity While it is possible that modified gravity might account for the accelerated expansion without the need for dark energy, it is more controversial whether it can also account for dark matter. In fact, evidence for dark matter comes from several different directions, and it seems difficult, if not impossible, to explain all of these by simply modifying the gravitational dynamics and without introducing a new kind of particle. Nevertheless, some of the observations on small scales associated with Dark Matter can be accounted for by Modified Newtonian Dynamics (MOND), which modifies Newton's theory of gravity below an acceleration scale a_0 [22]:

$$\vec{\nabla}.\Big[\mu(r)\Big(\frac{|\vec{\nabla}\phi|}{a_0}\Big)\Big] = 4\pi G\rho(r), \tag{1.75}$$

where $\mu(r)$ is a function of the radial coordinate which is linear for small r and flattens to a constant value for large r. While MOND fits better than CDM some of the observations on small scales, it completely fails on large scales and thus is not a viable alternative of dark matter.

However, its partial success on small scales has led to the idea that maybe there is something peculiar about the coupling between dark matter and the metric which modifies the CDM-like behaviour on small scales and leads to a MOND-like phenomenology. This idea was pursued in [23, 24]. Starting from the general action

$$S[g] = S_{EH}[g] + S_m[g,\Psi] + S_{DM}[g,\Phi] + S_{int}[g,\Psi,\Phi], \tag{1.76}$$

the argument is that in order to explain the success of MOND, dark matter must have a universal effect on baryons i.e. it must provide an effective metric along which baryons propagate. Therefore,

$$S_m[g,\Psi] + S_{int}[g,\Psi,\Phi] \approx S_m[g+h,\Psi]. \tag{1.77}$$

Here $\tilde{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu}$ is the physical metric to which normal matter couples minimally (i.e. the metric in the Jordan frame), which is different from the gravitational metric $g_{\mu\nu}$ (the metric in the Einstein frame) which appears in EFEs (in GR both metrics coincide). It is argued elsewhere [25] that the most general transformation that preserves causality and respects the weak equivalence principle is the disformal transformation

$$\tilde{g}_{\mu\nu} = A(\phi)g_{\mu\nu} + B(\phi)\nabla_\mu\phi\nabla_\nu\phi. \tag{1.78}$$

It is thus safe to assume that $h_{\mu\nu} = \nabla_\mu\phi\nabla_\nu\phi$. From (1.77), by expanding the RHS to first order in h and using the definition of the stress-energy tensor

$$T_m^{\mu\nu} \equiv \frac{2}{\sqrt{-g}}\frac{\delta S_m[g_{\mu\nu},\Psi]}{\delta g_{\mu\nu}}, \tag{1.79}$$

one can show that

$$S_{int}[g,\Psi,\Phi] = \frac{1}{2}\int \mathrm{d}^2x\sqrt{-g}T_m^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi. \tag{1.80}$$

Using EFEs and writing everything in terms of the physical metric $\tilde{g}_{\mu\nu}$, this becomes

$$S_{int}[\tilde{g}, \Psi, \Phi] = -\frac{1}{16\pi G} \int \mathrm{d}^4x \sqrt{-\tilde{g}} G^{\mu\nu}[\tilde{g}] \nabla_\mu \phi \nabla_\nu \phi. \tag{1.81}$$

The advantage of such a non-minimal coupling is that it causes baryons to propagate on an effective metric different from the one on which dark matter propagates, and this leads to an effective MOND-like phenomenology on small scales. Since MOND fits the density profiles of most galactic systems much better than CDM and in particular gives cored profiles, this is an improvement compared to the standard CDM scenario, which furthermore requires only a minimal modification of the gravitational interactions of dark matter.

Challenging the Cosmological Principle All possible deviations from ΛCDM so far involved a modification either of the matter content of the universe or the dynamical theory of gravity. However, it is also possible to challenge the Cosmological Principle - the assumption that the universe is spatially homogeneous and isotropic on large scales. In fact there are several studies which show that the Universe might have a preferred spatial direction [26–28].

Furthermore, there are cosmological models which violate the Cosmological Principle but are so far able to fit the data for some regions in their parameter space as well as ΛCDM. Such models are for example Bianchi models, Lemaitre-Tollman-Bondi models, and Swiss cheese models [1]. While the ΛCDM is preferred on the basis of naturalness and simplicity, these models are valid alternatives which cannot be excluded on the basis of the current data available.

1.5 Testing deviations from the ΛCDM model

Now that we have presented several possible modifications of the standard ΛCDM paradigm, an interesting and very important question is whether these modifications can be tested and experimentally distinguished from the ΛCDM. The most promising ways to test them come from two main directions - light observations and gravitational wave observations. Important observables here are the luminosity distance and the redshift for light, and the luminosity distance and speed for gravitational waves. These observables can give important information about possible deviations from the standard model.

GW170817 and its implications for Horndeski dark energy One major test of modified theories of gravity as a candidate to explain the accelerating expansion of the Universe comes from the detection of a gravitational signal by the advanced LIGO and advanced Virgo gravitational-wave detectors combined with the detection of a gamma-ray burst by the Fermi-GBM about 1.7 s later [29]. Both signals are identified as originating from a binary neutron star merger event about 40 Mpc (in luminosity distance) away from us. Since in alternative theories of gravity gravitational waves propagate with a modified speed compared to the speed of light, the short arrival time

difference between the two waves allows to put stringent constraints on modified gravity theories which attempt to explain the accelerated expansion without the need for dark energy.

The speed of gravitational waves in any modified theory of gravity can be written as [30]

$$c_g^2 = 1 + \alpha_g, \tag{1.82}$$

where α_g parametrises deviations from the speed of light. The predicted time delay that would accumulate between the arrival times of gravitational and electromagnetic waves can be written to first order in α_g as

$$\Delta t = \frac{2\alpha_g}{\ell}, \tag{1.83}$$

where ℓ is the distance from the Earth to the emission event. The actual arrival time difference is about 1.7s which translates into a constraint on α_g:

$$\Delta t \lesssim 1.7s \implies |\alpha_g| \lesssim 10^{-15}. \tag{1.84}$$

For the Horndeksi action (1.36)

$$\alpha_g = \frac{X[2G_{4,X} + 2G_{5,X} - (\ddot{\phi} - \dot{\phi}H)G_{5,X}]}{G_4 - 2XG_{4,X} - XG_{5,\phi} - \phi HXG_{5,X}}. \tag{1.85}$$

If we set $\alpha_g = 0$, then in order to avoid any unnatural fine-tuning, we must necessarily have $G_{4,X} = G_{5,\phi} = G_{5,X} = 0$. This implies that $G_5 = const.$ and $G_4 = f(\phi)$, and therefore $\mathcal{L}_4 = f(\phi)R$, $\mathcal{L}_5 = 0$. The only non-minimal coupling that survives is the Brans-Dicke coupling $f(\phi)R$. Thus we see that GW170817 puts strong constraints on Horndeski dark energy.

Luminosity distance measurements from standard candles In the standard treatment, data from supernovae of Type Ia and other standard candles is fitted by using a luminosity distance–redshift relation of the type (1.7), i.e. a theoretical expression for the luminosity distance which is derived by assuming that light propagates on a perfectly homogeneous and isotropic background. While this is a good approximation on large scales, it fails to be true on small scales. Therefore the light from distant stars could be sensitive to the small-scale lumps and inhomogeneities as it propagates towards us. Moreover, the assumptions of isotropy and homogeneity fail to hold in some alternative cosmological models which call for a more drastic deviation from the Cosmological Principle. In order to be able to investigate the role of inhomogeneities from cosmic structure in the cosmographic approach and also to test alternative cosmologies, it is necessary to have a more general expression for the luminosity distance with which to fit the data.

There have been numerous attempts to derive a generalised $d_L(z)$ relation ever since the paper of Sasaki [31]. In that paper, under suitable conditions, the following formula

for the luminosity distance in a perturbed geometry was derived:

$$d_L(z,\lambda_s) = \bar{d}_L(z)\left[1 + \left(\frac{a'}{a}\delta\eta\right)_o + \coth\left(\sqrt{-k}\lambda_s\right)\sqrt{-k}\delta\lambda_s - \frac{1}{2}\int_0^{\lambda_s}\delta\theta(\lambda)d\lambda\right], \quad (1.86)$$

where $\bar{d}_L(z)$ is the luminosity distance evaluated at the background, while $\delta\eta$, $\delta\lambda$ and $\delta\theta$ are the perturbations of the conformal time, the affine parameter and the expansion. Further progress was made in [32]. Generalised formulas for d_L (or some function of it, such as the magnitude or the fractional fluctuation) have also been derived in [33–37]. In [35] the authors compute the two-point correlation function of the luminosity distance while in [37] the authors compute the luminosity distance to second order in perturbations in the geodesic lightcone gauge and then transform to the Poisson gauge.

1.6 Motivation and overview of the present research

The present thesis aims to make a small contribution towards our understanding of the physics beyond the ΛCDM model.

Main contributions We will first discuss a slightly different model of dark matter as a Bose-Einstein condensate, in which a non-minimal coupling of the the type (1.81) forms during the process of condensation and its effect is added to the effect of the quantum pressure term. The non-minimal coupling opens a door to testing the model, as it modifies the speed of gravitational waves with respect to the speed of light. As a result, we can use the observed time delay between the gravitational wave and gamma ray burst signal in GW170817 to put a constraint on the model. We will then turn to the other side of the dark sector - dark energy. One of the best probes of dark energy is the luminosity distance and we will first discuss how it can be expressed in terms of a mathematical object known as the van Vleck determinant. We will then proceed to derive theoretical formulae for the luminosity distance in some spacetimes of cosmological interest and demonstrate how they can be used to do generalised cosmography. One of the most exciting deviations from the ΛCDM model is that dark energy might be inhomogeneous and the expansion of the universe can be anisotropic. However, such a deviation is very hard to test with present data. Still we can say something about the quality of future data that would be necessary in order to distinguish this scenario from the standard one.

Outline The structure of this thesis is the following. In Chapter 2 we start by discussing the model of dark matter as a BEC with a non-minimal coupling. We first look at the core-cusp problem of CDM, then we present the model and show how it can help to improve the core-cusp problem. In Chapter 3 we look at how the model can be tested with GW170817. We first discuss gravitational wave propagation in modified theories of gravity with a non-minimal coupling, then we parametrise the uncertainties of the BEC model and derive a formula for the predicted arrival time difference between the two waves. Comparing this with the actual time delay allows us to constrain the parameters

of the model. In Chapter 4 we turn to the notion of a cosmological distance and how it can help us to probe the expansion history of the universe. We present different definitions of a cosmological distance, and then focus on the luminosity distance. We discuss the Jacobi map and the van Vleck determinant, and how the luminosity distance can be expressed in terms of them. Finally, we discuss the cosmographic approach and why it is so useful. In Chapter 5 we continue with the discussion from the previous chapter by using the expression for the luminosity distance in terms of the van Vleck determinant and the Jacobi map in order to derive theoretical formulae for the luminosity distance in Conformally FLRW cosmologies and perturbed FLRW cosmologies. We also show how to generalise the cosmographic approach. Chapter 6 is more data analysis oriented - we perform tests of the isotropy of the universe with future forecast data. We first consider a very simple beyond-the-ΛCDM model where dark energy undergoes a phase transition and becomes anisotropic at late times. We derive an expression for the luminosity distance in this model and then investigate the quality of future data needed to distinguish this model from the ΛCDM. Finally, we conclude in Chapter 7.

2

Dark matter Bose-Einstein condensates

In this chapter we begin our investigation of models beyond the ΛCDM by considering the problems faced by the CDM paradigm and focussing on a particular alternative model in which dark matter is a scalar field which forms a Bose-Einstein condensate in high-density astrophysical regions, and in the process acquires a non-minimal coupling to the metric.

2.1 The core-cusp problem

As explained in the Introduction, the standard CDM paradigm faces some serious challenges on small (galactic) scales. Among these problems the most serious seems to be the core-cusp problem which arises from a mismatch between the equilibrium density profile of galaxy halos predicted with numerical N-body simulations and the actual observed density profiles of typical galaxies.

Numerical simulations with CDM predict that the density profile of a virialised dark matter halo is given by the Navarro-Frenk-White (NFW) profile [8, 38]

$$\rho_{NFW}(r) = \frac{\rho_s}{\frac{r}{r_s}(1+\frac{r}{r_s})^2}, \tag{2.1}$$

where r is the radius from the centre of the halo and ρ_s and r_s are constants specific for the galaxy of interest (thus they are free parameters in the simulation).

The NFW profile has a cusp at the centre - a region where the density grows without limit (Fig. 2.1). However, observations favour density profiles with a core at the centre. For example, the observed circular velocities in spirals, dwarf disks and low surface brightness systems are well fitted by an empirical Burkert profile [8, 39]

$$\rho_B(r) = \frac{\rho_0}{(1+\frac{r}{r_0})(1+(\frac{r}{r_0})^2)}, \tag{2.2}$$

where r_0 is the radius of the core and ρ_0 is the density at the centre (Fig. 2.2). Another profile that fits well observations of galaxy halos is the Einasto profile [40]

$$\rho_E(r) \propto \exp(-Ar^\alpha), \tag{2.3}$$

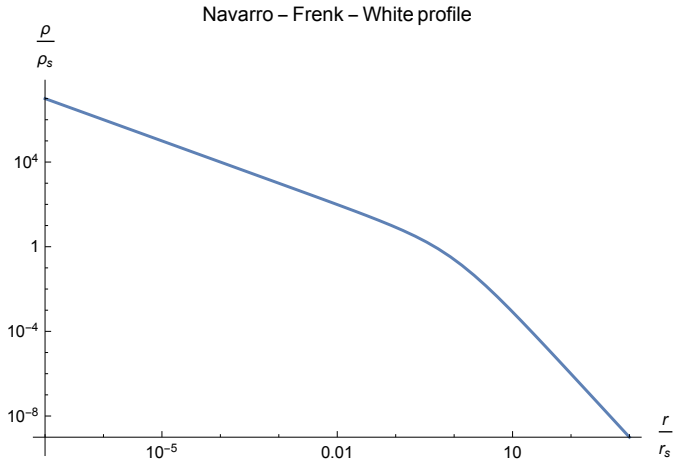

Figure 2.1: The Navarro-Frenk-White (NFW) density profile. The density grows indefinitely for small radii.

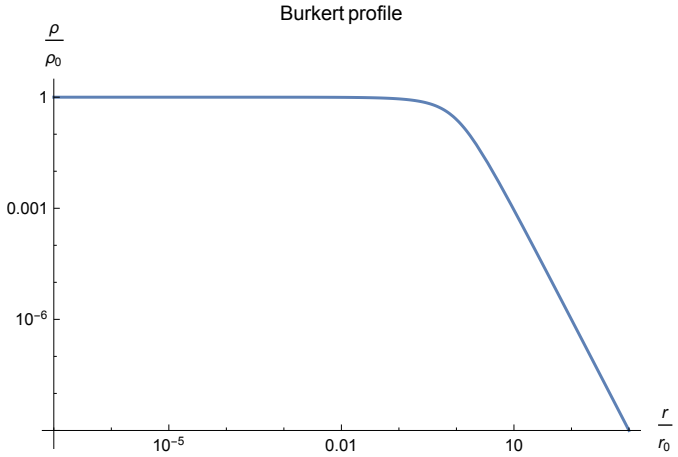

Figure 2.2: The Burkert density profile. The density tends to a constant value for small radii.

where A and α are constants. In either case the actual density towards the centre of the halo is much lower than the theoretically expected one. This discrepancy between CDM simulations, which predict a cusp, and observations, which favour a core, is called the core-cusp problem.

There are different approaches that attempt to solve or at least ameliorate the core-cusp problem (some of them were already discussed in the Introduction). One such approach is to introduce a baryon-feedback onto CDM particles i.e. to introduce interactions between baryons and dark matter. Another approach is to modify the gravitational dynamics on galactic scales (MOND) and thereby completely get rid of the need for dark matter. Yet another class of approaches is to consider dark matter as a very

light (low-mass) scalar field. Examples of the latter are fuzzy dark matter where the scalar field oscillates harmonically and this leads to modifications of the CDM-like behaviour on small length scales and BEC dark matter where the scalar field forms BEC in galactic regions and thus modifies its small-scale behaviour. One of the most attractive BEC approaches is the one where dark matter is modelled as a fully relativistic scalar field consistent with General Relativity and the non-relativistic limit is taken at the end. The formation of a macroscopic coherence length (the healing length) leads to the conjecture that there is a direct (non-minimal) coupling between dark matter and the metric. In this chapter we will review this model and consider how it could help to solve the core-cusp problem.

2.2 Non-minimally coupled BEC model

A relativistic Bose-Einstein condensate (BEC) is a relativistic gas of bosons in which most of the particles have undergone a phase transition and are in a condensate phase characterised by a macroscopic occupation number. The starting relativistic action in flat spacetime is

$$S = -\int \left[\frac{1}{2}\partial_\mu \hat{\phi}^\dagger \partial^\mu \hat{\phi} + \frac{1}{2}m^2 \hat{\phi}^\dagger \hat{\phi} + U(\hat{n}, \lambda)\right] d^4x, \tag{2.4}$$

where $\hat{\phi}(\vec{x}, t)$ is a relativistic scalar Bose field operator, $\hat{n}$ is defined by $\hat{n} = \hat{\phi}^\dagger \hat{\phi}$ (in the non-relativistic theory this has the interpretation of a number density, while in the relativistic theory it is simply a quantity related to the mass density $\hat{\rho}$ by $\hat{\rho} = m^2 \hat{n}$), λ is a dimensionless coupling constant and $U(\hat{n}, \lambda)$ is a self-interaction term of the form

$$U(\hat{n}, \lambda) = \frac{\lambda}{2}\hat{n}^2. \tag{2.5}$$

One could also add an external potential $V(\vec{x}, t)$ to the action (2.4). The condensation occurs below a critical temperature T_c. In the condensate phase the Bose field can be split into a classical complex scalar field ϕ describing the ground state, and quantum excitations $\hat{\varphi}$:

$$\hat{\phi} = \phi(\hat{1} + \hat{\varphi}). \tag{2.6}$$

The field ϕ obeys a relativistic Klein-Gordon equation of the form

$$\Box\phi - m^2\phi - U'\phi = 0, \tag{2.7}$$

where the prime denotes derivative with respect to the number density defined above. In the non-relativistic limit this reduces the Gross-Pitaevski equation (1.65) (with the difference that here the potential U does not have to be the one for point interactions) and for this reason ϕ is often called the "wave function" of the condensate. It is useful to express the condensate wave function and its complex conjugate in terms of the

hydrodynamic variables $\rho = m^2\phi^*\phi$ and u^μ in the so-called Madelung representation (this mirrors the procedure performed in the non-relativistic case):

$$\phi \equiv \frac{1}{m}\sqrt{\rho}e^{i\theta}, \tag{2.8}$$

$$u^\mu \equiv \frac{1}{m}\nabla^\mu\theta. \tag{2.9}$$

The 4-velocity u^μ is in general not normalised. One can instead define the normalised velocity $v^\mu \equiv u^\mu/\sqrt{-u^\mu u_\mu}$ but for our purposes this is not necessary and we will keep using u^μ in the subsequent formulae. [1] With these definitions the Klein-Gordon equation can be shown to be equivalent to two equations in terms of ρ and u^μ:

$$\nabla_\mu(\rho u^\mu) = 0, \tag{2.10}$$

$$u^\mu u_\mu = -1 + \Big(- 2U^{'}(\rho) + \frac{1}{m^2}\frac{\Box\sqrt{\rho}}{\sqrt{\rho}}\Big) \tag{2.11}$$

where $U'(\rho)$ is the derivative of the self-interaction potential U with respect to the mass density ρ and $V_q \equiv \frac{1}{m^2}\frac{\Box\sqrt{\rho}}{\sqrt{\rho}}$ is the quantum potential. The first equation is the continuity equation while the second is an equation for the norm of u^μ. It shows that the 4-velocity u^μ is generally not normalised but that for CDM for which both $U(\rho) = 0$ and $V_q = 0$ it reduces to the normalised velocity. Eqn. (2.11) can be turned into a dynamical equation by taking the covariant derivative:

$$u^\mu\nabla_\mu u_\nu = -\partial_\nu\left[U'(\rho)\Big(1 - \xi^2\frac{\Box\sqrt{\rho}}{\sqrt{\rho}}\Big)\right], \tag{2.12}$$

where $\xi^2 \equiv 1/(2m^2U'(\rho))$ is the healing length which is the length scale over which the density of the condensate returns to its bulk value when perturbed locally. Notice the appearance of a third derivative on the RHS of the equation due to the extra covariant derivative that we took.

If we consider a uniform square potential well, the density must vanish at the boundary and become constant towards the centre of the square well. In that case, the healing length would correspond to the length scale over which the density rises from 0 to that constant value (see [41] for more details). If the typical length and time scales of variation of the density ρ are much larger than ξ, then we can ignore the quantum pressure term in (2.12), in which case (2.10) and (2.12) are completely equivalent to the continuity and Euler equations of a relativistic perfect fluid.

Now we want to generalise the above formalism from flat to curved spacetime. We saw that the process of condensation is characterised by the formation of a typical length

[1] The quantity ρ is obtained by taking the vacuum expectation value of the operator $\hat{\rho} = m^2\hat{n}$ and can be interpreted as the rest mass density in the BEC frame. One can check that this is the correct interpretation by deriving the stress-energy tensor of the scalar field and contracting it with v^μ in order to obtain the energy density.

scale - the healing length ξ. If this length scale is comparable to the length scale set by the curvature of spacetime, i.e. the length scale over which the geometry of spacetime starts deviating from Minkowski, then the condensate would become sensitive to the global geometry. Thus one would expect the formation of a non-minimal coupling - a direct coupling between the curvature and the Bose field with a coupling constant directly related to the healing length of the BEC. Unfortunately, in the absence of a rigorous mathematical theory of condensation in a curved spacetime, one cannot say for sure what the form of this coupling would be. However, we can impose several requirements that narrow the range of possible couplings.

The requirement to have second order field equations leads to a coupling which is a subclass of the Horndeski Lagrangian (1.36). We also require that the coupling contains a dimensionful coupling constant (thus ϕR would not work for example) in order to justify why this coupling is present for DM but not for baryons (because baryons have no macroscopic coherence length scale, while DM has - the healing length). Furthermore, as we discussed in the Introduction, a coupling of the form $G_{\mu\nu}\nabla^{\mu}\phi\nabla^{\nu}\phi$ in the Jordan frame leads to an effective modified gravity behaviour on small scales which, in the Einstein frame, causes baryons to propagate on an effective metric different from that on which DM propagates and thus leads to an effective MOND-like phenomenology. Based on this reasoning it was conjectured in [42] that the non-minimal coupling developed during the process of condensation is $L^2G_{\mu\nu}\nabla^{\mu}\phi^{\dagger}\nabla^{\nu}\phi$ where L is a coupling constant with dimensions of length. This coupling satisfies all the above requirements and in Appendix B we show that it is equivalent to a subpart of the Horndeski Lagrangian (1.36).

Thus the action of the Bose field in the condensate phase is taken to be a sum of three terms

$$S = S_{EH} + S_{\phi} + S_{NMC}, \tag{2.13}$$

where

$$S_{EH} = \frac{1}{16\pi G}\int R\sqrt{-g}d^4x \tag{2.14}$$

is the standard Einstein-Hilbert term,

$$S_{\phi} = -\int\left[\frac{1}{2}g^{\mu\nu}\nabla_{\mu}\hat{\phi}^{\dagger}\nabla_{\nu}\hat{\phi} + \frac{1}{2}m^2\hat{\phi}^{\dagger}\hat{\phi} + U(\hat{n},\lambda_i)\right]\sqrt{-g}d^4x \tag{2.15}$$

is the minimally coupled part of the Bose field and

$$S_{NMC} = \int L^2G_{\mu\nu}\nabla^{\mu}\hat{\phi}^{\dagger}\nabla^{\nu}\hat{\phi}\sqrt{-g}d^4x \tag{2.16}$$

is the non-minimally coupled part of the Bose field. Varying the above action with respect to $g^{\mu\nu}$ gives the gravitational field equation

$$G_{\mu\nu} = 8\pi G\big[T^{\phi}_{\mu\nu} + T^{NMC}_{\mu\nu}\big], \tag{2.17}$$

where $T^{\phi}_{\mu\nu}$ is the stress-energy tensor associated with the minimally coupled field[2]

$$T^{\phi}_{\mu\nu} = \frac{1}{2}\nabla_\mu\hat{\phi}^\dagger\nabla_\nu\hat{\phi} + \frac{1}{2}\nabla_\mu\hat{\phi}\nabla_\nu\hat{\phi}^\dagger - g_{\mu\nu}(\frac{1}{2}g^{\alpha\beta}\nabla_\alpha\hat{\phi}^\dagger\nabla_\beta\hat{\phi} + \frac{1}{2}m^2\hat{\phi}\hat{\phi}^\dagger + U(\hat{\phi}\hat{\phi}^\dagger), \quad (2.18)$$

and the contribution arising from the variation of the non-minimal coupling is packed in the pseudo-stress-energy tensor $T^{NMC}_{\mu\nu}$:

$$\begin{aligned} T^{NMC}_{\mu\nu} = 8\pi G L^2\Big[& XG_{\mu\nu} - \frac{1}{2}\Box\hat{\phi}^\dagger\Box\hat{\phi}g_{\mu\nu} + \frac{1}{2}\nabla_\alpha\nabla_\beta\hat{\phi}^\dagger\nabla^\alpha\nabla^\beta\hat{\phi}g_{\mu\nu} \\ & + \frac{1}{2}R\nabla_\mu\hat{\phi}^\dagger\nabla_\nu\hat{\phi} + 3\nabla^\beta\nabla_\mu\hat{\phi}^\dagger\nabla_\beta\nabla_\nu\hat{\phi} + \nabla_\mu\nabla_\nu\hat{\phi}^\dagger\Box\hat{\phi} - \nabla^\lambda\hat{\phi}^\dagger\nabla_\mu\hat{\phi}R_{\lambda\nu} \\ & - \nabla^\lambda\hat{\phi}^\dagger\nabla_\nu\hat{\phi}R_{\lambda\mu} - R_{\mu\beta\nu\lambda}\nabla^\beta\hat{\phi}^\dagger\nabla^\lambda\hat{\phi} + R_{\lambda\beta}\nabla^\lambda\hat{\phi}^\dagger\nabla^\beta\hat{\phi}g_{\mu\nu} + c.c.\Big], \end{aligned} \quad (2.19)$$

where c.c. stands for the complex conjugate. It is useful to manipulate this equation by imposing both a fluid limit and a Newtonian limit.

As we saw before, for the fluid limit we need to express $T^{\phi}_{\mu\nu}$ and $T^{NMC}_{\mu\nu}$ in terms of the fluid variables ρ and u^μ and we assume that ρ varies slowly enough in order to ignore the quantum pressure term in (2.12).

For the Newtonian limit we expand the metric as the flat spacetime metric plus a first-order scalar mode perturbation Φ and we assume small time gradients and small velocities compared to the speed of light. The full calculation can be seen in [42]. At the end the field equation (2.17) reduces to a modified Poisson equation of the form

$$\nabla^2\Phi = 4\pi G(\rho - L^2\nabla^2\rho). \quad (2.20)$$

2.3 Phenomenology of the model

In the numerical simulations which lead to the cuspy profile (2.1) dark matter is modelled as dust, i.e. a pressureless collection of particles interacting only via standard (Newtonian) gravity. However, in our model dark matter is a scalar field which forms a Bose-Einstein condensate and therefore gravitates in the non-relativistic limit via the modified Poisson equation (2.20). This would lead to a different prediction for the equilibrium density profile. In order to find it, we need to solve a set of three coupled partial differential equations in terms of three variables Φ, ρ and $\vec{v}$. The first is, of course, the modified Poisson equation (2.20), the other two are the fluid equations (2.10) and (2.12) in the non-relativistic limit:

$$\nabla^2\Phi = 4\pi G(\rho - L^2\nabla^2\rho), \quad (2.21)$$

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}.(\rho\vec{v}) = 0, \quad (2.22)$$

[2]We must stress that the relation between the non-condensate phase where the field is minimally coupled and the condensate phase where a non-minimal coupling exists is not equivalent to the relation between a UV theory and an IR theory in the EFT sense. In particular, both regimes exist below the Planck scale and are not related to each other by a RG flow, rather by a thermodynamic phase transition.

$$m\frac{\partial\vec{v}}{\partial t} = -m(\vec{v}.\vec{\nabla})\vec{v} - m\vec{\nabla}\Phi + \vec{\nabla}\Big[U^{'}(\rho)\Big(1 - \xi^2\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}}\Big)\Big]. \tag{2.23}$$

Up to the extra term in the Poisson equation, these are the same equations as for the non-relaitivistic BEC discussed in the Introduction. Unfortunately, solving these three equations analytically is not possible. A numerical simulation would be required in order to evolve ρ and $\vec{v}$ from some specified initial conditions. However, we can still argue from the structure of these equations that they would tend to smooth out any cusps. In particular, the quantum pressure term $\xi^2\vec{\nabla}\Big(U^{'}(\rho)\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}}\Big)$ in (2.23) would prevent the building up of large gradients of the density.

We can actually find an approximate relation between ξ and m such that the cusp within galaxy halos is smoothed out. For that we consider the regime after the halo has relaxed to its equilibrium density profile. In that regime the gradients of the density and the velocities are negligible. This means that in (2.21) and (2.23) we can neglect both the term from the non-minimal coupling proportional to L^2 and the quantum pressure term proportional to ξ^2. Looking for static solutions $\vec{v} = 0$ leads to the same density profile as (1.73):

$$\rho(r) = \rho_0\frac{\sin(kr)}{kr}, \tag{2.24}$$

where $k = \sqrt{\frac{Gm^3}{\hbar^2 a}}$ and a is the scattering length. The radius of the halo is given by

$$R = \pi\sqrt{\frac{\hbar^2 a}{Gm^3}}. \tag{2.25}$$

On the other hand, for a non-relativistic BEC the healing length is related to the scattering length by [43]

$$\xi = \frac{1}{\sqrt{4\pi a n}}, \tag{2.26}$$

where $n \approx \bar{\rho}/m$ is the number density of particles in the condensate phase. Combining (2.25) and (2.26) allows us to express the healing length as a function of the mass:

$$\xi = \frac{\hbar}{m}\frac{\pi}{R}\frac{1}{\sqrt{4\pi G\bar{\rho}}}. \tag{2.27}$$

This relation is essentially obtained by requiring that the density distribution of the BEC in the non-relativistic minimally coupled limit is of the same size as the size of the typical galaxy halo. Since the density distribution of the BEC does not have the cusp at the centre, one can think of this as the healing length necessary to solve the core-cusp problem (of course, as mentioned before in order to really demonstrate that the core-cusp problem is solved, one would have to run numerical simulations).

In order to find the coupling constant L which favours a smooth core, we need to find a relation between L and ξ. Since the non-minimal coupling is present when ξ is of the same order as l_c the length scale of the curvature, it is the ratio ξ/l_c which controls the

strength of the non-minimal coupling. Since, l_c already appears in the Einstein tensor, L has to be of the same order as ξ. Indeed, taking $L = \xi$ we have that

$$\xi^2 G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi \sim \left(\frac{\xi}{l_c}\right)^2 g_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi, \tag{2.28}$$

and therefore the term becomes important only when $\xi \sim l_c$. Therefore, from now on we will fix $L = \xi$. We need to stress that this identification between the two length scales does not follow from the mathematics of the model or from any rigorous mathematical theory of condensation in a curved spacetime. Instead, it is based on intuition and the desire to consider the simplest possible model. So even though we assume that for now, one must entail the possibility that L and ξ could be related in some more complicated way. Writing the coupling constant as $\varsigma \equiv 1/L$ and using (2.27) gives a relation between ς and m:

$$\varsigma = m\frac{R}{\pi}\sqrt{4\pi G\rho}. \tag{2.29}$$

Equation (2.29) gives the coupling constant ς which favours a smooth core as a function of the mass m.

Field oscillations and CDM limit The complex scalar field ϕ should behave as CDM, i.e. pressureless dust, on large scales in order to fit with cosmological observations. As we noted before, this implies that in this case the 4-velocity u^μ is normalised and therefore there exists a frame where it takes the form $u^\mu = (1, 0, 0, 0)$. Because of the isotropy on large scales, the time direction in u^μ necessarily coincides with the cosmic time direction as defined in FLRW. From (2.9) this implies that $\dot{\theta} = m$ and therefore the scalar field can be written as

$$\phi = A(r)\mathrm{e}^{i(mt+\alpha_0)}, \tag{2.30}$$

where $A \equiv |\phi|$ is the amplitude of oscillations - we write it as $A(r)$ to emphasise that it depends on r only. The energy of oscillations is

$$\rho = \frac{1}{2}|\dot{\phi}|^2 + \frac{1}{2}m^2|\phi|^2 = m^2A^2 = m^2\phi^*\phi, \tag{2.31}$$

consistent with the hydrodynamic definition of the density.

For later convenience, it would also be useful to split the complex field in terms of its real and imaginary components:

$$\phi = \phi_1 + i\phi_2, \tag{2.32}$$

where each component oscillates according to:

$$\phi_1 = A(r)\cos(mt + \alpha_0), \tag{2.33}$$

$$\phi_1 = A(r)\sin(mt + \alpha_0). \tag{2.34}$$

These oscillations guarantee that the pressure averages out to zero and therefore that the field behaves as CDM on large scales. We will treat the CDM-like behaviour as a zero-level approximation for the evolution of the field inside galaxies and clusters.

Summary The ΛCDM model, while very successful in explaining dark matter observations on large scales, experiences problems in fitting observations on small (galactic) scales. Most notorious of these small-scale problems is the core-cusp problem - numerical N-body simulations predict a very different density profile in galaxy halos from the one observed. As a step towards a solution of the core-cusp problem and the reconciliation of dark matter observations on small and large scales, in this chapter we considered the model of dark matter as a BEC. We argued that the resultant profile fits better observations in galaxies due to the effect of the quantum pressure term which smooths out large gradients in the density and prevents the formation of a cusp. We derived a relation between the healing length and the mass of the field that needs to be satisfied so that the cusp can be smoothed out in agreement with observations. The formation of a non-zero coherence length (the healing length) allows the condensate to couple directly to the curvature and thus leads to the appearance of an extra non-minimal coupling in the action which modifies the gravitational dynamics on small scales. We have also argued that in the CDM limit, in which the non-minimal coupling is suppressed, the field oscillates harmonically - something which would be very important later on when we shall parametrise deviations from this oscillatory behaviour by taking into account the possible damping due to the non-minimal coupling. Having reviewed the model of dark matter as a non-minimally coupled BEC, let us now turn to the question of how gravitational wave propagation is modified within this model.

3

Testing the BEC model with gravitational waves

In the previous chapter we developed the model of dark matter as a Bose-Einstein condensate with a non-minimal coupling to the metric. In this chapter we want to test this model by using the signals from the event GW170817. If dark matter is really non-minimally coupled in high-density regions, then as the gravitational wave and the gamma-ray burst propagate through these regions, the speed of gravitational waves would be modified, which would result in a non-zero time lag between the two signals. By comparing the predicted time delay from our model to the observed time delay, we can put constraints on the model [44]. We begin by deriving a formula for the speed of the gravitational waves in a theory with the action (2.13) both for the cases of a timelike and spacelike gradient of the scalar field. Then, after parametrising the model to take into account several uncertainties, we derive a formula for the total time lag predicted by the model. Comparing this to the observed time delay, we put constraints on the parameters m and ς. Finally, we comment on the validity of our results.

3.1 GW propagation in modified gravity

It is well known that in several modified theories of gravity, in particular Horndeski Scalar-Tensor theories, gravitational waves propagate with a speed different from that of light. It is therefore of no surprise that in our model, where a non-minimal coupling forms after a phase transition, this phenomenon is also present. We now outline the mathematical procedure to calculate the speed of gravitational waves. We choose to work at the level of the field equations, although another possibility would have been to work at the level of the action (see, for example, [45] and [46] for such an approach).

Starting from the gravitational part of the action (2.13) expanded in terms of the condensate wave function ϕ:

$$S_{EH} = \int \Big[\frac{1}{16\pi G}R + L^2 G_{\mu\nu}\nabla^\mu\phi^*\nabla^\nu\phi\Big]\sqrt{-g}d^4x \tag{3.1}$$

it takes a tedious but straightforward calculation to obtain the gravitational field equa-

tion

$$\frac{1}{16\pi G}G_{\mu\nu} + L^2\Big[-\frac{1}{2}XG_{\mu\nu} + \frac{1}{4}\Box\phi^*\Box\phi g_{\mu\nu} - \frac{1}{4}\nabla_\alpha\nabla_\beta\phi^*\nabla^\alpha\nabla^\beta\phi g_{\mu\nu} \\ - \frac{1}{4}R\nabla_\mu\phi^*\nabla_\nu\phi + \frac{3}{2}\nabla^\beta\nabla_\mu\phi^*\nabla_\beta\nabla_\nu\phi - \frac{1}{2}\nabla_\mu\nabla_\nu\phi^*\Box\phi + \frac{1}{2}\nabla^\lambda\phi^*\nabla_\mu\phi R_{\lambda\nu} \\ + \frac{1}{2}\nabla^\lambda\phi^*\nabla_\nu\phi R_{\lambda\mu} + \frac{1}{2}R_{\mu\beta\nu\lambda}\nabla^\beta\phi^*\nabla^\lambda\phi - \frac{1}{2}R_{\lambda\beta}\nabla^\lambda\phi^*\nabla^\beta\phi g_{\mu\nu} + c.c.\Big] = 0. \tag{3.2}$$

where the term in the brackets is proportional to $T^{NMC}_{\mu\nu}$ evaluated at the condensate wavefunction ϕ (i.e. ignoring quantum excitations). Eqn. (3.2) reduces to the Einstein Field equation in vacuum when the non-minimal coupling constant is sent to zero: $L \to 0$. We expand both the metric and the scalar field (condensate wave function) in terms of a background plus a first order perturbation, use local flatness to write the metric background as Minkowski

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{3.3}$$

$$\phi = \bar{\phi} + \varphi, \tag{3.4}$$

impose the de Donder gauge

$$\partial_\mu h^{\mu\nu} = \frac{1}{2}\partial_\nu h, \tag{3.5}$$

and require that the trace vanishes

$$h \equiv \eta^{\mu\nu}h_{\mu\nu} = 0. \tag{3.6}$$

The latter is a restriction of the solution, not a gauge fixing. Since we are interested only in the propagation of tensor modes, and since tensor and scalar modes decouple from each other in linear theory, we can without loss of generality impose the above restriction, thereby removing any scalar modes involving the trace from the equation [47]. The final result is the linearised field equation[1]

$$\frac{1}{16\pi G}\Box h_{\mu\nu} - L^2\Big[\frac{1}{4}\Box h_{\mu\nu}\partial_\rho\bar{\phi}^*\partial^\rho\bar{\phi} - \frac{1}{2}\Box h_{\rho\mu}\partial_\nu\bar{\phi}^*\partial^\rho\bar{\phi} \\ - \frac{1}{2}\Box h_{\rho\nu}\partial_\mu\bar{\phi}^*\partial^\rho\bar{\phi} + \frac{1}{2}\Box h_{\rho\sigma}\partial^\rho\bar{\phi}^*\partial^\sigma\bar{\phi}\eta_{\mu\nu} + \frac{1}{2}(-\partial_\mu\partial_\nu h_{\rho\sigma} \\ + \partial_\mu\partial_\rho h_{\nu\sigma} + \partial_\nu\partial_\rho h_{\mu\sigma} - \partial_\rho\partial_\sigma h_{\mu\nu})\partial^\rho\bar{\phi}^*\partial^\sigma\bar{\phi} + c.c.\Big] = 0. \tag{3.7}$$

The real fields ϕ_1 and ϕ_2 appear on the same footing in the Lagrangian and thus satisfy similar equations of motion. Therefore, the gradients of both ϕ_1 and ϕ_2 must be of the same nature - spacelike or timelike. Thus from now on we will only speak about the gradient of $\bar{\phi}$. There are two cases that we need to consider - first, when the gradient of $\bar{\phi}$ is timelike, and second when it is spacelike.

[1]Some steps of that and the previous computation were performed using the xAct package of Mathematica.

3.1.1 The Case of a Timelike Gradient of $\bar{\phi}$

If the gradient of $\bar{\phi}$ is timelike, we can boost to a frame where it has the form[2]

$$\bar{\phi}_\mu = (\dot{\bar{\phi}}, 0, 0, 0). \tag{3.8}$$

We perform a $3+1$ decomposition of the perturbation $h_{\mu\nu}$ (again we are interested only in the propagation of tensor modes):

$$h_{00} = 0, \qquad h_{0i} = h_{i0} = 0, \qquad \partial^i h_{ij} = \delta^{ij} h_{ij} = 0. \tag{3.9}$$

This allows to rewrite (3.7) in the form of a wave equation

$$-\ddot{h}_{ij} + c_g^2 \nabla^2 h_{ij} = 0, \tag{3.10}$$

from where it is straightforward to extract the speed of propagation of gravitational waves - c_g:

$$c_g^2 = \frac{1 + 8\pi G L^2 |\dot{\bar{\phi}}|^2}{1 - 8\pi G L^2 |\dot{\bar{\phi}}|^2}. \tag{3.11}$$

3.1.2 The Case of a Spacelike Gradient of $\bar{\phi}$

If the gradient of $\bar{\phi}$ is spacelike, then we can boost to a frame where it has the form

$$\bar{\phi}_\mu = (0, \vec{\nabla}\bar{\phi}). \tag{3.12}$$

We separate the space gradient into components parallel and perpendicular to the propagation vector $\vec{k}$:

$$|\vec{\nabla}\bar{\phi}|^2 = |\bar{\phi}_\parallel|^2 + |\bar{\phi}_\perp|^2. \tag{3.13}$$

Performing the $3+1$ decomposition as before and considering a single mode of wave vector $\vec{k}$ and frequency ω,

$$h_{ij} = A_{ij} e^{i(\vec{k}.\vec{x} - \omega t)}, \tag{3.14}$$

allows to find the dispersion relation

$$\omega^2 - c_g^2 k^2 = 0, \tag{3.15}$$

from where, again, it is straightforward to read the speed of propagation:

$$c_g^2 = \frac{1 + 8\pi G L^2 |\bar{\phi}_\parallel|^2 + 8\pi G L^2 |\bar{\phi}_\perp|^2}{1 - 8\pi G L^2 |\bar{\phi}_\parallel|^2 + 8\pi G L^2 |\bar{\phi}_\perp|^2}. \tag{3.16}$$

[2]Technically, we need to do this for both components of the complex field. However, since the evolution of the two components is similar, the same boost would bring both components to approximately the same form.

One can also rewrite the last formula in terms of the angles ι_1 and ι_2 between the wave vector $\vec{k}$ and the gradient of each component of the complex scalar field, $\vec{\nabla}\bar{\phi}_1$ and $\vec{\nabla}\bar{\phi}_2$:

$$c_g^2 = \frac{1 + 8\pi GL^2|\vec{\nabla}\bar{\phi}_1|^2 + 8\pi GL^2|\vec{\nabla}\bar{\phi}_2|^2}{1 - 8\pi GL^2|\vec{\nabla}\bar{\phi}_1|^2\cos(2\iota_1) - 8\pi GL^2|\vec{\nabla}\bar{\phi}_2|^2\cos(2\iota_2)}. \tag{3.17}$$

In our case the scalar field is only radially dependent, Eqn. (2.30), which implies that the two gradients are aligned $\iota_1 = \iota_2 = \iota$, so the expression for the speed reduces to

$$c_g^2 = \frac{1 + 8\pi GL^2|\vec{\nabla}A|^2}{1 - 8\pi GL^2|\vec{\nabla}A|^2\cos(2\iota)}. \tag{3.18}$$

A very important point to note is that Equations (3.16) and (3.18) imply that the speed depends on the direction of propagation of the wave or more accurately on the angle subtended between the wave vector and the spatial gradient of the field. In fact, when $\vec{k}$ is orthogonal to $\vec{\nabla}A$, the speed of gravity is equal to the speed of light. As we vary the angle the speed of gravity increases and it reaches its maximal possible value when $\vec{k}$ is aligned with $\vec{\nabla}A$.

Another important point is that both (3.11) and (3.16) allow for superluminal propagation. While the fundamental theory - General Relativity with a minimally-coupled scalar field, respects all the Lorenz symmetries, as the phase transition occurs and a non-minimal coupling forms, the Lorentz invariance is spontaneously broken. The reason behind that is that the BEC selects a preferred frame of reference and the gradient of the scalar field selects a preferred direction in spacetime. This Lorentz violation is not more drastic than what happens in theories of dark energy with non-minimal couplings and is also reminiscent of the Scharnhorst effect in optics [48]. There higher order QED corrections modify the speed of light at low wavelengths as it travels in the Casimir vacuum between two parallel tiny plates so that the speed in the direction orthogonal to the plates is greater than c. Because of the boundary condition set by the plates, the ground state breaks the Lorentz invariance and this leads to superluminal propagation even though the fundamental theory from which the effect is derived still respects all Lorentz symmetries. In fact, as shown in [49] the Scharnhorst effect does not lead to any causal paradoxes. This serves as an argument that also in the BEC model the superluminal propagation is not a problem and is consistent with all the assumptions that we have made.

It is an interesting question whether the modification of the propagation speed is the same for waves of all wavelengths. In fact, the answer is "no" and the easiest way to see that is through the dispersion relation $\omega^2 = c_g^2 k^2$. If we write the speed of gravity in the compact form

$$c_g^2 = c^2\Big(1 + \frac{L^2\alpha}{\delta^2}\Big) \tag{3.19}$$

where δ is the length scale over which $\bar{\phi}$ changes and α is a dimensionless parameter which depends on the Planck length and the strength of the field, then the dispersion

relation can be written as (keeping factors of c just for clarity)

$$\omega^2 = c^2k^2 + c^2\Big(\frac{4\pi^2\alpha}{\delta^2}\Big)\frac{k^2}{\kappa^2}. \tag{3.20}$$

where in the second term on the RHS we have introduced $\kappa \equiv \frac{2\pi}{L}$ in order to make the scaling between k and κ and hence λ and L more apparent. It is obvious that (keeping the gradient fixed) for $\kappa \gg k$, hence $\lambda \gg L$, this reduces to the standard relativistic dispersion relation, while a significant deviation from the standard relation occurs when $\lambda \sim L$ or $\lambda \ll L$. Thus the speed of propagation is modified only for wavelengths smaller than the healing length (which we have identified with the coupling constant). Since in our case the wavelength is much shorter than the healing length, as verified later, there will be a corresponding modification of the speed. On the other hand λ has to be larger than the scattering length a since on scales smaller than a individual particles can be discerned and the condensate description breaks. Since the scalar field constituting dark matter is very weakly self-interacting, the scattering length has to be very small - much smaller than the wavelength of the observed gravitational waves. We elaborate more on this later in the chapter.

An intriguing peculiarity of (3.11) and (3.16) is that they allow the denominator to become zero and therefore the speed to blow up. This is not a serious worry in our case because the terms proportional to the gradients of the scalar field can a posteriori be verified to be extremely small. Actually, when we later estimate the predicted difference in the arrival time of electromagnetic and gravitational waves, we will only need the linearly expanded versions of (3.11) and (3.16), which can be compactly written as

$$c_g^2 \approx 1 + 2\Delta c_g, \tag{3.21}$$

where $\Delta c_g \equiv \frac{c_g - c}{c}$. For the cases of timelike and spacelike gradients, Δc_g is given by

$$\Delta c_g \approx 8\pi G L^2 |\dot{\bar{\phi}}|^2, \tag{3.22}$$

$$\Delta c_g \approx 8\pi G L^2 |\bar{\phi}_{\parallel}|^2. \tag{3.23}$$

In the case when the scalar field is only radially dependent and its spatial gradient is aligned with the wave vector $\vec{k}$, the latter expression can be rewritten as

$$\Delta c_g \approx 8\pi G L^2 \vec{\nabla} A \cdot \vec{\nabla} A = \frac{8\pi G L^2}{m^2} \vec{\nabla}\sqrt{\rho} \cdot \vec{\nabla}\sqrt{\rho}. \tag{3.24}$$

Nevertheless, the potential divergence of (3.11) and (3.16) shows that these formulas cannot be applied at arbitrary large gradients of the scalar field. It is viable that there is a feedback mechanism built inside the model which prevents the building of such gradients. Indeed, for a BEC the formation of large gradients is prevented by the the quantum pressure term, but a similar mechanism might hold even for a genuine modification of General Relativity.

3.2 Parametrisation of the model and astrophysical setting

We currently do not have a rigorous mathematical theory of condensation in a curved spacetime. As a result, there are several uncertainties in the model of DM BEC. For example, we do not know how strong the gravitational potential wells need to be so that the condensate forms. Therefore, several scenarios are possible. The condensate might form only within galaxy halos, or it might also form outside galaxy halos. It might even be possible that the whole halo of a cluster of galaxies is in the condensed phase. We capture the uncertainty about the level at which the condensate forms by introducing a free parameter β which is defined in the next subsection.

A second source of uncertainty is whether the scalar field (i.e. the condensate wave function) oscillates or not around the minima of its potential. As we have argued before, in order for the scalar field to mimic CDM on large scales, the field has to oscillate. However, the presence of the non-minimal coupling $G_{\mu\nu}\nabla^{\mu}\phi^{*}\nabla^{\nu}\phi$ changes the evolution of the field with respect to the standard case. It is viable that this coupling would serve as a damping force which would tend to dissipate the oscillatory energy. In order to calculate the precise effect of the non-minimal coupling we would have to numerically solve for the evolution of the scalar field and also know details about the dynamics of the galaxy immediately after the condensation has happened and before is has reached equilibrium. Since this goes beyond the scope of this work, we are going to capture the uncertainty about whether the oscillations are damped or not by introducing two extra free parameters - γ_1 and γ_2. Including m - the mass scale of the field, and $\varsigma \equiv \frac{1}{L}$ the scale of the non-minimal coupling, we have in total five free parameters - $m, \varsigma, \beta, \gamma_1, \gamma_2$. We will express the total predicted arrival time difference between GWs and GRBs that accumulates along the way between the source and the observer in terms of them.

In order to achieve that analytically, we work at an order of magnitude level. We also make several other simplifying assumptions and approximations. We assume that the density distribution is given by the Burkert profile (2.2), i.e. we assume that the quantum pressure has already smoothed out any cusps and the system has relaxed to its equilibrium profile (Ref. [42] shows that one can obtain a profile similar to the Burkert one within this model). The dominant effect on the time delay of GRBs with respect to GWs comes from the halos of the Milky Way and the host galaxy and eventually from the halos of clusters between us and the event (all other halos are too far away from the physical path of the waves). We assume that the GWs and GRBs pass through the centre of each halo. At first, this might seem like a drastic assumption but it actually leads to the maximal possible difference in the arrival time of the two signals and is good enough for putting a constraint on the theory. Nonetheless, later we comment on the case where the waves pass at an impact parameter b.

In order to perform an order of magnitude estimation of the difference in the arrival time, we also assume that in the case of spatial gradients the GW propagates with the maximal possible speed (i.e. the speed is constant and is calculated at the radius where the gradient of ϕ is largest) only inside the core of the halo (this agrees with a numerical calculation of the full arrival time difference performed for different values

of the parameters ς and m), while in the case of time gradients, we later show that the modification of the speed of GWs is proportional to the average DM density of the halo. In both cases we work to first order in Δc_g as given by Equations (3.22) and (3.23) respectively (a posteriori checks can verify that these terms are very small and numerical checks confirm that for the spatial case the first order calculation gives the same result as the full calculation).

We calculate the contribution to the arrival time difference coming from a single galaxy halo assuming that the galaxy is typical i.e. that it has the following parameters [50]:

Dark matter density at the centre	$\rho_{0,g} = 3 \times 10^{-22} kg/m^3$
Average dark matter density in the halo	$\bar{\rho}_{gh} = 10^{-23} kg/m^3$
Radius of the core	$r_{0,g} = 15kpc = 4.5 \times 10^{20} m$
Radius of the halo	$R_{h,g} = 200kpc = 6 \times 10^{21} m$

Similarly, when we take into account the possible contributions from clusters, we assume that they are typical with the following parameters:

Dark matter density at the centre	$\rho_{0,cl} = 9 \times 10^{-23} kg/m^3$
Average dark matter density in the halo	$\bar{\rho}_{cl} = 3 \times 10^{-24} kg/m^3$
Virial radius	$R_{h,cl} = 1.2 \times 10^{23} m$
Radius of the core	$r_{0,cl} = 2.4 \times 10^{22} m$

In order to calculate the average dark matter density for a galaxy halo and a cluster halo, we use the following formula which is obtained from numerical simulations [51,52]:

$$\bar{\rho} \approx 200 \rho_{cr}(z_{vir}), \tag{3.25}$$

where $\rho_{cr}(z_{vir})$ is the critical density of the Universe evaluated at z_{vir} and z_{vir} is the redshift at which the galaxy or cluster virialized.

In order to obtain $r_{0,cl}$ from $R_{h,cl}$ we use the fact that they are related by [38]

$$r_{0,cl} = \frac{R_{h,cl}}{C_p}, \tag{3.26}$$

where the concentration parameter C_p is approximately a constant (it is very weakly mass-dependent) and for a typical cluster of galaxies: $C_p \approx 5$.

Note that the above given $\rho_{0,cl}$ is obtained from $\rho_{0,g}$ by assuming they are related by

$$\frac{\rho_{0,cl}}{\rho_{0,g}} = \frac{\bar{\rho}_{cl}}{\bar{\rho}_{gh}}. \tag{3.27}$$

It is not certain whether this assumption holds - it is still debatable whether halos of clusters have a well-defined core and what their central density is. In any case, it is unlikely that this ambiguity will affect the order of magnitude constraints that we will later obtain.

The distance between the source and the observer is $\ell = 40$ Mpc. If we draw a line, between us and the source about half of the line would pass through a cluster. Since the typical virial radius of a cluster halo is about 4 Mpc, then statistically there are about 5 clusters or $N = 10$ half-clusters between us and the source.

Parametrising the model The first source of uncertainty is whether the condensation occurs only within galaxy halos or also outside galaxy halos within the cluster halo. We started from the assumption that there is condensation within galaxies. What can we say about the physics at the scale of galaxy clusters?

A priori, there is a range of different possibilities. It might be that there is no dark matter in the condensate phase outside galaxies, it might be that in some parts of the clusters where the gravitational fields are stronger the dark matter field has condensed or it might be that the whole cluster is in the condensate phase. In fact, we can capture all the different cases by introducing a term $\tilde{L}G_{\mu\nu}\langle\phi\rangle^{\mu}\langle\phi\rangle^{\nu}$ in the effective Lagrangian at cluster scales. Here $\langle\phi\rangle$ is the average of the field over cluster scales and $\tilde{L}$ is a new coupling constant which we keep as a free parameter that represents our uncertainty as to what exactly happens on cluster scales.

We allow $\tilde{L}$ to have a range from 0, which corresponds to the case where there is no condensation anywhere outside galaxies, to $\tilde{\xi} = \alpha\xi$ which corresponds to the case where the whole cluster is in the condensate phase. Here α is an empirical constant which takes into account the fact that a BEC in the whole cluster would have a different healing length than a BEC in a galaxy. In order to calculate α we need to take into account that the healing length of a condensate is related to the average matter density and radius of a halo by (2.27). This implies

$$\alpha = \left(\frac{\bar{\rho}_{cl}}{\bar{\rho}_{gh}}\right)^{-1/2} \frac{R_{h,g}}{R_{h,cl}} \approx 0.1. \tag{3.28}$$

We now define the parameter β by

$$\beta \equiv \frac{\tilde{L}}{\alpha\xi} \tag{3.29}$$

so that it has a range $(0, 1)$.

The second source of uncertainty is whether the field ϕ, which corresponds to the condensate wave function, oscillates or not. This is important because any oscillatory behaviour would lead to large time gradients, which would contribute significantly to the modification of the speed of gravitational waves in a cosmological medium. The scalar field that constitutes dark matter would tend to oscillate. However, we saw that in our case the non-minimal coupling to the curvature could potentially serve as a damping force which would tend to relax this oscillation. The question is whether this damping really happens and how effective it is.

We can parametrise this uncertainty by introducing another free parameter γ_1. In the extreme case where there is no damping there would be simple harmonic oscillations with period proportional to the inverse of the mass of the scalar field. We model the effect of the damping by still assuming simple harmonic oscillations but with a longer period T_{eff} and therefore a smaller effective mass m_{eff}. While this is not strictly true, it is good enough for the purposes of the present investigation. We define the dimensionless parameter γ_1 by

$$\gamma_1 \equiv \frac{m_{\text{eff}}}{m}. \tag{3.30}$$

We see that it has limiting values $\gamma_1 = 0$ corresponding to completely damped oscillations and $\gamma_1 = 1$ corresponding to completely undamped (free) oscillations.

Next we need to take into account that if the condensation happens within clusters as well, then the oscillations might be damped to a different extent outside galaxy halos. In order to capture this uncertainty, we introduce another free parameter γ_2 which is defined in the same way as γ_1 but for oscillations within clusters, leaving γ_1 to represent only our uncertainty about oscillations within galaxies.

3.3 Constraining the model

Deriving the formula for the total arrival time difference The modification of the speed of propagation of gravitational waves with respect to the speed of light is proportional to the gradients of the scalar field (condensate wave-function). Larger gradients imply larger modification which in turn implies larger time lag between the two waves. Therefore the total time lag, the difference in arrival time which is detected at the observer, accumulates only when the waves pass through regions with large gradients of the field. There are four separate contributions to the arrival time difference that we need to consider - from spatial gradients within galaxies, from time gradients within galaxies, from spatial gradients within clusters, from time gradients within clusters. We will now derive formulas for each in turn.

The contribution from spatial gradients in galaxies We first look at the contribution from the spatial gradients in galaxies. We make use of all the assumptions and approximations stated earlier. The DM density distribution inside a galaxy halo is given by the Burkert density profile

$$\rho(r) = \frac{\rho_{0,g}}{(1 + \frac{r}{r_{0,g}})(1 + (\frac{r}{r_{0,g}})^2)}. \tag{3.31}$$

Substituting that inside (3.24) one obtains:

$$\Delta c_g(r) = \frac{\rho_{0,g}}{4M_p^2\varsigma^2 m^2 r_{0,g}^2} F\Big(\frac{r}{r_{0,g}}\Big). \tag{3.32}$$

where M_p is the reduced Planck mass defined by

$$M_p \equiv 1/\sqrt{8\pi G} = 2.4 \times 10^{27} eV \tag{3.33}$$

and F is a function defined by

$$F(x) := \frac{(1 + 2x + 3x^2)^2}{(1 + x)^3(1 + x^2)^3}. \tag{3.34}$$

The total time lag between electromagnetic and gravitational waves can be written as an integral over the line of sight

$$\begin{aligned}\Delta t &= 2\int_0^{\ell/2} dr \Delta c_g(r) \\ &= \frac{\rho_{0,g}}{2M_p^2\varsigma^2 m^2 r_{0,g}} \int_0^{\ell/2r_{0,g}} F(x)dx. \end{aligned} \tag{3.35}$$

Since we assume that the wave propagates with the maximally modified speed over a distance equal to the diameter of the core, this reduces to

$$\Delta t_g = \frac{\rho_{0,g}}{4M_p^2\varsigma^2 m^2 r_{0,g}} F\Big(\frac{r_{\max,g}}{r_{0,g}}\Big). \tag{3.36}$$

where $r_{\max,g}$ is the radius at which the gradient is maximal: $r_{\max,g} = 1.5 \times 10^{20} m$.

The contribution from time gradients in galaxies Next we calculate the contribution from time gradients in galaxy halos. Since galaxies are virialised objects, they don't evolve very much. There are very small time gradients from rotation and fluxes of radiation and heat. However, they are mostly associated with the visible matter inside halos and their influence on the evolution of the dark matter field is negligible. The main contribution to the time gradient of the condensate field comes from possible oscillations of the field. Now we calculate this contribution in terms of the free parameter γ_1. According to our assumptions, the field oscillates as a free field with effective mass $m_{\rm eff}$:

$$\bar{\phi} = A(r)\mathrm{e}^{i(m_{\rm eff}t+\alpha_0)}. \tag{3.37}$$

Therefore,

$$|\dot{\bar{\phi}}|^2 = m_{\rm eff}^2 A(r)^2 = m_{\rm eff}^2 \bar{\phi}^2 = \frac{m_{\rm eff}^2 \bar{\rho}_{gh}}{m^2} = \gamma_1^2 \bar{\rho}_{gh}, \tag{3.38}$$

where $\bar{\phi}^2 \propto \bar{\rho}_{\rm gh}$ is necessary since the field oscillates everywhere within the halo. The fractional modification of the speed of gravitational waves (written in terms of Newton's constant G) is

$$\Delta c_g = \frac{8\pi G}{\varsigma^2}|\dot{\bar{\phi}}|^2 = \frac{8\pi G \bar{\rho}_{\rm gh}\gamma_1^2}{\varsigma^2}, \tag{3.39}$$

where we make use of the assumed first order relation between the density and the scalar field. This immediately gives us the formula for the arrival time difference:

$$\Delta t_{g,osc} = 2R_{h,g}\Delta c_g = \frac{16\pi G R_{h,g}\bar{\rho}_{gh}\gamma_1^2}{\varsigma^2}. \tag{3.40}$$

The contribution from spatial gradients in clusters The dark matter density distribution inside a cluster halo follows the Burkert profile but with different values of the central density $\rho_{0,cl}$ and of the radius of the core $r_{0,cl}$:

$$\rho(r) = \frac{\rho_{0,cl}}{(1+\frac{r}{r_{0,cl}})(1+(\frac{r}{r_{0,cl}})^2)}. \tag{3.41}$$

As a consequence, the maximal value of the gradient of the scalar field is reached at $r_{max,cl} = 7.7 \times 10^{21} m$. The difference between the arrival times in this case is given by the same formula as (3.36) but with the parameters of the cluster halo and with one major difference - the prefactor that we insert is, in accordance with our assumptions, $N \times r_{0,cl}$ where $N = 10$ is the number of half-clusters between us and the source:

$$\Delta t_{cl} = \frac{N\alpha^2\beta^2\rho_{0,cl}}{4M_p^2\varsigma^2 m^2 r_{0,cl}} F\Big(\frac{r_{max,cl}}{r_{0,cl}}\Big). \tag{3.42}$$

The contribution from time gradients in clusters The case of the time delay between the two waves arising from time gradients within clusters of galaxies mirrors almost exactly the case of time gradients within galaxies except that it is now proportional to the average dark matter density within clusters - $\bar{\rho}_{cl}$ and that the distance over which the GW propagates with modified speed is $\ell/2$:

$$\Delta t_{cl,osc} = \frac{4\pi G\ell\bar{\rho}_{cl}\alpha^2\beta^2\gamma_2^2}{\varsigma^2}. \tag{3.43}$$

The total arrival time difference Gathering all contributions together, we obtain a formula for the total predicted arrival time difference in terms of the five free parameters:

$$\begin{aligned}\Delta t_{tot} &= \Delta t_g + \Delta t_{g,osc} + \Delta t_{cl} + \Delta t_{cl,osc} \\ &= \frac{\rho_{0,g}}{4M_p^2\varsigma^2 m^2 r_{0,g}} F\Big(\frac{r_{\max,g}}{r_{0,g}}\Big) + \frac{16\pi G R_{h,g}\bar{\rho}_{gh}\gamma_1^2}{\varsigma^2} \\ &+ \frac{N\alpha^2\beta^2\rho_{0,cl}}{4M_p^2\varsigma^2 m^2 r_{0,cl}} F\Big(\frac{r_{max,cl}}{r_{0,cl}}\Big) + \frac{4\pi G\ell\bar{\rho}_{cl}\alpha^2\beta^2\gamma_2^2}{\varsigma^2}.\end{aligned} \tag{3.44}$$

It is easy to see how the total arrival time difference scales with the three parameters which parametrise our uncertainty - β, γ_1 and γ_2. It is proportional to the square of each parameter. Since each parameter has a range $(0, 1)$, the terms containing these parameters become important when the corresponding parameter gets close to 1. This also means that we cannot a priori ignore any of these terms, since each of them becomes important in some regime. However, we can identify several physically important limiting cases and testing the model in each of these cases gives us an idea of the overall constraint.

Testing the Limiting Cases There are two limiting values - 0 and 1, for each of the three parameters - β, γ_1 and γ_2, which naively would give a total of eight limiting cases. However, two of them, $\{\beta = 0, \gamma_1 = 0, \gamma_2 = 0\}$ and $\{\beta = 0, \gamma_1 = 1, \gamma_2 = 0\}$ are redundant, since they correspond to no condensation outside galaxy halos, which in turn would make it irrelevant if the field there oscillates or not (in fact, it probably will, since there is no non-minimal coupling there to suppress the oscillations).

Of the remaining six limiting cases, there are two classes of cases which lead to different results. It turns out that the constraint is not sensitive to the value of β, i.e. in the end it is not important whether the condensation happens only inside galaxy halos

or also inside cluster halos. Instead the constraint is sensitive to the values of γ_1 and γ_2, i.e. it matters whether the oscillations are suppressed everywhere or they happen freely somewhere.

No field oscillations The first class of cases is when there are no oscillations anywhere in the condensate phase which corresponds to the cases $\{\beta = 0, \gamma_1 = 0\}$ and $\{\beta = 1, \gamma_1 = 0, \gamma_2 = 0\}$. Then the gradients are far too small to seriously affect the time difference between gravitational and electromagnetic waves. Calculating the constraint in $\varsigma - m$ space at an order of magnitude level, we obtain

$$\varsigma \times m \gtrsim 10^{-52} eV^2. \tag{3.45}$$

Fig.3.1 shows a constraint plot in the $\varsigma - m$ parameter space. The red line corresponds to the relation (2.29) i.e. we identify the coupling constant L with the healing length ξ and demand that ξ is of the order necessary to fit the size of non-relativistic BEC halo with the size of a typical DM halo. Thus it provides an independent constraint in the parameter space originating from the requirement that the density profile of the BEC fits with the observed density profile within galaxies. Combining this with the time of flight constraint, we get separate constraints on ς and m

$$m \gtrsim 10^{-24} eV, \qquad \varsigma \gtrsim 10^{-28} eV. \tag{3.46}$$

Undamped field oscillations The second class of cases is when the oscillations happen somewhere - it could be either in galaxy halos, in cluster halos or in both. This corresponds to the cases $\{\beta = 0, \gamma_1 = 1\}$, $\{\beta = 1, \gamma_1 = 1, \gamma_2 = 0\}$, $\{\beta = 1, \gamma_1 = 0, \gamma_2 = 1\}$ and $\{\beta = 1, \gamma_1 = 1, \gamma_2 = 1\}$. Calculating the constraint in $\varsigma - m$ space at an order of magnitude level, we obtain a strong constraint on ς

$$\varsigma \gtrsim 10^{-25} eV, \tag{3.47}$$

and a weak constraint on m coming from the spatial gradients of ϕ as calculated in the previous case. Fig. 3.2 shows a constraint plot in the $\varsigma - m$ parameter space. Again combining the time of flight constraint with the requirement that the BEC halo fits in size with a typical galaxy halo (the red line) leads to a separate constraint on m

$$m \gtrsim 10^{-21} eV. \tag{3.48}$$

These constraints allow us to verify that $\lambda \ll L$ and therefore the GW is sensitive to the non-minimal coupling. Indeed, for GW170817, $\lambda \sim 10^5 - 10^7 m$, while (3.47) implies that the region $L \gtrsim 10^{18} m$ is allowed. For small L such that $L \sim \lambda$ the GW would no longer be sensitive to the non-minimal coupling and there would be no corresponding time delay between the two waves. However, since this region is allowed anyway, it does not change the constraint.

At first it might seem surprising that the contribution to the arrival time difference from cluster halos is of the same order as the contribution from galaxy halos especially in

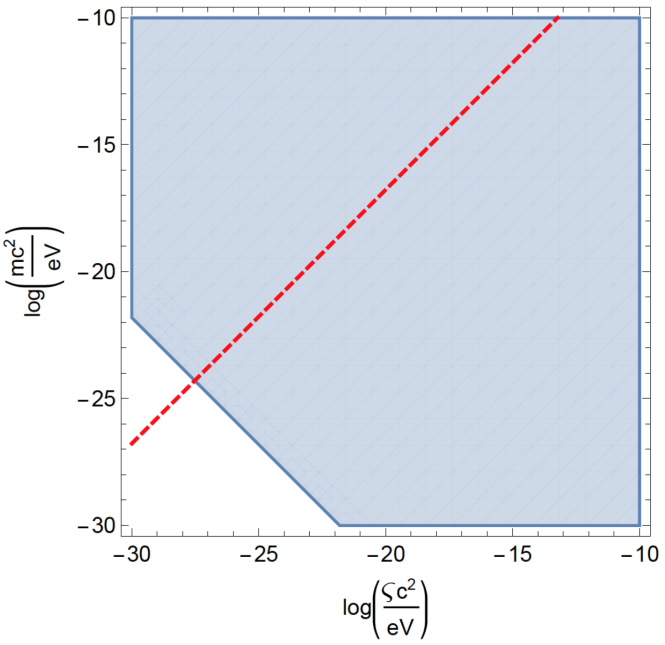

Figure 3.1: A constraint on the $\varsigma - m$ parameter plane when the oscillations are completely suppressed everywhere. The arrival time difference between gravitational and electromagnetic waves is solely due to spatial gradients. Models which lie in the blue region are consistent with the observed time difference $\Delta t \lesssim 1.7s$. The red line corresponds to values of ς and m for which the BEC halo fits the size of a typical galaxy halo.

the case where the difference arises from the oscillations of the field. Naively, one would expect the contribution from cluster halos to be much larger due to the larger distance within a condensed region that the waves have to propagate (the factor of ℓ in the fourth term of (3.44) compared to the factor of $R_{h,g}$ in the second term). However, it turns out that this surplus in the propagation distance is exactly compensated by the fact that the healing length and hence the non-minimal coupling is of different size outside galaxy halos compared to inside due to the different dark matter density (the factor of α^2 in the fourth term). As a consequence, it is completely irrelevant for the time-of-flight constraint whether the condensation happens only inside galaxy halos or also outside. What matters only is whether the oscillations are suppressed or not. As we saw this uncertainty is captured by the parameter γ_1 (and γ_2 for oscillations outside galaxy halos). Formula (3.44) shows that the arrival time difference depends quadratically on γ_1 as it varies from 0 to 1. Thus as the oscillations are gradually switched on, larger and larger regions of the parameter space will get forbidden by the oscillations. Figures 3.1 and 3.2 show the limiting cases of $\gamma_1 = 0$ and $\gamma_1 = 1$.

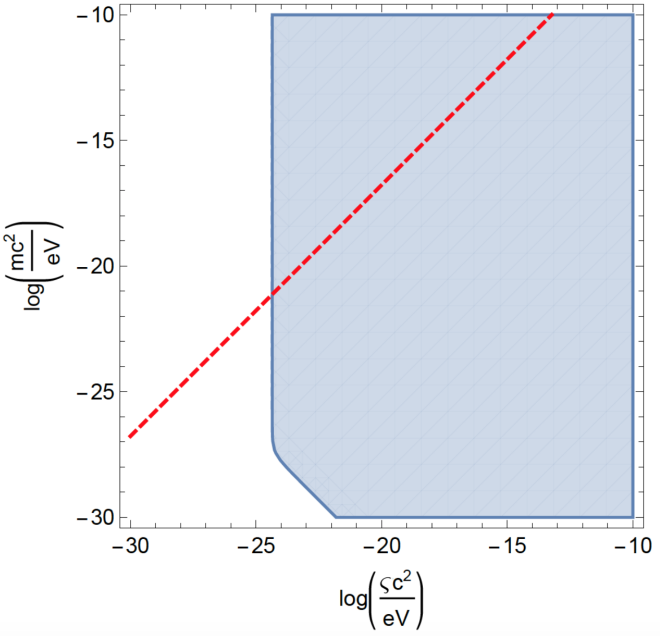

Figure 3.2: A constraint on the $\varsigma - m$ parameter plane when the field oscillates somewhere in galaxy or cluster halos. The dominant contribution to the arrival time difference is from the time gradients. Models which lie in the blue region are consistent with the observed time difference $\Delta t \lesssim 1.7s$. The red line corresponds to values of ς and m for which the BEC halo fits the size of a typical galaxy halo.

Waves at an impact parameter b So far we have only considered the case where the waves pass through the centre of the halo of a galaxy or a cluster. This is good enough for putting a constraint on the model because it maximises the total time difference accumulated along the path of the two waves. But how does the total arrival time difference change as we shift the path of the waves away from the centre of the halo while still preserving the spherical symmetry? This would only make a difference, if the arrival time difference is dominated by spatial gradients i.e. oscillations are largely suppressed. Then we can derive a formula for the arrival time difference in terms of the impact parameter b, which is defined as the closest distance between the path of the waves and the centre of the halo. Working again to first order in $\frac{1}{M_p^2\varsigma^2}|\frac{d\bar{\phi}}{dr}|^2$, we obtain the formula which is just a slight modification of (3.35) above

$$\Delta t(b) = \frac{2}{M_p^2\varsigma^2}\frac{l^2}{b^2+l^2}\int_0^{l/2}\left|\frac{d\bar{\phi}}{dr}\right|^2_{r=(b^2+x^2)^{1/2}} dx. \tag{3.49}$$

The difference between the arrival times of the two signals decreases quickly as b is increased. For example, the nearest galaxy to the path of the waves, other than the Milky Way or the host galaxy, is at $b \sim 10^{22}m$ [53]. Taking values of ς and m for which the model barely passes the test in the case of completely suppressed oscillations:

$m \sim 10^{-24} eV$ and $\varsigma \sim 10^{-28} eV$ (i.e. the point where the red dashed line intersects the boundary between the allowed and forbidden regions in Fig.3.1) gives an additional time lag from the nearest galaxy of $\Delta t \sim 10^{-5} s$. This justifies our decision to ignore possible contributions to the arrival time difference from other galaxies.

Other constraints on the mass The relation between the mass and the coupling constant (2.29) which corresponds to the red dashed line in the plots suggests that we can increase both m and ς indefinitely while still fitting the profile of a galaxy with that of a BEC halo. However, both m and ς depend on the scattering length a via (2.25) and (2.26) and the scattering length is not a completely free parameter in our model. Indeed, we can obtain an independent constraint on the mass m by considering the possible range of a.

The scattering length is a constant which characterises the strength of the interactions between the bosons in the condensate. Since the interactions have to be repulsive, a has to be positive.[3] The case $a = 0$ would correspond to no interactions. This is very unlikely since in order to form a condensate, a weak interaction between the bosons is necessary (though see [55] for the possibility of forming a BEC without interactions). The case $a \lesssim l_p$ where l_p is the Planck length would correspond to an interaction weaker than the gravitational one. This also seems unlikely since in that case gravitational interactions between the particles would dominate over their repulsive interactions. On the other hand, a cannot become very large because this would imply strongly interacting dark matter which would be at odds with the observations. In particular since the lower limit for the range of observed wavelengths of the gravitational waves is $\lambda_{min} \sim 10^5\ m$ and since a scattering length of that same order would imply very strong interactions of dark matter which is excluded by the observations, we can safely conclude that $a \ll \lambda$ and thus gravitational waves propagate safely in the regime where the condensate description holds. In fact, there is a hierarchy of length scales which goes like this:

$$l_p < a < \lambda < \xi. \tag{3.50}$$

A possible, though highly overestimated, upper bound for a is given by the scattering length of Rubidium atoms [20]: $a_{Rb} = 5.77 \times 10^{-9} m$. If we impose

$$l_p < a < a_{Rb}, \tag{3.51}$$

we obtain from Equations (2.25) and (2.29) a constraint on the mass m given by

$$5.07 \times 10^{-10}\,\mathrm{eV} < m < 0.36\,\mathrm{eV}. \tag{3.52}$$

If we compare this interval to the time-of-flight constraints, we see that the interval lies deep into the allowed region of the parameter space both in the case of no oscillations and of free oscillations.

[3] Incidentally, this implies that the boson which forms the condensate in this model cannot be the QCD axion for which the interactions are attractive [54]. However, it could be a generalised axion particle.

Summary This chapter was devoted to putting a constraint on the model where dark matter forms a non-minimally coupled BEC by using the time delay between gravitational and electromagnetic waves emitted during a binary neutron star merger. We calculated the modification of the speed of propagation of gravitational waves with respect to the speed of light both for timelike (Eqn. (3.11)) and spacelike (Eqn. (3.16)) gradients of the condensate wavefunction and we argued that such a modification would be present whenever the wavelength of the gravitational waves falls in the range $a < \lambda < \xi$ – a condition which is satisfied in our case. Then we proceeded to estimate the actual time delay predicted by the model. Due to the uncertainty in the behaviour of the condensate wavefunction and the scale at which the condensation happens we introduced 3 additional parameters and expressed the expected time delay in terms of them and the mass of the field m and the strength of the non-minimal coupling ς. The total time delay was a sum of 4 contributions – the spatial and time gradients of the field in galaxies and galaxy clusters respectively (Eqn. (3.44)). Comparing the predicted time delay with the observed one allowed us to put constraints on m and ς (Figures 3.1 and 3.2). We found out that the constraint is independent of whether the condensation occurs in galaxy halos only or in clusters too, but it is quite sensitive to whether the scalar field oscillates or whether the oscillations are suppressed. Finally, we compared our constraint on the mass with other independent constraints.

4

The luminosity distance as a probe of the geometry of spacetime

In this chapter we will overview the notion of a distance in Cosmology. We will see that there are different types of distances and they are all defined in terms of light emitted by the source and detected by the observer. Throughout the chapter we work in a very general context without restricting to a particular Cosmology and show how the luminosity distance can be calculated by using two independent mathematical objects - the Jacobi map and the van Vleck determinant [56,57]. We explore Cosmography, which is an approach to test the evolution of the Universe in a model-independent way. Only later do we focus on FLRW metrics to illustrate in a more specific context some of the techniques used. This chapter should be thought of as a preparation for the next one in which we will apply these techniques to more sophisticated cosmological models.

4.1 Cosmological distances - definitions and uses

Defining a notion of a distance in Cosmology is a challenge for several reasons. First, spacetime is curved and so our Euclidean notion of a distance breaks. Second, spacetime is not static, rather it is dynamical and generally there is no background geometry to serve as a ruler. Third, we want a distance which is measurable by light observations and which captures our intuition that objects at larger distances from us are more separated from us. It turns out that there are several such notions of a cosmological distance. While they reduce to the same Euclidean distance in a flat spacetime, in a curved spacetime they are truly different and in some occasions are related to one another through well-known transformations. We will now go through the definitions and subtleties of some of them, namely those that will play an important role in our analysis later on: the luminosity distance, the area distance, the angular diameter distance, the redshift and the affine parameter distance.

The luminosity and area distances Consider a spacetime $(\mathcal{M}, g_{\mu\nu})$ and a point source emitting light at the source event S (Fig. 4.1). An extended observer located at O receives the light emitted by S. The luminosity distance is defined in terms of the energy flux measured by the observer, F_o, and the (absolute) luminosity at the source, L_s, assumed to be known by the observer [32, 58]:

$$F_o = \frac{L_s}{4\pi\, d_L^2}. \tag{4.1}$$

Measuring the photon energy flux F_o is straightforward, while estimating the source luminosity L_s is trickier and model-dependent but achievable. Once both quantities are known, an estimate of the observed luminosity distance is given by

$$d_L = \sqrt{\frac{L_s}{4\pi\, F_o}}. \tag{4.2}$$

One can easily see that in a Minkowski spacetime this reduces to the standard notion of a distance. On the theoretical side, the intrinsic luminosity of S is related to the flux F measured by O by the integral [1]

$$L = \int_{S^2} (1+z)^2 F \, \mathrm{d}A, \tag{4.3}$$

where S^2 is the 2-sphere centred at the source S and passing through the observer O, while z is the redshift of the light. If the source radiates isotropically, we can write (4.3) as a differential relation

$$F \, \mathrm{d}A_o = \frac{L}{4\pi} \frac{\mathrm{d}\Omega_s}{(1+z)^2}, \tag{4.4}$$

where $\mathrm{d}A_o$ is an area element at the observer and $\mathrm{d}\Omega_s$ is the infinitesimal solid angle at the source (see Fig. 4.1). Therefore, the luminosity distance can be written in terms of elements of the geometry in the following way:

$$d_L = (1+z)\sqrt{\frac{\mathrm{d}A_o}{\mathrm{d}\Omega_s}}. \tag{4.5}$$

Later the quantity $\frac{\mathrm{d}A_o}{\mathrm{d}\Omega_s}$ will be related to the metric thus allowing to make a theoretical prediction for the luminosity distance in any given cosmological model.

Often instead of measuring the photon energy flux it is easier to count the number of photons received by the observer and to estimate the number of photons emitted by the source over the total lifetime of the event (or over some specified interval, for example from peak luminosity to half-maximum). Then in terms of integrated photon number flux, and photon number luminosity:

$$(F_\#)_o = \frac{(L_\#)_s}{4\pi(dA_o/d\Omega_s)}. \tag{4.6}$$

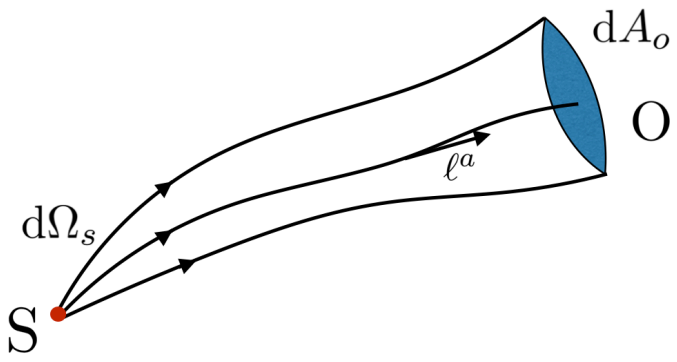

Figure 4.1: A congruence of light rays emitted at a point source S and received by an extended observer O. The luminosity distance between S and O is given by $d_L = (1+z)\sqrt{\mathrm{d}A_o/\mathrm{d}\Omega_s}$, while the area distance is $d_{area} = \sqrt{\mathrm{d}A_o/\mathrm{d}\Omega_s}$. Here ℓ^a is a tangent vector to a geodesic in the congruence.

This eliminates the two explicit redshift factors, and so allows one to empirically define the "area distance" (see figure 4.1):

$$d_{area} = \sqrt{\frac{(L_\#)_s}{4\pi\,(F_\#)_o}} = \sqrt{\frac{dA_o}{d\Omega_s}}. \tag{4.7}$$

The area distance differs from the luminosity distance by the absence of the redshift factor $1+z$ and is often more convenient to work with. Due to the similarity between the two distances, any formula for the luminosity distance that we derive later can be recast in terms of the area distance simply by eliminating the redshift factor.

The luminosity and area distances are defined for objects of a known and fixed intrinsic luminosity. Most often they are measured by performing observations in the electromagnetic spectrum. However, the notion of luminosity distance (and area distance) can also be defined for other types of waves. Recently, the gravitational wave luminosity distance has become quite popular because some binary compact object mergers serve as standard sirens - they have a fixed gravitational wave luminosity [59–61]. While the electromagnetic luminosity distance can be only used to directly test the geometry of the Universe, the gravitational luminosity distance also carries information about the gravitational dynamics. More importantly, the two notions of a distance are truly different - they agree for General Relativity, but not for most modified theories of gravity. Hence, the gravitational wave luminosity distance is particularly adept at testing gravity on cosmological scales. From now on our interest will be mainly in the standard electromagnetic luminosity distance even though most of the results can be adapted for the gravitational wave luminosity distance as well.

The angular diameter distance and the distance duality relation A different notion of a distance between a source and an observer is the angular diameter distance

defined in terms of the physical size of the source and the angle subtended at the observer (Fig. 4.2):

$$d_{angular} = \sqrt{\frac{dA_s}{d\Omega_o}}. \tag{4.8}$$

The angular diameter distance can be measured for objects of known size called standard rulers (an example is the BAO). While the luminosity and angular diameter distances are completely independent, there is a relation between them called the "Etherington distance duality" relation. The derivation of this relation rests on a number of assumptions, namely (1) photon number is conserved, (2) gravity is described by a metric theory, (3) photons travel on unique null geodesics. Given these assumptions, one can show that [1]

$$\sqrt{\frac{dA_o}{d\Omega_s}} = (1+z)\sqrt{\frac{dA_s}{d\Omega_o}}, \tag{4.9}$$

or writing that in terms of the distances defined so far:

$$d_L = (1+z)^2\, d_{angular}; \qquad d_{area} = (1+z)\, d_{angular}. \tag{4.10}$$

Exotic physics could in principle violate one or more of the assumptions above and for this reason observational tests of the distance duality relation are an active area of research [62–64].

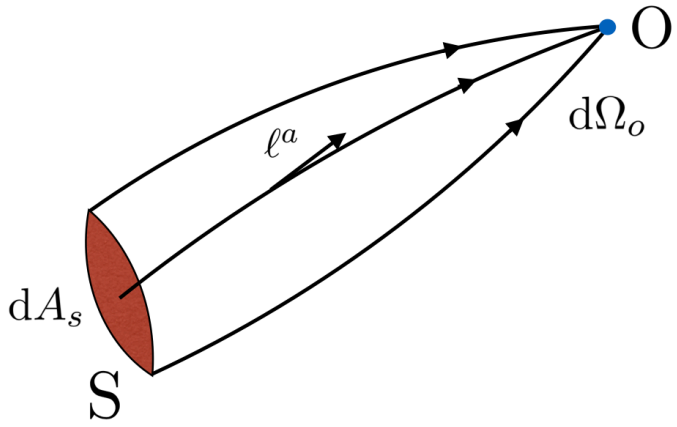

Figure 4.2: A congruence of light rays emitted at an extended source S and received by a point observer O. The angular diameter distance between S and O is given by $d_{angular} = \sqrt{dA_s/d\Omega_o}$. Here ℓ^a is a tangent vector to a geodesic in the congruence. Under suitable technical conditions the Etherington distance duality relation yields $\sqrt{dA_o/d\Omega_s} = (1+z)\sqrt{dA_s/d\Omega_o}$.

The redshift A different notion of a distance is the observed redshift of the source. To define the redshift and to isolate the different contributions, we consider a single null

geodesic (the photon trajectory), affinely parameterized by λ, and carefully distinguish the null tangent 4-vector, the null 4-wave-vector, and the null 4-momentum:

$$\ell^\alpha = \frac{dx^\alpha}{d\lambda}; \qquad k^\alpha = \tilde{\omega}\,\ell^\alpha; \qquad p^\alpha = \hbar\, k^\alpha = \hbar\tilde{\omega}\,\ell^\alpha = \tilde{E}\,\ell^\alpha. \tag{4.11}$$

Since the photon 4-momentum is, by definition, parallel transported along the null trajectory we have $\ell^\alpha \nabla_\alpha p^\beta = 0$, (no external forces act on the photon, it is in free-flight). Since the tangent vector is chosen to be affinely parameterized, we have $\ell^\alpha \nabla_\alpha \ell^\beta = 0$. Consequently the scalar $\tilde{E} = \hbar\tilde{\omega}$, (and hence $\tilde{\omega}$ itself), is constant along the null trajectory. However $\tilde{E}$ is not the locally measured energy, and $\tilde{\omega}$ is not the locally measured frequency. In general for an observer of 4-velocity V^α one has

$$E = -g_{\alpha\beta}V^\alpha p^\beta = \tilde{E}\,(-g_{\alpha\beta}V^\alpha \ell^\beta); \qquad \omega = -g_{\alpha\beta}V^\alpha k^\beta = \tilde{\omega}\,(-g_{\alpha\beta}V^\alpha \ell^\beta). \tag{4.12}$$

So E and ω can change by purely geometric factors along the photon trajectory, as they should.

Now let the source have timelike 4-velocity $(V_s)^\alpha$, and the observer have timelike 4-velocity $(V_o)^\alpha$. Then the total redshift is given by [32, 65]

$$1 + z = \frac{(g_{\alpha\beta}\; p^\alpha V^\beta)_s}{(g_{\alpha\beta}\; p^\alpha V^\beta)_o} = \frac{(g_{\alpha\beta}\; k^\alpha V^\beta)_s}{(g_{\alpha\beta}\; k^\alpha V^\beta)_o} = \frac{(g_{\alpha\beta}\; \ell^\alpha V^\beta)_s}{(g_{\alpha\beta}\; \ell^\alpha V^\beta)_o}. \tag{4.13}$$

Note that the total redshift is purely geometrical, and by definition automatically frequency independent (achromatic). Let us now introduce two fiducial 4-velocities, $(W_s)^\alpha$ and $(W_o)^\alpha$, at the source and observer and write

$$1 + z = \frac{(g_{\alpha\beta}\; \ell^\alpha W^\beta)_s}{(g_{\alpha\beta}\; \ell^\alpha W^\beta)_o}\; \frac{(g_{\alpha\beta}\; \ell^\alpha V^\beta)_s}{(g_{\alpha\beta}\; \ell^\alpha W^\beta)_s}\; \frac{(g_{\alpha\beta}\; \ell^\alpha W^\beta)_o}{(g_{\alpha\beta}\; \ell^\alpha V^\beta)_o}. \tag{4.14}$$

These fiducial 4-velocities W^α might represent, for instance, the local rest frame of the CMB, or the local rest frame of the Hubble flow. This now factorizes the total redshift into an overall cosmological/gravitational contribution, plus two peculiar velocity contributions. First, the factor

$$1 + z_c = \frac{(g_{\alpha\beta}\; \ell^\alpha W^\beta)_s}{(g_{\alpha\beta}\; \ell^\alpha W^\beta)_o} \tag{4.15}$$

represents the combined effects (as seen by fiducial observers) of cosmological expansion plus possible local variations in the gravitational field. Second, the factors

$$1 + z_p = \frac{(g_{\alpha\beta}\; \ell^\alpha V^\beta)}{(g_{\alpha\beta}\; \ell^\alpha W^\beta)} = \gamma(1 - \hat{\ell}\cdot\vec{v}) \tag{4.16}$$

represent the effect of peculiar velocities of source/observer 4-velocities V^α relative to the fiducial background W^α. (In the absence of any choice of fiducial observer W^α one cannot even begin to define the notion of "peculiar velocity".) Here we have gone to

Riemann normal coordinates at both source and observer, so $g_{\alpha\beta} \to \eta_{\alpha\beta}$, and have gone to the fiducial rest frame $W^\alpha \to (1;\vec{0})$, with $V^\alpha \to \gamma(1;\vec{v}^i)$ and $\ell^\alpha \propto (1;\hat{\ell}^i)$ to cast the peculiar redshifts in the $\gamma(1-\hat{\ell}\cdot\vec{v})$ form. (Note that in order to find the last expression for the peculiar redshift, we did not have to normalise the null vector, i.e. $\ell^\alpha \propto (1;\hat{\ell}^i)$ simply defines a direction and implies $\ell^\alpha = \ell^0(1;\hat{\ell}^i)$. But ℓ^0 drops out of the calculation and so the result is independent of the normalization chosen.) Overall we have the general result

$$1+z = (1+z_c)\,\frac{[\gamma(1-\hat{\ell}\cdot\vec{v})]_s}{[\gamma(1-\hat{\ell}\cdot\vec{v})]_o} = (1+z_c)\,\frac{1+z_{p,s}}{1+z_{p,o}} = (1+z_c)\,(1+z_D), \tag{4.17}$$

where $1+z_D$ is the Doppler contribution to the redshift. This neatly splits the total redshift into cosmological/gravitational contributions plus peculiar motion contributions. This version of the redshift equation in principle allows for arbitrarily high peculiar velocities, and arbitrarily high cosmological/gravitational redshifts.[1] (Note that peculiar velocities need not necessarily be intrinsically small. Ultimately it is an observational question as to just how small they are.)

Affine parameter distance Consider the spray of affinely parameterised null geodesics, with null tangent 4-vector $\ell^\alpha = dx^\alpha/d\lambda$, emitted from the source into some solid angle $d\Omega_s$ (Fig. 4.1). While the null affine parameters are well defined (up to constant rescaling) along each individual null geodesic, there is a priori no connection between the affine parameters on distinct null geodesics. Introduce such a connection between distinct null geodesics by enforcing (at the source and for all elements of the geodesic spray):

$$(g_{\alpha\beta}\,\ell^\alpha W^\beta)_s = -1. \tag{4.18}$$

Going (temporarily) to Riemann normal coordinates at the source, $(g_s)_{\alpha\beta} \to \eta_{\alpha\beta}$, and going to the fiducial rest frame $(W_s)^\alpha \to (1;\vec{0})$, this implies (at the source and for all elements of the geodesic spray) that we can set

$$\ell^\alpha = (1;\hat{\ell}^i). \tag{4.19}$$

Note that $\hat{\ell}^i$ will vary as one selects different elements of the geodesic spray; the geodesics in the geodesic spray are moving more-or-less in the same direction, they are not moving exactly in the same direction. That is, the affine parameters for all the null geodesics of interest are normalised by asserting that, at the source, the affine parameter along any of the elements of the null geodesic spray equals the proper time of the fiducial observer at the apex of that geodesic spray. That is, the fiducial observer at the source sees an

[1] Often the peculiar velocities are known or assumed to be small, $|\vec{v}| \ll 1$, (we have set $c \to 1$), in which case one has the simple perturbative result $1+z = (1+z_*)\,\Big(1-\hat{\ell}\cdot[\vec{v}_s-\vec{v}_o]+O(v^2)\Big)$. But such a low-peculiar-velocity approximation is by no means necessary. Sometimes one sees the Doppler factor written as $\sqrt{(1+\hat{\ell}\cdot\vec{v})/(1-\hat{\ell}\cdot\vec{v})}$, but this is actually *wrong* as it incorrectly ignores the transverse Doppler effect. (Though it does give the correct low-peculiar-velocity limit.)

outgoing (partial) light cone, and using local flatness, uses his/her time coordinate to naturally fix the near-apex affine parameter on each of the null geodesics making up the (partial) light cone. With this normalisation convention the affine parameter distance $\Delta\lambda = \lambda_o - \lambda_s$ equals the distance the photon would have travelled in Minkowski space if the Riemann tensor were forced to zero.[2] The affine parameter distance will play an important role later on when we express the luminosity distance in terms of the van Vleck determinant.

Other distances There are other notions of cosmological distances such as the paralax distance, the lookback time and the transverse comoving distance (see [66] for a review). While they are useful in different circumstances, for our purposes it is enough to work with the ones defined so far.

Having defined different notions of a cosmological distance which can be measured by light observations made by observers on Earth, it is of great interest how one can compute theoretically a cosmological distance given a particular model. We focus on the luminosity distance and we show how this can be done using two mathematical objects - the Jacobi map and the van Vleck determinant.

4.2 The Jacobi map

The Jacobi map is a powerful tool to analyse the propagation of light in a curved spacetime. Roughly speaking, it maps solid angles subtended by light geodesics at the source to the area of the locus of these geodesics at the observer. The Jacobi determinant can be used to compute the luminosity distance of a source in any spacetime. We now develop this result in more rigour.

The Jacobi determinant and the luminosity distance At a most basic level the Jacobi determinant can be understood as the Jacobian of the transformation from cartesian to angular coordinates at the observer. To be more specific, consider a congruence of null geodesics starting from a source S and ending at an extended observer O (Fig. 4.3). The locus of end points for each geodesic at O form a section of a 2-dimensional spacelike surface. More precisely, this surface is obtained by intersecting the section of the null hypersurface comprising the light geodesics with a spacelike hypersurface at the event O. There are two types of coordinates that one can introduce on this 2-dimensional surface - angular coordinates $\xi^i = (\theta, \phi)$ and cartesian coordinates $x^i = (x, y)$. The 2-dimensional Jacobi matrix is defined as

$$J^i_j \equiv \frac{\partial x^i}{\partial \xi^j} \tag{4.20}$$

[2]With this normalisation convention the cosmological/gravitational contribution to the redshift simplifies to $1 + z_c = [(g_{\alpha\beta}\, \ell^\alpha W^\beta)_o]^{-1}$. This apparent simplification is less useful than one might at first imagine since one has to propagate the chosen normalisation from source to observer.

This implies that for an area element at the observer dA_o and an angular element at the source $d\Omega_s$

$$dA_o = |\det(J)| d\Omega_s. \tag{4.21}$$

Hence, $dA_o/d\Omega_s = |\det(J)|$ and by (4.5) the luminosity distance is given by

$$d_L = |\det(J)|^{1/2}(1+z). \tag{4.22}$$

The Jacobi map At a more geometric level, the Jacobi map can be understood as arising from the geodesic deviation equation [32]. Consider again the congruence of light rays emitted by the source as shown in Fig. 4.3. It can be parametrised as

$$x^\alpha = f^\alpha(\lambda, y_i), \tag{4.23}$$

where λ is the affine parameter along each light ray, and the y_i parametrise neighbouring rays. Focussing on a single one-parameter family of light rays, one has

$$x^\alpha = f^\alpha(\lambda, y). \tag{4.24}$$

The tangent vector and the wave vector are defined as

$$\ell^\alpha = \frac{\partial f^\alpha}{\partial \lambda}, \qquad k^\alpha = \tilde{\omega}\, \ell^\alpha, \tag{4.25}$$

where $\tilde{\omega}$ is just a constant with dimension $[L^{-1}]$. The geodesic deviation vector is defined as

$$Y^\alpha = \frac{\partial f^\alpha}{\partial y}. \tag{4.26}$$

For a point source all light rays intersect at S and therefore $Y^\alpha_s = 0$. The geodesic deviation equation for the family of geodesics is

$$\frac{D^2 Y^\alpha}{d\lambda^2} = R^\alpha{}_{\beta\gamma\delta}\, \ell^\beta \ell^\gamma\, Y^\delta. \tag{4.27}$$

The equation is linear and therefore the solution at O is a linear combination of the initial values at the source S (this is a nontrivial result – see for instance [67]). Since $Y^\alpha_s = 0$, we must have

$$Y^\rho_o = \mathcal{J}^\rho_\alpha(O,S) \frac{DY^\alpha_s}{d\lambda}, \tag{4.28}$$

where $\mathcal{J}^\rho_\alpha(O,S)$ is the 4-dimensional Jacobi map. It is useful to define the following infinitesimal vectors:

$$\delta x^\alpha := Y^\alpha\, \delta y, \tag{4.29}$$

which can be thought of as pointing from one geodesic to a neighbouring one along the family, and

$$\delta\theta^\alpha := \frac{DY^\alpha}{d\lambda}\, \delta y, \tag{4.30}$$

which connects one geodesic to a neighbouring one at the source and whose magnitude is the angular separation between the two geodesics at the source. Then (4.28) becomes

$$\delta x_o^\mu = \mathcal{J}_\alpha^\mu(O,S)\ \delta\theta_s^\alpha. \tag{4.31}$$

Thus the Jacobi map maps initial directions around the source to vectors transversal to the photon beam at the observer position.

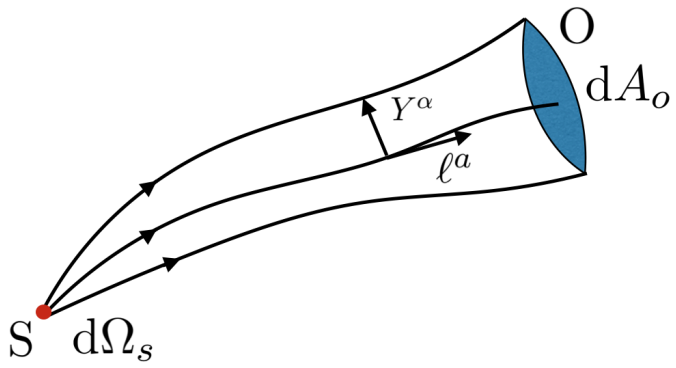

Figure 4.3: A congruence of light rays emitted at a point source S and received by an extended observer O. The luminosity distance between S and O is given by $d_L = (1+z)\sqrt{dA_o/d\Omega_s}$, while ℓ^α is a tangent vector to a geodesic in the congruence while Y^α is a transverse vector connecting different geodesics of the congruence.

The Jacobi map defined in (4.31) is a 4-dimensional map from the tangent space $T_s(\mathcal{M})$ to $T_o(\mathcal{M})$. However the vectors δx_o^μ and $\delta\theta_s^\alpha$ live in 2-dimensional subspaces of the tangent spaces at O and S, normal to the four-velocities of the observer and the source, U_o and U_s, respectively, and normal to the photon direction at O and S (see Appendix D). To find the true Jacobi map, we need to project onto these subspaces

$$J(O,S) := P_o \mathcal{J} P_s, \tag{4.32}$$

where P_o and P_s are the projectors

$$(P_s)^\mu_\nu = (\delta^\mu_\nu + U^\mu U_\nu - n^\mu n_\nu)_s, \qquad (P_o)^\mu_\nu = (\delta^\mu_\nu + U^\mu U_\nu - n^\mu n_\nu)_o, \tag{4.33}$$

and where n_s and n_o are normalized spacelike vectors pointing in the photon direction in the reference frames of the source and the observer

$$n_o = \big(\ell + (\ell\cdot U)U\big)_o, \qquad n_s = \big(\ell + (\ell\cdot U)U\big)_s. \tag{4.34}$$

Here $J(O,S)$ is a 2-dimensional map from a subspace of $T_s(\mathcal{M})$ to a subspace of $T_o(\mathcal{M})$ (it is identical to the Jacobi matrix defined in (4.20)). It follows from its definition in (4.31) and (4.32) that

$$\frac{dA_o}{d\Omega_s} = |\det J(O,S)| \tag{4.35}$$

(see general discussion in [56], and other sources [32–37, 68, 69]). Thus the luminosity distance is given by

$$d_L(S,O) = |\det J(O,S)|^{\frac{1}{2}}(1+z). \tag{4.36}$$

Solving for the Jacobi map We now want to solve the geodesic deviation equation (4.27) in order to find the Jacobi map. In practice it is easier to transform the geodesic deviation equation, which is a 2nd-order differential equation for Y^μ, into two coupled 1st-order differential equations for δx^μ and $\delta\theta^\mu$. (This is similar to what one does in Hamiltonian mechanics where a single 2nd-order differential equation for a given dynamical variable is transformed into two 1st-order differential equations for the dynamical variable and its conjugate momentum.) It follows from (4.29), (4.30) and (4.27) that

$$\frac{\mathrm{D}(\delta x^\mu)}{\mathrm{d}\lambda} = \delta\theta^\mu; \qquad \frac{\mathrm{D}(\delta\theta^\mu)}{\mathrm{d}\lambda} = R^\mu_{\nu\alpha\beta}\,\ell^\nu\ell^\alpha\delta x^\beta. \tag{4.37}$$

Equivalently

$$\frac{\mathrm{d}(\delta x^\alpha)}{\mathrm{d}\lambda} = C^\alpha_\nu(\lambda)\,\delta x^\nu + \delta\theta^\alpha; \qquad \frac{\mathrm{d}(\delta\theta^\alpha)}{\mathrm{d}\lambda} = A^\alpha_\nu(\lambda)\,\delta x^\nu + C^\alpha_\nu(\lambda)\,\delta\theta^\alpha; \tag{4.38}$$

where

$$C^\alpha_\nu(\lambda) := -\Gamma^\alpha_{\mu\nu}\ell^\mu; \qquad A^\alpha_\nu(\lambda) := R^\alpha_{\rho\mu\nu}\ell^\rho\ell^\mu. \tag{4.39}$$

In order to find the Jacobi map, this system must be solved consistently with the initial conditions

$$\delta x^\alpha(\lambda_s) = 0; \qquad (\ell^\alpha\delta\theta_\alpha)(\lambda_s) = (U^\alpha_s\delta\theta_\alpha)(\lambda_s) = 0. \tag{4.40}$$

Similar equations have also been derived in [32]. We will use these equations at the end of this section and in the next section in order to solve for the luminosity distance in some cosmologically interesting spacetimes.

We saw how the Jacobi map and its determinant in particular can be used to track the geodesic deviation of light rays in any spacetime and therefore to calculate the luminosity distance. Similar calculations can be performed by another widely used mathematical object called the van Vleck determinant.

4.3 The van Vleck determinant

Definition and applications The van Vleck determinant is an ubiquitous object that appears in many different areas of Theoretical Physics [70]. Consider a classical system of n degrees of freedom described by an action

$$S = \int \mathcal{L}(\dot{\vec{q}}, \vec{q})dt \tag{4.41}$$

Minimising the action ($\delta S < 0$) with respect to variations of the path between two points in configuration space $(\vec{q}_i, t_i)$ and $(\vec{q}_f, t_f)$ leads to the classical path γ along which the equations of motion hold. Now we can evaluate the action over γ:

$$S_\gamma = \int_\gamma \mathcal{L}(\dot{\vec{q}}, \vec{q})dt, \tag{4.42}$$

and take the second derivative of the resultant scalar with respect to each of the components of $\vec{q}_i$ and $\vec{q}_f$: $\frac{\partial^2 S_\gamma}{\partial(q_i)_k \partial(q_f)_l}$, that is we consider variations of the initial and final state of the system and how the scalar function S_γ changes with respect to that. The result is an $n \times n$ matrix whose determinant is called the van Vleck determinant. More precisely, the van Vleck determinant is defined in the following way:

$$\Delta_\gamma(\vec{q}_i, t_i, \vec{q}_f, t_f) \equiv (-1)^n \det\Big[\frac{\partial^2 S_\gamma}{\partial(q_i)_k \partial(q_f)_l}\Big] \tag{4.43}$$

It has applications in various areas of Theoretical Physics such as the WKB approximation in Quantum Mechanics, Wormhole physics and the regularisation of the stress-energy tensor in a curved spacetime [71, 72].

Geodesic interval and geodesic focussing The van Vleck determinant is also applied to geodesic deviation and geodesic focussing [73]. In order to see how, we first define the geodesic distance between two events S and O connected by a geodesic γ:

$$s_\gamma(x_s, x_o) = \begin{cases} \int_\gamma \mathrm{d}\tau & \text{timelike} \\ \int_\gamma \mathrm{d}s & \text{spacelike} \\ 0 & \text{null} \end{cases} \tag{4.44}$$

where τ is the proper time along a timelike geodesic and s is a similar parameter along a spacelike geodesic. Then the geodesic interval between S and O is given by

$$\sigma_\gamma(x_s, x_o) := \pm\frac{1}{2}[s_\gamma(x_s, x_o)]^2 \tag{4.45}$$

with a $+$sign for a spacelike geodesic and a $-$sign for a timelike geodesic. Then the van Vleck determinant is given by an expression similar to (4.43):

$$\Delta_{vV}(x_s, x_o) := -\frac{\det\{\nabla_\mu^{x_s} \nabla_\nu^{x_o} \sigma_\gamma(x_s, x_o)\}}{\sqrt{g(x_s) g(x_o)}} \tag{4.46}$$

where $g(x_s)$ is the determinant of the metric at S. The factor in front of the determinant is $(-1)^d$ where d is the number of spatial dimensions. Here we consider only the case $d = 3$.

As we will see later, the van Vleck determinant captures information about focussing or defocussing of a geodesic congruence or, in different words, it parametrises deviations from the inverse square law which is obeyed by light geodesics in a flat spacetime. More precisely, if $\Delta_{vV} = 1$, the inverse square law is obeyed. However, if $\Delta_{vV} > 1$, the geodesic congruence is focussed, while if $\Delta_{vV} < 1$, the geodesic congruence is defocussed. This property of the van Vleck determinant is useful for calculating the luminosity distance.

The luminosity distance in terms of the van Vleck determinant Now we consider a congruence of geodesics emanating from point S - the source (Fig. 4.3). This means that for a fixed source the van Vleck determinant becomes a function of the affine parameter λ along the congruence. It can also be shown (see [73]) that the van Vleck determinant obeys a first-order differential equation

$$\frac{d\Delta_{vV}}{d\lambda} = \left(\frac{2}{\lambda - \lambda_s} - \theta\right)\Delta_{vV}, \tag{4.47}$$

where λ_s is the affine parameter at the source and θ is the expansion of the congruence:

$$\theta = \frac{1}{dA/d\Omega_s}\frac{d(dA/d\Omega_s)}{d\lambda}. \tag{4.48}$$

The solution of (4.47) at point O - the observer, is given by

$$\Delta_{vV}(\lambda_o, \lambda_s) = \frac{(\lambda_o - \lambda_s)^2}{(dA_o/d\Omega_s)}. \tag{4.49}$$

Therefore, Eqn. (4.5) immediately implies that there is a relation between the luminosity distance and the van Vleck determinant:

$$d_L = (1+z)(\lambda_o - \lambda_s)\Delta_{vV}^{-1/2}. \tag{4.50}$$

Eqn. (4.49) also implies a relation between the Jacobi determinant and the van Vleck determinant:

$$|\det J(O,S)| = (\lambda_o - \lambda_s)^2\Delta_{vV}^{-1}. \tag{4.51}$$

Thus Eqn.(4.50) is equivalent to Eqn.(4.36). However, in (4.50) the affine parameter distance is separate from the focussing/defocussing contributions while in (4.36) the two terms are lumped together. Thus, (4.50) has a more clear physical interpretation.

Limitations of the luminosity distance Despite the usefulness of the notion of the luminosity distance and its wide applications in cosmology and cosmography, there are circumstances where the luminosity distance fails to give an accurate representation of the separation between the source and the observer. This happens for example, near a conjugate point of the null geodesic congruence emanating from the source. A conjugate point is a point where the transverse vector field Y^μ vanishes: $Y^\mu = 0$. At such a point the Jacobi determinant is zero, while the van Vleck determinant blows up: $\det(J) = 0$, $\Delta_{vV} = \infty$. As a result, an observer located at the conjugate point would infer a vanishing luminosity distance of the source: $d_L(S,O) = 0$, while an observer located close to a conjugate point would infer a highly misrepresented value of the luminosity distance. An example of a geometry which leads to such an effect is FLRW with $k = 1$. The spacial slices are 3-spheres and any two antipodal points on these 3-spheres would be conjugate points. This means that a source and an observer located at antipodal points, i.e. where the metric distance between them is maximal, would have vanishing luminosity distance. This failure of the luminosity distance at conjugate points is actually an opportunity because it allows for tests of the topology of the Universe. So far there is no observational evidence for a non-trivial topology of our observable patch of the Universe.

4.4 Cosmography in FLRW

Now we will use the results for the luminosity distance derived so far in order to show how the geometry of the Universe can be tested in a model-independent way. First, however, we need to establish a general result about conformal transformations. Many of the cosmological spacetimes that we are interested in can be related to simpler spacetimes through conformal transformations (local rescalings) of the metric. Therefore, it would be extremely valuable, if we know how the main quantities of interest - luminosity distance, redshift, affine parameter distance, etc., transform under conformal transformations.

Conformal transformations We define a conformal transformation (or Weyl transformation) to be a local rescaling of the metric:

$$g_{\mu\nu} = \exp(2\Phi)\hat{g}_{\mu\nu}, \tag{4.52}$$

where $\Phi(x)$ is an arbitrary function of space and time (note that we choose to write the general function in front of the metric as an exponential because it will be computationally useful later on). The behaviour of null geodesics under conformal transformations implies specific transformation properties for the luminosity distance. Under the conformal transformation (4.52) the paths of the null geodesics are unaffected, while the affine parameters are non-trivially related by [72, 74, 75] (see Appendix C for details)

$$d\lambda = \exp(2\Phi)d\hat{\lambda}, \tag{4.53}$$

which implies the transformation law for the null vector ℓ^μ

$$\ell^\mu = \exp(-2\Phi)\hat{\ell}^\mu. \tag{4.54}$$

Area transformations are straightforward while angles remain invariant

$$dA_o = \exp(2\Phi_o)d\hat{A}_o, \qquad d\Omega_s = d\hat{\Omega}_s. \tag{4.55}$$

This implies

$$\frac{dA_o}{d\Omega_s} = \exp(2\Phi_o)\,\frac{d\hat{A}_o}{d\hat{\Omega}_s}, \tag{4.56}$$

which immediately leads to a transformation law for the area distance:

$$d_{area} = \sqrt{\frac{dA_o}{d\Omega_s}} = \exp(\Phi_o)\,\sqrt{\frac{d\hat{A}_o}{d\Omega_s}} = \exp(\Phi_o)\,\hat{d}_{area}. \tag{4.57}$$

For a timelike vector $V^\mu = \frac{dx^\mu}{d\tau}$ with proper time τ the transformation laws are the following:

$$d\tau = \exp(\Phi)d\hat{\tau}, \qquad V^\mu = \exp(-\Phi)\hat{V}^\mu. \tag{4.58}$$

Therefore,

$$\begin{aligned} g_{\mu\nu}\ell^\mu V^\nu &= (\exp(2\Phi)\hat{g}_{\mu\nu})(\exp(-2\Phi)\hat{\ell}^\mu)(\exp(-\Phi)\hat{V}^\nu) \\ &= \exp(-\Phi)(\hat{g}_{\mu\nu}\hat{\ell}^\mu\hat{V}^\nu). \end{aligned} \tag{4.59}$$

This implies that the redshift transforms as

$$\begin{aligned} 1+z &= \frac{(g_{\mu\nu}\ell^\mu V^\nu)_s}{(g_{\mu\nu}\ell^\mu V^\nu)_o} \\ &= \exp(\Phi_o - \Phi_s)\frac{(\hat{g}_{\mu\nu}\hat{\ell}^\mu\hat{V}^\nu)_s}{(\hat{g}_{\mu\nu}\hat{\ell}^\mu\hat{V}^\nu)_o} \\ &= \exp(\Phi_o - \Phi_s)\ (1+\hat{z}). \end{aligned} \tag{4.60}$$

Hence the luminosity distance transforms as

$$\begin{aligned} d_L &= (1+z)\sqrt{\frac{dA_o}{d\Omega_s}} \\ &= \exp(2\Phi_o - \Phi_s)(1+\hat{z})\sqrt{\frac{d\hat{A}_o}{d\hat{\Omega}_s}} \\ &= \exp(2\Phi_o - \Phi_s)\hat{d}_L. \end{aligned} \tag{4.61}$$

In applications one would typically take $g_{\mu\nu}$ to be the physical spacetime metric while $\hat{g}_{\mu\nu}$ would be some (not directly physical) conformal deformation of spacetime useful for calculational purposes. Observe that the Jacobi matrix has the simple transformation law

$$J^i{}_j = \exp(\Phi_o)\ \hat{J}^i{}_j; \qquad \det(J) = \exp(2\Phi_o)\ \det(\hat{J}). \tag{4.62}$$

However the transformation law for the affine parameter distance (and therefore also for the van Vleck determinant) is more subtle and is given by an integral, not by a simple local rescaling:

$$\Delta\hat{\lambda} = \hat{\lambda}_o - \hat{\lambda}_s = \int_s^o \exp[-2\Phi(\lambda')]\ d\lambda'. \tag{4.63}$$

The fact that the affine parameter distance (and hence the van Vleck determinant) does not simply rescale under conformal transformations is actually potentially useful, not a hindrance – it implies that the van Vleck determinant might actually simplify much more radically under conformal transformations.

The $d_L - z$ relation in FLRW spacetimes A particularly simple example of conformal transformations of the spacetime metric comes from considering FLRW cosmologies. Consider the usual FLRW cosmology with a spacetime metric:

$$ds^2_{FLRW} = -dt^2 + a(t)^2\left\{\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\ d\phi^2)\right\}, \tag{4.64}$$

where $k \in \{-1, 0, +1\}$, r is dimensionless and a has units of distance. Introduce a new time parameter (conformal time) by defining $dt/a(t) = d\eta$ and then subsequently reparametrise $a(t)$ to $a(\eta)$ to write

$$ds^2_{FLRW} = a(\eta)^2 \left[-d\eta^2 + \frac{dr^2}{1-kr^2} + r^2 d\Omega \right], \tag{4.65}$$

where $d\Omega = d\theta^2 + \sin^2\theta \, d\phi^2$. Now the term in braces

$$ds^2_E = -d\eta^2 + \frac{dr^2}{1-kr^2} + r^2 d\Omega \tag{4.66}$$

is just the spacetime metric for the Einstein static universe, which we shall use as a computational aid, the real physical universe not being static. Then

$$d_{L,FLRW} = \left(\frac{a_o^2}{a_s} \right) d_{L,E}, \tag{4.67}$$

where $d_{L,E}$ is the luminosity distance in the Einstein static universe. It is dimensionless and unphysical but extremely easy to calculate. For a source placed at $r = 0$, and observer located at r, the flux is simply

$$F = \frac{1}{(1+z_E)^2} \, \frac{L}{4\pi r^2}, \tag{4.68}$$

where z_E is the redshift in the Einstein static universe. It depends only on the peculiar velocities (there are no local inhomogeneities or cosmological effects) so

$$d_{L,E} = (1+z_E) r = \left(\frac{1+z_{p,s}}{1+z_{p,o}} \right) r. \tag{4.69}$$

In order to determine r in terms of conformal time, consider a radial null geodesic for which

$$d\eta = \frac{dr}{\sqrt{1-kr^2}}, \tag{4.70}$$

so we have

$$\Delta\eta = \eta_o - \eta_s = \int_0^r \frac{d\tilde{r}}{\sqrt{1-k\tilde{r}^2}} = \frac{\arcsin\left(\sqrt{k} r\right)}{\sqrt{k}}, \tag{4.71}$$

which implies

$$r = \frac{\sin\left(\sqrt{k}\Delta\eta\right)}{\sqrt{k}}. \tag{4.72}$$

For each of the three different values of k, this reduces to

$$r = \begin{cases} \sin(\Delta\eta) & \text{for } k = 1 \\ \sinh(\Delta\eta) & \text{for } k = -1 \\ \Delta\eta & \text{for } k = 0. \end{cases} \tag{4.73}$$

Eqn.(4.72) implies

$$\frac{dA_o}{d\Omega_s} = r^2 = \frac{\sin^2\left(\sqrt{k}\Delta\eta\right)}{k}. \tag{4.74}$$

The affine null parameter is simply the conformal time $\lambda_E = \eta$ which allows us extract the van Vleck determinant:

$$\Delta_{vV,E} = \frac{k(\Delta\eta)^2}{\sin^2\left(\sqrt{k}\Delta\eta\right)}. \tag{4.75}$$

Then for $k = 0$ (flat spatial slices) $\Delta_{vV,E} \equiv 1$, so there is no deviation from the inverse square law. For $k = -1$ (hyperbolic spatial slices) $\Delta_{vV,E} = \frac{(\Delta\eta)^2}{\sinh^2(\Delta\eta)} < 1$. That is, one has defocussing induced suppression of the inverse square law. Finally, for $k = 1$ (hyperspherical spatial slices) $\Delta_{vV,E} = \frac{(\Delta\eta)^2}{\sin^2(\Delta\eta)} > 1$, one has focussing induced enhancement of the inverse square law.

Now going to the physical FLRW spacetime one has

$$\begin{aligned} d_{L,FLRW} &= \left(\frac{a_o^2}{a_s}\right)\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right) r \\ &= a_o(1+z_c)\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right)\frac{\sin\left(\sqrt{k}\Delta\eta\right)}{\sqrt{k}}. \end{aligned} \tag{4.76}$$

This expression cleanly separates the various distinct physical contributions to the luminosity distance in a general FLRW spacetime. Of course, we have a very similar expression for the area distance:

$$d_{area,FLRW} = a_o\frac{\sin\left(\sqrt{k}\Delta\eta\right)}{\sqrt{k}}. \tag{4.77}$$

The luminosity distance depends on the redshift in two distinct ways - first, through the redshift factor $1 + z = (1 + z_c)\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right)$ that appears at the front, and second, through the conformal time interval $\Delta\eta$ that appears in the sine function. In fact, the first step in any cosmographic analysis is to expand the $\Delta\eta$ as a power series of the cosmological redshift z_c:

$$\Delta\eta = P(1+z_c) \qquad \text{with the normalisation} \qquad P(1) = 0. \tag{4.78}$$

We next show how this is performed in FLRW.

Cosmographic expansion in FLRW spacetimes Cosmographic analyses make good physical sense whenever the cosmological spacetime can be sliced by spacelike hypersurfaces which can be factored into an overall "size of the universe" (depending only on some convenient global time parameter t, possibly some proper time measured

by some class of fiducial observers) multiplied by something that depends on the "shape" of the spatial slices. That is, take

$$ds^2 = -N(t,\vec{x})^2\,dt^2 + a(t)^2\,[g_{shape}(t,\vec{x})]_{ij}\,dx^i dx^j. \tag{4.79}$$

This form of the metric is a variant on the notion of "synchronous gauge" and might be called "pre-synchronous gauge", or "conformally synchronous gauge".[3]

Whenever such a decomposition makes sense one can further construct a "conformal time" coordinate $d\eta = dt/a(t)$ and use this to recast the spacetime metric as

$$ds^2 = a(\eta)^2\left\{-N(\eta,\vec{x})^2\,d\eta^2 + \,[g_{shape}(\eta,\vec{x})]_{ij}\,dx^i dx^j\right\}. \tag{4.80}$$

As long as this can be done (and this is a rather mild constraint on the cosmology), one can undertake a cosmographic analysis either in terms of the t-time derivatives as in (1.5)

$$H = \frac{\dot{a}}{a}; \qquad q = -\frac{1}{H^2}\frac{\ddot{a}}{a}; \qquad j = \frac{1}{H^3}\frac{\dddot{a}}{a} \tag{4.81}$$

or in terms of η-time derivatives

$$\mathcal{H} = \frac{a'}{a}; \qquad \mathcal{Q} = -\frac{1}{\mathcal{H}^2}\frac{a''}{a}; \qquad \mathcal{J} = \frac{1}{\mathcal{H}^3}\frac{a'''}{a}. \tag{4.82}$$

Here we do it for the latter case but the former case is very similar. Expanding the scale factor in a truncated Taylor series around the "observer" conformal time η_o, the conformal time equivalent of the present epoch, we obtain

$$a(\eta) = a(\eta_o)\left[1 + \mathcal{H}_o(\eta-\eta_o) - \frac{\mathcal{H}_o^2\mathcal{Q}_o}{2}(\eta-\eta_o)^2 + \frac{\mathcal{H}_o^3\mathcal{J}_o}{6}(\eta-\eta_o)^3 + O(\eta-\eta_o)^4\right], \tag{4.83}$$

and, using $1 + z_c = a_o/a(\eta)$, we can derive an expansion of z_c in terms of $\mathcal{H}_o(\eta-\eta_o)$. We find

$$\begin{aligned} z_c(\eta) = &-[\mathcal{H}_o(\eta-\eta_o)] + \frac{2+\mathcal{Q}_o}{2}[\mathcal{H}_o(\eta-\eta_o)]^2 - \frac{\mathcal{J}_o + 6\mathcal{Q}_o + 6}{6}[\mathcal{H}_o(\eta-\eta_o)]^3 \\ &+ O\Big([\mathcal{H}_o(\eta-\eta_o)]^4\Big). \end{aligned} \tag{4.84}$$

Reverting the series, we obtain

$$\mathcal{H}_o(\eta-\eta_o) = -\,z_c + \frac{2+\mathcal{Q}_o}{2}z_c^2 - \frac{3\mathcal{Q}_o^2 + 6\mathcal{Q}_o + 6 - \mathcal{J}_o}{6}z_c^3 + O(z_c^4). \tag{4.85}$$

[3]Observe that the phrase "synchronous gauge", where $N(t,\vec{x}) = 1$, is somewhat of a misnomer. When enforced globally it enforces the existence of a timelike geodesic vorticity-free congruence $V = dt$. The conformally synchronous gauge is less restrictive, only requiring the existence of a timelike vorticity-free congruence $V = N^{-1}\,dt$, that is not necessarily geodesic. Note we also want $\partial_t \det([g_{shape}(t,\vec{x})]_{ij})$ to be perturbatively small.

The analogous expression in terms of the t-time is

$$H_o(t-t_o) = -z_c + \frac{2+q_o}{2}z_c^2 - \frac{3q_o^2+6q_o+6-j_o}{6}z_c^3 + O(z_c^4). \tag{4.86}$$

Such perturbative expansions can in principle be carried out to arbitrarily high order, and their usefulness is limited only by the extent to which we can measure, estimate, or theoretically predict the Hubble, deceleration, jerk, and higher-order parameters. The key point is that these cosmographic series make sense under very generic conditions, whenever one is able to peel off an "overall size" and a natural "global time" for the universe. These cosmographic series will generically only be part of the full analysis, (for instance they ignore peculiar velocities and the effect of local clumping), but if the "overall size" $a(t)$ or equivalently $a(\eta)$ is chosen appropriately, they can easily be the dominant feature contributing to the luminosity distance.

Now we turn to the specific case of a FLRW spacetime, and in particular we restrict ourselves to the case $k = 0$ in agreement with observations (the cases $k = \pm 1$ can be handled similarly). Inserting the expression (4.85) in (4.76) with $k = 0$ we obtain

$$d_L(z) = \frac{a_o}{\mathcal{H}_o}\left[z - \frac{\mathcal{Q}_o}{2}z_c^2 + \frac{3\mathcal{Q}_o^2+3\mathcal{Q}_o-\mathcal{J}_o}{6}z_c^3 + O(z_c^4)\right]. \tag{4.87}$$

Note that

$$\frac{a_o}{\mathcal{H}_o} = \frac{a_o}{(da/d\eta)_o/a_o} = \frac{a_o}{(da/dt)_o} = \frac{1}{H_o}, \tag{4.88}$$

where H_o is the usual Hubble parameter measured by the astronomers. That is, in any FLRW cosmology

$$d_L(z_c) = \frac{1}{H_o}\left[z_c - \frac{\mathcal{Q}_o}{2}z_c^2 + \frac{3\mathcal{Q}_o^2+3\mathcal{Q}_o-\mathcal{J}_o}{6}z_c^3 + O(z_c^4)\right]. \tag{4.89}$$

Furthermore, $\mathcal{Q}_o$ and $\mathcal{J}_o$ can be converted to q_o and j_o which gives

$$d_L(z_c) = \frac{1}{H_0}\Big[z_c + \frac{1-q_0}{2}z_c^2 - \frac{1-q_0-3q_0^2+j_0}{6}z_c^3 + O(z_c^4)\Big]. \tag{4.90}$$

Peculiar velocities From the above we see that in a $k = 0$ FLRW universe without peculiar velocities

$$d_L = \frac{a_o}{\mathcal{H}_o}(1+z_c)P(z_c), \tag{4.91}$$

where $P(z_c)$ is the specific polynomial

$$P(z_c) = z_c - \frac{2+\mathcal{Q}_o}{2}z_c^2 + \frac{3\mathcal{Q}_o^2+6\mathcal{Q}_o+6-\mathcal{J}_o}{6}z_c^3 + O(z_c^4). \tag{4.92}$$

If one now adds peculiar velocities, then the only change is that we substitute the prefactor $(1+z_c)$ with the total redshift $(1+z) = (1+z_c)(1+z_D)$

$$d_L = \frac{a_o}{\mathcal{H}_o}(1+z)P(z_c), \tag{4.93}$$

where $(1+z_D)$ is the Doppler contribution to the redshift which arises due to the peculiar velocities of the source and the observer:

$$1 + z_D = \frac{1 + z_{p,s}}{1 + z_{p,o}}. \tag{4.94}$$

Then, assuming that peculiar velocities, and hence z_D, are small, we have

$$z_c = \frac{1+z}{1+z_D} - 1 \approx z - (1+z)z_D + O(z_D^2). \tag{4.95}$$

Therefore

$$d_L = \frac{a_o}{\mathcal{H}_o}(1+z)P\left(z - (1+z)z_D + O(z_D^2)\right), \tag{4.96}$$

implying

$$d_L = \frac{a_o}{\mathcal{H}_o}\left\{(1+z)P(z) - (1+z)^2 P'(z)z_D + O(z_D^2)\right\}. \tag{4.97}$$

This gives an explicit formula for estimating the potential effect of peculiar velocities on luminosity distance. The fractional size of the effect is easily seen to be

$$\frac{\Delta d_L}{d_L} = -(1+z)\frac{P'(z)}{P(z)}z_D + O(z_D^2). \tag{4.98}$$

Evaluating explicitly the polynomial $P(z)$ to $O(z^3)$, we can find an expression for d_L to $O(z^2)$ and $O(z_D)$

$$\begin{aligned} d_L = \frac{a_o}{\mathcal{H}_o}\Bigg[&- z_D + \Big(1 + \mathcal{Q}_o z_D\Big)z \\ &- \left(\frac{\mathcal{Q}_o}{2} + \frac{3\mathcal{Q}_o^2 + 2\mathcal{Q}_o - \mathcal{J}_o}{2}z_D\right)z^2 + O(z^3) + O(z_D^2)\Bigg]. \end{aligned} \tag{4.99}$$

As a further application we might consider a situation where on average the peculiar Doppler shifts are zero: $\langle z_D \rangle = 0$. Then on average

$$\langle d_L \rangle = \frac{a_o}{\mathcal{H}_o}\left[z - \frac{\mathcal{Q}_o}{2}z^2 + O(z^3) + O(z_D^2)\right], \tag{4.100}$$

and so

$$d_L - \langle d_L \rangle = -\frac{a_o\, z_D}{\mathcal{H}_o}\left[1 - \mathcal{Q}_o z + \left(\frac{3\mathcal{Q}_o^2 + 2\mathcal{Q}_o - \mathcal{J}_o}{2}\right)z^2 + O(z^3) + O(z_D)\right]. \tag{4.101}$$

This could be used, in principle, to estimate peculiar Doppler redshifts z_D (and so peculiar velocities) at various values of total redshift z. This would be done by first neglecting peculiar Doppler redshifts to naively fit $d_L(z)$ to the supernova data, thereby determining the cosmographic coefficients, and then binning the supernovae into small redshift bins to observationally determine $d_L - \langle d_L \rangle$.

The $d_L(z)$ relation in the ΛCDM model The cosmographic approach developed so far does not require the specification of a theory of gravity. It rests only on the assumption that the metric is of the form (4.79). In the case of FLRW, no use of the Friedmann equations is made in order to derive an expression for the luminosity distance as a power series of the redshift. The assumption of homogeneity and isotropy, together with the definition of the luminosity distance, redshift and the cosmographic parameters together imply (4.89).

Now we want to look at the luminosity distance in a model which requires the specification of a theory of gravity, namely the ΛCDM model which is based on the General Theory of Relativity. The Friedmann equations are used to relate the Hubble function to the matter content of the universe in the form of the parameters $\Omega_{r,0}$, $\Omega_{m,0}$, $\Omega_{\Lambda,0}$, $\Omega_{k,0}$:

$$H^2 = H_0^2\Big[\Omega_{r,0}\Big(\frac{a_0}{a}\Big)^4 + \Omega_{m,0}\Big(\frac{a_0}{a}\Big)^3 + \Omega_{k,0}\Big(\frac{a_0}{a}\Big)^2 + \Omega_{\Lambda,0}\Big]. \tag{4.102}$$

Setting $a_0 = 1$ and going to conformal time, this implies

$$d\eta = H_0^{-1}\Big[\Omega_{r,0} + \Omega_{m,0}a + \Omega_{k,0}a^2 + \Omega_{\Lambda,0}a^4\Big]^{-1/2}da, \tag{4.103}$$

or writing everything in terms of the cosmological redshift z_c:

$$d\eta = -H_0^{-1}\Big[\Omega_{r,0}(1+z_c)^4 + \Omega_{m,0}(1+z_c)^3 + \Omega_{k,0}(1+z_c)^2 + \Omega_{\Lambda,0}\Big]^{-1/2}dz_c. \tag{4.104}$$

Integrating from the source S to the observer O gives

$$\eta_o - \eta_s = -H_0^{-1}\int_{1+z_c}^{1}\Big[\Omega_{r,0}x^4 + \Omega_{m,0}x^3 + \Omega_{k,0}x^2 + \Omega_{\Lambda,0}\Big]^{-1/2}dx. \tag{4.105}$$

Therefore, the luminosity distance is

$$\begin{aligned} d_L &= (1+z_c)(\eta_o - \eta_s) \\ &= \frac{1+z_c}{H_0}\int_1^{1+z_c}\Big[\Omega_{r,0}x^4 + \Omega_{m,0}x^3 + \Omega_{k,0}x^2 + \Omega_{\Lambda,0}\Big]^{-1/2}dx. \end{aligned} \tag{4.106}$$

As before peculiar velocities can be introduced by the substitution $(1+z_c) \to (1+z)$ in the prefactor. Eqn. (4.106) is far more specific than (4.90), because it relies on a particular cosmological model, whereas (4.90) is model-independent. The lesson from this is that while luminosity distance observations are very useful for testing and constraining the geometry of the Universe, they can also be indirectly used to test the theory of gravity and the matter content of the Universe.

Summary In this chapter we looked at several different definitions of a distance in Cosmology, focussing in particular on the luminosity distance which can be measured for a class of astronomical objects with a fixed intrinsic luminosity known as standard candles (or more precisely standardisable candles). The luminosity distance is a very useful tool

to test the geometry of the Universe because it can be computed theoretically for a particular metric and then compared to observations. On the theoretical side, there are two tools that are invaluable in the computation of the luminosity distance - the van Vleck determinant and the Jacobi map. Furthermore, since the luminosity distance scales in a well-known way under conformal transformations of the metric, we can use these tools to calculate the luminosity distance for a variety of spacetimes. The most useful approach to test the geometry of the Universe is the cosmographic approach where the theoretical luminosity distance is expanded as a power series of the redshift. Observations of both luminosity distance and redshift then allow to constrain the cosmographic parameters, H_0, q_0, s_0, etc., and hence reconstruct at least the most recent expansion history of the Universe. If one introduces a specific cosmological model, one can in addition test the gravitational dynamics or/and the matter content of the Universe. Next we are going to use these techniques to explore more sophisticated cosmological models, namely Conformally FLRW models and perturbed FLRW models. We will also look at a specific toy model where a more generalised cosmographic expansion can be performed.

5

Beyond the ΛCDM model

5.1 Motivation for going beyond the ΛCDM model

In the Introduction we explained why the ΛCDM model might not be the final word as far as classical cosmology is concerned. One of the most exciting possible deviations from the standard cosmological model is if dark energy is not simply a cosmological constant or a vacuum energy, but a dynamical field which is changing with time and is possibly inhomogeneous in space [13, 76–79]. Such a modification would influence light propagation in the Universe and therefore would have an observable effect on some of the distances discussed in the previous section, in particular the luminosity distance.

Even within the standard cosmological model the role of inhomogeneities from cosmic structure on the propagation of light and their effect on the inference of the cosmographic parameters guiding the expansion of the Universe are pretty much an open question. It is usually assumed that light propagates on the FLRW background completely oblivious to any inhomogeneities. The reason for that is that the Universe is statistically homogeneous and hence the effect of different inhomogeneities would cancel out. However, it would still be interesting to explore even within the ΛCDM framework how the standard cosmographic approach can be modified and adapted to a more realistic inhomogeneous Universe and whether this would lead to a significant change in our estimation of the cosmographic parameters.

Taking these two considerations into account, in this chapter we will explore two deviations from the FLRW background on which electromagnetic waves are usually assumed to propagate [56, 57]. The first is a Conformal FLRW spacetime in which the FLRW background is conformally deformed by an arbitrary large conformal factor in a way consistent with CMB physics. The second is FLRW with linear scalar perturbations. The latter can be thought of as arising solely due to cosmic structure thus in principle allowing for tests of the impact of inhomogeneities on the estimation of cosmological parameters entirely within the paradigm of ΛCDM, or they can also be thought of as arising due to inhomogeneities in the distribution of dark energy thus allowing to probe models beyond the ΛCDM.

5.2 Conformal FLRW

The CFLRW spacetimes A conformal FLRW spacetime (CFLRW spacetime) is any spacetime whose metric is conformally related to a FLRW one with an arbitrary position-dependent conformal factor [65]. This in turn implies that any CFLRW spacetime is conformal to an Einstein static spacetime:

$$ds^2_{CFLRW} = g_{\alpha\beta}\, dx^\alpha dx^\beta = f(x)^2 \left\{ -d\eta^2 + \frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\; d\phi^2) \right\}, \qquad (5.1)$$

where $f(x)$ is an arbitrary function of spacetime.

The physical reason that CFLRW spacetimes are important is because of the observed smoothness of the CMB: In view of the fact that the CMB is smooth to $O(10^{-5})$, the conformal mode is the only deviation from FLRW that has the slightest chance of becoming non-perturbatively large between last scattering and the current epoch without grossly distorting the CMB [65]. More precisely, if we write $f(x) \equiv \exp[\Theta(\eta, \vec{x})]$ then in order to be consistent with CMB observations, we demand that at the last scattering surface, at $z \approx 1100$, the variation of Θ is sufficiently small: $\Delta\Theta_{ls} \leq 10^{-5}$, while it can be arbitrarily large for $z \ll 1100$. The null geodesics in this model coincide with the null geodesics in FLRW spacetime and therefore the angular distribution of the CMB is preserved - the only signature of the conformal factor is an isotropic, observer-dependent temperature shift:

$$T_{loc} = T_c \exp[-\Theta_o] \qquad (5.2)$$

where T_{loc} is the local CMB temperature measured by the observer, T_c is the cosmological temperature of the CMB which would have been inferred in the absence of any conformal deformation, and Θ_o is the value of Θ_o for the observer.

The mathematical reason that CFLRW spacetimes are important is because (from the point of view of luminosity distance, area distance, and photon propagation) they are almost as easy to analyse as the FLRW spacetimes. Although the CFLRW cosmologies are relatively easy to analyse mathematically, to date only relatively crude bounds have been put on the conformal factor [65]. (More complicated deviations from FLRW are generally treated perturbatively, though there is an ongoing debate as to the significance of non-perturbative effects. See for instance [80–85].)

The analysis of photon propagation in the Einstein static (unphysical) reference spacetime carries through (completely unchanged) as in the previous subsection, and going to the physical CFLRW spacetime one has

$$d_{L,CFLRW} = f_o\,(1+z_*)\,\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right)\,\frac{\sin\left(\sqrt{k}\Delta\eta\right)}{\sqrt{k}}, \qquad (5.3)$$

where f_o is the value of the function $f(x)$ evaluated at the spacetime coordinates of the observer, while z_* is the cosmological/local gravity contribution to the redshift, which arises from local conformal inhomogeneities: $1 + z_* = f_o/f_s$. One could factorise the conformal factor as $f(x) = a(\eta)\tilde{f}(x)$, i.e. into a cosmological contribution $a(\eta)$

characterising the overall change in volume of the spatial slices, and a second contribution $\tilde{f}(x)$ characterising volume-preserving conformal distortions of the spatial slices [65]. Then the cosmological and local contributions to the redshift further factorise as

$$(1+z_*) = \frac{f_o}{f_s} = \frac{a_o}{a_s}\,\frac{(\tilde{f})_o}{(\tilde{f})_s} = (1+z_c)(1+z_{loc}). \tag{5.4}$$

Collecting all of this leads to the general result

$$d_{L,CFLRW} = a_o\,\tilde{f}_o(1+z_c)\,(1+z_{loc})\,\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right)\,\frac{\sin\left(\sqrt{k}\,\Delta\eta\right)}{\sqrt{k}}. \tag{5.5}$$

This formula cleanly separates the various distinct physical contributions to the luminosity distance in a general CFLRW spacetime. Note in particular, that the factor $f_o = a_o\tilde{f}_o$ is common to all the objects one might look at. So there is no real loss of generality, if we simply absorb this into the definition of distance and assert that

$$d_{L,CFLRW} \propto (1+z_c)\,(1+z_{loc})\,\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right)\,\frac{\sin\left(\sqrt{k}\,\Delta\eta\right)}{\sqrt{k}}. \tag{5.6}$$

This is equivalent to considering a luminosity modulus instead of a luminosity distance, $\mu_{L,CFLRW} = \ln(d_{L,CFLRW})$, and agreeing to ignore a common offset.

Cosmography in CFLRW spacetimes Adopting a modified cosmographic analysis, one can invert $a(\eta)$ to find $\eta(a)$, and use $a_o/a_s = 1+z_c$ to rewrite the conformal time as a function of the cosmological contribution to the redshift:

$$\Delta\eta = h(1+z_c). \tag{5.7}$$

This implies

$$d_{L,CFLRW} = a_o\,(1+z_c)(1+z_{loc})\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right)\,\frac{\sin\left[\sqrt{k}\,h(1+z_c)\right]}{\sqrt{k}}. \tag{5.8}$$

This quite formal result seems to be as far as one can go without making some approximations and resorting to perturbation theory.

General spacetimes In view of the results presented above we can argue that non-perturbatively the best we can hope for in any completely general spacetime is that

$$d_L = a_o\,(1+z)\,F(z_c); \qquad d_{area} = a_o\,F(z_c); \qquad F(0) = 0; \tag{5.9}$$

for some function $F(z_c)$ of the cosmological contribution to the total redshift

$$1+z = (1+z_c)\,(1+z_{loc})\,\left(\frac{1+z_{p,s}}{1+z_{p,o}}\right). \tag{5.10}$$

5.3 Perturbed FLRW

Now we look at a linearly perturbed FLRW metric with 2 scalar modes in the Newtonian gauge (Appendix A)

$$\mathrm{d}s^2 = a^2(\eta)\Big[-(1+2\Psi(\vec{x},\eta))\mathrm{d}\eta^2 + (1+2\Phi(\vec{x},\eta))\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\Big], \tag{5.11}$$

where Ψ and Φ are the so called Bardeen potentials. Even though we work in a particular gauge, Ψ and Φ are gauge-invariant quantities. We ignore vector and tensor perturbation - the former are experimentally very small while the latter propagate away. This is a sensible thing to do since at first order the different types of perturbations do not get mixed under gauge transformations. Assuming that the stress-energy tensor is anisotropic and using Einstein's equations sets $\Phi = -\Psi$. However, here we will make no assumptions of that sort and so we will keep Φ and Ψ as general, though perturbatively small, functions of spacetime.

From now on all quantities are expressed to first order in terms of the Bardeen potentials. To first order the metric (5.11) can be cast in the form

$$\mathrm{d}s^2 = f^2(\eta,\vec{x})\Big[-(1+2\xi)\mathrm{d}\eta^2 + \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\Big], \tag{5.12}$$

where the overall conformal factor is

$$f(\eta,\vec{x}) = a(\eta)(1+2\Phi)^{\frac{1}{2}} \approx a(\eta)(1+\Phi), \tag{5.13}$$

and

$$\xi = \Psi - \Phi. \tag{5.14}$$

Calculating the redshift Now look at the simplified one-mode metric

$$\mathrm{d}\hat{s}^2 = -(1+2\xi)\mathrm{d}\eta^2 + \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j, \tag{5.15}$$

which is simply background Minkowski space plus a perturbation: $\hat{g}_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. We require that the 4-velocities of the source and the observer are normalised $\hat{U}^\mu\hat{U}^\nu\hat{g}_{\mu\nu} = -1$ and we again first consider the case of zero peculiar velocities. This implies that to first order

$$\hat{U}^\mu_s = \hat{U}^\mu_o = (1-\xi,\vec{0}). \tag{5.16}$$

The source emits light which travels on null geodesics with wave vector $\hat{k}^\mu_s$. The emission frequency is given by

$$\hat{\omega}_s := -\hat{g}_{\mu\nu}\hat{k}^\mu_s\hat{U}^\nu_s = -\tilde{\omega}\hat{g}_{\mu\nu}\hat{\ell}^\mu_s\hat{U}^\nu_s = \tilde{\omega}(1+\xi_s), \tag{5.17}$$

where we use the fact that locally at the source spacetime is approximately flat. So $\hat{\ell}^\mu_s \approx \hat{\bar{\ell}}^\mu_s = (1,\vec{n})$ where the bar here and thereafter will denote the background value of a given object. The observed frequency is similarly given by

$$\hat{\omega}_o = \tilde{\omega}\,\hat{\ell}^0_o\,(1+\xi_o). \tag{5.18}$$

Then the redshift is given by

$$1+\hat{z}=\frac{\hat{\omega}_s}{\hat{\omega}_o}=\frac{1+\xi_s}{\hat{\ell}^0_o(1+\xi_o)}=(\hat{\ell}^0_o)^{-1}(1+\xi_s-\xi_o). \tag{5.19}$$

In order to calculate this redshift, we need to relate the tangent vector of the light ray at the position of the observer $\hat{\ell}^\mu_o$ to the tangent vector at the source $\hat{\ell}^\mu_s \approx (1,\vec{n})$. This can be done via the geodesic equation

$$\frac{\mathrm{d}\hat{\ell}^\mu}{\mathrm{d}\hat{\lambda}}=-\hat{\Gamma}^\mu_{\rho\sigma}\hat{\ell}^\rho\hat{\ell}^\sigma, \tag{5.20}$$

which to first order becomes

$$\frac{\mathrm{d}\hat{\ell}^{(1)\mu}}{\mathrm{d}\hat{\lambda}}=-\hat{\Gamma}^\mu_{\rho\sigma}\hat{\bar{\ell}}^\rho\hat{\bar{\ell}}^\sigma, \tag{5.21}$$

where the background connection vanishes: $\hat{\bar{\Gamma}}^\mu_{\rho\sigma}=0$, because the background space is Minkowski. The Christoffel symbols can be easily calculated from the metric (5.15)

$$\hat{\Gamma}^0_{00}=\xi_{,\eta}; \tag{5.22}$$

$$\hat{\Gamma}^0_{0i}=\hat{\Gamma}^0_{i0}=\hat{\Gamma}^i_{00}=\xi_{,i}; \tag{5.23}$$

$$\hat{\Gamma}^0_{ij}=\hat{\Gamma}^k_{0i}=\hat{\Gamma}^k_{i0}=\hat{\Gamma}^k_{ij}=0; \tag{5.24}$$

Hence, the solution of the geodesic equation is given by

$$\hat{\ell}^{(1)0}_o-\hat{\ell}^{(1)0}_s=-\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\xi_{,\eta}+2\vec{\nabla}\xi.\vec{n})\mathrm{d}\hat{\lambda}; \tag{5.25}$$

$$\hat{\ell}^{(1)i}_o-\hat{\ell}^{(1)i}_s=-\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}\xi_{,i}\mathrm{d}\hat{\lambda}. \tag{5.26}$$

Equation (5.25), and the fact that $\hat{\ell}^\mu_s\approx(1,\vec{n})$, together imply that

$$\hat{\ell}^0_o=1-\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\xi_{,\eta}+2\vec{\nabla}\xi\cdot\vec{n})\mathrm{d}\hat{\lambda}. \tag{5.27}$$

Therefore, the redshift to first order becomes

$$1+\hat{z}=1-(\xi_o-\xi_s)+\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\xi_{,\eta}+2\vec{\nabla}\xi\cdot\vec{n})\mathrm{d}\hat{\lambda}. \tag{5.28}$$

We can put that in a more useful form by changing variables from the affine parameter $\hat{\lambda}$ to the conformal time η. Using (5.27) we have that to first order

$$\mathrm{d}\hat{\lambda}=\mathrm{d}\eta\left[1+\int_{\hat{\lambda}_s}^{\hat{\lambda}}(\xi_{,\eta}+2\vec{\nabla}\xi\cdot\hat{n})\mathrm{d}\hat{\lambda}'\right]. \tag{5.29}$$

We also use that to first order

$$\xi_{,\eta} = \frac{\mathrm{d}\xi}{\mathrm{d}\eta} - \vec{\nabla}\xi \cdot \vec{n}. \tag{5.30}$$

This then gives us the following expression for the redshift

$$1 + \hat{z} = 1 + \int_{\eta_s}^{\eta_o} \vec{\nabla}\xi \cdot \vec{n}\mathrm{d}\eta. \tag{5.31}$$

or, equivalently,

$$1 + \hat{z} = 1 + \xi_o - \xi_s - \int_{\eta_s}^{\eta_o} \xi_{,\eta}\mathrm{d}\eta. \tag{5.32}$$

Now we can find the redshift in the full perturbed FLRW metric

$$1 + z = \frac{a_o}{a_s}\left(1 + \Phi_o - \Phi_s + \int_{\eta_s}^{\eta_o} \vec{\nabla}\xi \cdot \vec{n}\mathrm{d}\eta\right), \tag{5.33}$$

or, equivalently,

$$1 + z = \frac{a_o}{a_s}\Big(1 + \Psi_o - \Psi_s - \int_{\eta_s}^{\eta_o} \xi_{,\eta}\mathrm{d}\eta\Big). \tag{5.34}$$

The redshift is a product of different contributions

$$1 + z = (1 + z_c)(1 + z_{gr})(1 + z_{ISW}). \tag{5.35}$$

Here

$$1 + z_c = \frac{a_o}{a_s}, \tag{5.36}$$

is the cosmological redshift due to the overall expansion of the universe, and

$$1 + z_{gr} = \sqrt{\frac{1 + 2\Psi_o}{1 + 2\Psi_s}} \approx 1 + \Psi_o - \Psi_s, \tag{5.37}$$

is the gravitational redshift due to the potential wells of the source and the observers. Finally,

$$1 + z_{ISW} = 1 - \int_{\eta_s}^{\eta_o} \xi_{,\eta}\mathrm{d}\eta \; = 1 - \int_{t_s}^{t_o} \xi_{,t} dt \tag{5.38}$$

is the gravitational redshift caused by changing potential wells along the path of the light — an integrated Sachs–Wolfe effect [86]. Eqn. (5.35) gives the total redshift without the Doppler redshift arising due to the peculiar velocities of the source and the observer. It is trivial to include the Doppler redshift in the analysis - (5.35) is modified to

$$1 + z = (1 + z_D)(1 + z_c)(1 + z_{gr})(1 + z_{ISW}), \tag{5.39}$$

where the Doppler contribution to the redshift is

$$1 + z_D = \frac{\gamma_s(1 - \vec{v}_s.\vec{n})}{\gamma_o(1 - \vec{v}_o.\vec{n})}, \tag{5.40}$$

and where $\gamma = (1 - |\vec{v}|^2)^{-\frac{1}{2}}$ and $\vec{v}_s$, $\vec{v}_o$ are the peculiar velocities of the source and the observer.

We can also adapt this redshift calculation to determine the affine parameter distance in terms of the difference in conformal time between the source and the observer. From the above, the relationship between affine parameter and conformal time is

$$\begin{aligned} \mathrm{d}\hat{\lambda} &= \mathrm{d}\eta \left[1 + \int_{\hat{\lambda}_s}^{\hat{\lambda}} (\xi_{,\eta} + 2\vec{\nabla}\xi \cdot \hat{n})\mathrm{d}\hat{\lambda}'\right] \\ &= \mathrm{d}\eta \left[1 + 2(\xi - \xi_s) - \int_{\eta_s}^{\eta} \xi_{,\eta'}\mathrm{d}\eta'\right] \\ &= \mathrm{d}\eta \left[1 + 2(\xi - \xi_s) + z_{ISW}(\eta_s) - z_{ISW}(\eta)\right], \end{aligned} \tag{5.41}$$

where

$$z_{ISW}(\eta) = -\int_{\eta}^{\eta_o} \xi_{,\eta'}\mathrm{d}\eta'. \tag{5.42}$$

Integrating the above expression gives

$$\hat{\lambda}_o - \hat{\lambda}_s = (\eta_o - \eta_s)\ \left[1 + 2(\langle\xi\rangle - \xi_s) + z_{ISW} - \langle z_{ISW}\rangle\right]. \tag{5.43}$$

Here $\langle\xi\rangle$ and $\langle z_{ISW}\rangle$ are simply averages along the line of sight:

$$\langle\xi\rangle := \frac{1}{\eta_o - \eta_s}\int_{\eta_s}^{\eta_o} \xi\mathrm{d}\eta; \tag{5.44}$$

$$\langle z_{ISW}\rangle := \frac{1}{\eta_o - \eta_s}\int_{\eta_s}^{\eta_o} z_{ISW}(\eta)\mathrm{d}\eta. \tag{5.45}$$

While $\langle\xi\rangle$ and $\langle z_{ISW}\rangle$, (and ξ_s and z_{ISW} for that matter), might be difficult to measure, they do at least have clear physical interpretations.

The Jacobi and van Vleck determinants The Jacobi map and Jacobi determinant can be calculated using the formalism developed in the previous chapter. We present here the final result for the Jacobi determinant and defer the full calculation to Appendix E. The Jacobi determinant in the unphysical metric (5.15) is given by:

$$(\det \hat{J})^{\frac{1}{2}} = (\hat{\lambda}_o - \hat{\lambda}_s)\left\{1 - \frac{1}{2}\frac{1}{\hat{\lambda}_o - \hat{\lambda}_s}\int_{\hat{\lambda}_s}^{\hat{\lambda}_o} (\hat{\lambda}_o - \hat{\lambda})(\nabla^2\xi - n^i n^j \xi_{,ij})(\hat{\lambda} - \hat{\lambda}_s)\mathrm{d}\hat{\lambda}\right\}. \tag{5.46}$$

Since the Jacobi and the van Vleck approaches are equivalent as demonstrated in the previous chapter, we must necessarily have that

$$(\det \hat{J})^{\frac{1}{2}} = \hat{\Delta}_{vV}^{-\frac{1}{2}}\ (\hat{\lambda}_o - \hat{\lambda}_s). \tag{5.47}$$

We will now show that this is indeed the case by a direct calculation of the van Vleck determinant.

In the weak field limit the van Vleck determinant is approximated by [71–73]

$$\begin{aligned}\hat{\Delta}_{vV} &\approx \exp\left[\frac{1}{\hat{\lambda}_o-\hat{\lambda}_s}\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\hat{\lambda}_o-\hat{\lambda})(\hat{R}_{\mu\nu}\hat{\ell}^\mu\hat{\ell}^\nu)(\hat{\lambda}-\hat{\lambda}_s)\mathrm{d}\hat{\lambda}\right]\\ &\approx 1+\left[\frac{1}{\hat{\lambda}_o-\hat{\lambda}_s}\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\hat{\lambda}_o-\hat{\lambda})(\hat{R}_{\mu\nu}\hat{\ell}^\mu\hat{\ell}^\nu)(\hat{\lambda}-\hat{\lambda}_s)\mathrm{d}\hat{\lambda}\right].\end{aligned} \tag{5.48}$$

The components of the Ricci tensor to first order are

$$\hat{R}_{00}=\nabla^2\xi; \qquad \hat{R}_{0i}=0; \qquad \hat{R}_{ij}=-\xi_{,ij}. \tag{5.49}$$

Since $\hat{\bar{R}}_{\mu\nu}=0$, only the term $\hat{R}^{(1)}_{\mu\nu}\hat{\bar{\ell}}^\mu\hat{\bar{\ell}}^\nu$ will contribute to first order in the expression (5.48). We have

$$\hat{R}^{(1)}_{\mu\nu}\,\hat{\bar{\ell}}^\mu\hat{\bar{\ell}}^\nu=(\nabla^2\xi-n^in^j\xi_{,ij}), \tag{5.50}$$

and therefore

$$\hat{\Delta}_{vV}=1+\frac{1}{\hat{\lambda}_o-\hat{\lambda}_s}\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\hat{\lambda}_o-\hat{\lambda})(\nabla^2\xi-n^in^j\xi_{,ij})(\hat{\lambda}-\hat{\lambda}_s)\mathrm{d}\hat{\lambda}. \tag{5.51}$$

Hence,

$$\hat{\Delta}_{vV}^{-\frac{1}{2}}=1-\frac{1}{\hat{\lambda}_o-\hat{\lambda}_s}\frac{1}{2}\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\hat{\lambda}_o-\hat{\lambda})(\nabla^2\xi-n^in^j\xi_{,ij})(\hat{\lambda}-\hat{\lambda}_s)\mathrm{d}\hat{\lambda}. \tag{5.52}$$

We see that (5.47) is satisfied and so the two approaches are equivalent as expected.

The luminosity distance in perturbed FLRW Now we finish the calculation of the luminosity distance in perturbed FLRW. We can express the Jabobi determinant (5.46) in terms of the conformal time by using the fact that

$$\frac{\mathrm{d}\eta}{\mathrm{d}\hat{\lambda}}=1-\int_{\hat{\lambda}_s}^{\hat{\lambda}}(\xi_{,\eta}+2\vec{\nabla}\xi\cdot\vec{n})\mathrm{d}\hat{\lambda}', \tag{5.53}$$

and hence to linear order

$$\mathrm{d}\hat{\lambda}=\mathrm{d}\eta\left(1+\int_{\eta_s}^{\eta}(\xi_{,\eta}+2\vec{\nabla}\xi\cdot\vec{n})\mathrm{d}\eta'\right). \tag{5.54}$$

The resulting expression for the Jacobi determinant is:

$$\begin{aligned}(\det\hat{J})^{\frac{1}{2}} &= (\eta_o-\eta_s)+\int_{\eta_s}^{\eta_o}\xi\mathrm{d}\eta+\int_{\eta_s}^{\eta_o}(\eta_o-\eta)(\vec{\nabla}\xi\cdot\vec{n})\mathrm{d}\eta-\xi_s(\eta_o-\eta_s)\\ &\quad-\frac{1}{2}\int_{\eta_s}^{\eta_o}(\eta_o-\eta)(\nabla^2\xi-n^in^j\xi_{,ij})(\eta-\eta_s)\mathrm{d}\eta,\end{aligned} \tag{5.55}$$

where again we have replaced a double integral by a single integral. Hence, the luminosity distance in two-mode perturbed (Φ, ξ) FLRW cosmology is given by

$$d_L(\eta_s,\eta_o,\vec{n}) = \frac{f_o^2}{f_s}\,\hat{d}_L \tag{5.56}$$

$$= \frac{f_o^2}{f_s}\,(\det\hat{J})^{\frac{1}{2}}\,(1+\hat{z}) \tag{5.57}$$

$$\begin{aligned} = \frac{a_o^2}{a_s}\,\Big[&(\eta_o-\eta_s)+2\Phi_o(\eta_o-\eta_s)-\Psi_s(\eta_o-\eta_s)\\ &+(\eta_o-\eta_s)\int_{\eta_s}^{\eta_o}\vec{\nabla}\xi\cdot\vec{n}\mathrm{d}\eta+\int_{\eta_s}^{\eta_o}\xi\mathrm{d}\eta+\int_{\eta_s}^{\eta_o}(\eta_o-\eta)(\vec{\nabla}\xi\cdot\vec{n})\mathrm{d}\eta\\ &-\frac{1}{2}\int_{\eta_s}^{\eta_o}(\eta_o-\eta)(\nabla^2\xi-n^in^j\xi_{,ij})(\eta-\eta_s)\mathrm{d}\eta\Big]. \end{aligned} \tag{5.58}$$

This formula shows the dependance of the luminosity distance measured by an observer O as a function of the conformal time of the source η_s in a given direction $\vec{n}$. In the next section we will present a simple toy model where we will use the above formula. For now will try to re-cast it to the extent possible in terms of various contributions to the redshift. For instance, by recognising that $d_{L,FLRW}=\frac{a_o^2}{a_s}(\eta_o-\eta_s)$ is the luminosity distance in FLRW without peculiar velocities, one can write

$$\begin{aligned} d_L(\eta_s,\eta_o,\vec{n}) \;=\;& d_{L,FLRW}(z_c)\Big[1+2\Phi_o-\Psi_s+\int_{\eta_s}^{\eta_o}\vec{\nabla}\xi\cdot\vec{n}\mathrm{d}\eta\\ &+\frac{1}{\eta_o-\eta_s}\int_{\eta_s}^{\eta_o}\xi\mathrm{d}\eta+\frac{1}{\eta_o-\eta_s}\int_{\eta_s}^{\eta_o}(\eta_o-\eta)(\vec{\nabla}\xi\cdot\vec{n})\mathrm{d}\eta\\ &-\frac{1}{2}\frac{1}{\eta_o-\eta_s}\int_{\eta_s}^{\eta_o}(\eta_o-\eta)(\nabla^2\xi-n^in^j\xi_{,ij})(\eta-\eta_s)\mathrm{d}\eta\Big]. \end{aligned} \tag{5.59}$$

There are several other ways of usefully repackaging the luminosity distance in the two-mode perturbed (Φ, ξ) FLRW cosmology we are considering. For instance, using (5.43), we have that

$$(\det\hat{J})^{\frac{1}{2}} = (\eta_o-\eta_s)\;\left[1+2(\langle\xi\rangle-\xi_s)+z_{ISW}-\langle z_{ISW}\rangle\right]\hat{\Delta}_{vV}^{-\frac{1}{2}}, \tag{5.60}$$

and substituting that inside (5.57) we obtain

$$\begin{aligned} d_L \;=\;& d_{L,FLRW}(z_c)\,(1+\Phi_o)\,(1+z_{gr})\,(1+z_{ISW})\,[1+2(\langle\xi\rangle-\xi_s)+z_{ISW}-\langle z_{ISW}\rangle]\\ &\times\left\{1-\frac{1}{2}\frac{1}{\eta_o-\eta_s}\int_{\eta_s}^{\eta_o}(\eta_o-\eta)(\nabla^2\xi-n^in^j\xi_{,ij})(\eta-\eta_s)\mathrm{d}\eta\right\}. \end{aligned} \tag{5.61}$$

The $(1+\Phi_o)$ factor is relatively uninteresting, since it only depends on what is happening at the observer, it is common to all observations — at worst it is a rescaling to marginalise over. These various ways of looking at the luminosity distance, we do feel, give us a somewhat better handle on the fundamental physics. Equations (5.59) and (5.61) are now manifestly of the form

$$d_L(\eta_s,\eta_o,\vec{n}) = d_{L,FLRW}(z_c)\times\{1+(\text{perturbatively small})\}\,. \tag{5.62}$$

5.4 Generalised cosmography

We now consider a simple toy model where the Bardeen potentials depend sinusoidally on conformal time and are independent of space

$$\Psi = -\Phi = \epsilon\, \sin(\kappa\eta), \tag{5.63}$$

where ϵ and κ are constants and ϵ is perturbatively small. Initially we shall neglect peculiar velocities, but subsequently show how to put them back in. We choose this particular toy model because it is tractable, and because, even though the Bardeen potentials in this case are independent of space, it serves to illustrate the basic principles behind generalising the cosmographic approach to an inhomogeneous universe. In order to analyse the real universe, one would need to consider more sophisticated models in which the Bardeen potentials depend also on space, and which would therefore be more computationally cumbersome.

Toy model without peculiar velocities Equations (5.58) and (5.33) become

$$d_L = \frac{a_o^2}{a_s}\Big[\Delta\eta + \epsilon\Big(-2\sin(\kappa\eta_o)\Delta\eta - \sin(\kappa\eta_s)\Delta\eta - 2\frac{\cos(\kappa\eta_o)}{\kappa} + 2\frac{\cos(\kappa\eta_s)}{\kappa}\Big)\Big]; \tag{5.64}$$

and

$$1 + z_s = \frac{a_o}{a_s}\Big[1 + \epsilon\Big(-\sin(\kappa\eta_o) + \sin(\kappa\eta_s)\Big)\Big]. \tag{5.65}$$

Now we derive a cosmographic series for d_L in terms of z. The coefficients to leading order are expected to be the same as in (4.87) plus corrections of order ϵ. The cosmographic parameters are defined in the same way as before – equations (4.82), and we make use of the following relation, valid for any conformal time η,

$$1 + z(\eta) = \frac{a_o}{a(\eta)}\Big[1 + \epsilon\Big(-\sin(\kappa\eta_o) + \sin(\kappa\eta)\Big)\Big]. \tag{5.66}$$

Expanding $a(\eta)$ and $\sin(\eta)$ as a series in terms of $(\eta - \eta_o)$ inside (5.66), we obtain a series for $z(\eta)$ in terms of $(\eta - \eta_o)$

$$\begin{aligned} z(\eta) = &\Big[-\mathcal{H}_o + \epsilon\Big(\kappa\cos(\kappa\eta_o)\Big)\Big](\eta - \eta_o) \\ &+ \Big[\mathcal{H}_o^2\Big(\frac{2 + \mathcal{Q}_o}{2}\Big) + \epsilon\Big(-\kappa\cos(\kappa\eta_o)\mathcal{H}_o - \kappa^2\frac{\sin(\kappa\eta_o)}{2}\Big)\Big](\eta - \eta_o)^2 \\ &+ \Big[-\mathcal{H}_o^3\Big(\frac{\mathcal{J}_o + 6\mathcal{Q}_o + 6}{6}\Big) + \epsilon\Big(\kappa\cos(\kappa\eta_o)\mathcal{H}_o^2(\frac{2 + \mathcal{Q}_o}{2}) \\ &+ \kappa^2\frac{\sin(\kappa\eta_o)}{2}\mathcal{H}_o - \kappa^3\frac{\cos(\kappa\eta_o)}{6}\Big)\Big](\eta - \eta_o)^3 + O(\eta - \eta_o)^4. \end{aligned} \tag{5.67}$$

Reverting this series, we find

$$\eta - \eta_o = A_1\, z + A_2\, z^2 + A_3\, z^3 + O(z^4), \tag{5.68}$$

where

$$A_1 = -\frac{1}{\mathcal{H}_o} + \epsilon\Big[-\kappa\frac{\cos(\kappa\eta_o)}{\mathcal{H}_o^2}\Big]; \tag{5.69}$$

$$\begin{aligned} A_2 = \frac{1}{\mathcal{H}_o}\Big(\frac{2+\mathcal{Q}_o}{2}\Big) \\ + \epsilon\Big[-\kappa\frac{\cos(\kappa\eta_o)}{\mathcal{H}_o^2} - \kappa^2\frac{\sin(\eta_o)}{2\mathcal{H}_o^3} + \kappa\frac{3\cos(\eta_o)}{\mathcal{H}_o^2}\Big(\frac{2+\mathcal{Q}_o}{2}\Big)\Big]; \end{aligned} \tag{5.70}$$

$$\begin{aligned} A_3 = -\frac{1}{\mathcal{H}_o}\left(\frac{6+3\mathcal{Q}_o^2+6\mathcal{Q}_o-\mathcal{J}_o}{6}\right) \\ + \epsilon\frac{1}{\mathcal{H}_o^2}\Big[\kappa\cos(\kappa\eta_o)\Big(\frac{-6-9\mathcal{Q}_o+5\mathcal{Q}_o^2-2\mathcal{J}_o^2}{2}\Big) \\ + \kappa^3\frac{\cos(\kappa\eta_o)}{6\mathcal{H}_o^2} + \kappa\frac{\sin(\kappa\eta_o)}{\mathcal{H}_o}\left(\frac{3+2\mathcal{Q}_o}{2}\right)\Big]. \end{aligned} \tag{5.71}$$

We also have

$$\Delta\eta = \eta_o - \eta_s = -A_1 z_s - A_2 z_s^2 - A_3 z_s^3 + O(z_s^4). \tag{5.72}$$

This allows us to expand $\sin(\eta_s)$, $\cos(\eta_s)$ and $\frac{a_o}{a_s}$ as functions of z_s. We find

$$\begin{aligned} \sin(\kappa\eta_s) = \sin(\kappa\eta_o) + \Big[\kappa\cos(\kappa\eta_o)A_1\Big] z_s \\ + \left[\kappa\cos(\kappa\eta_o)A_2 - \kappa^2\frac{\sin(\kappa\eta_o)A_1^2}{2}\right] z_s^2 \\ + \left[\kappa\cos(\kappa\eta_o)A_3 - \kappa^2\sin(\kappa\eta_o)A_1A_2 - \kappa^3\frac{\cos(\kappa\eta_o)}{6}A_1^3\right] z_s^3 + O(z_s^4); \end{aligned} \tag{5.73}$$

while

$$\begin{aligned} \cos(\kappa\eta_s) = \cos(\kappa\eta_o) + \Big[-\kappa\sin(\kappa\eta_o)A_1\Big] z_s \\ + \left[-\kappa\sin(\kappa\eta_o)A_2 - \frac{\kappa^2\cos(\kappa\eta_o)A_1^2}{2}\right] z_s^2 \\ + \left[-\kappa\sin(\kappa\eta_o)A_3 - \kappa^2\cos(\kappa\eta_o)A_1A_2 + \kappa^3\frac{\sin(\kappa\eta_o)}{6}A_1^3\right] z_s^3 + O(z_s^4); \end{aligned} \tag{5.74}$$

and

$$\begin{aligned}\frac{a_o}{a_s} = 1 &+ \left[1 - \epsilon\kappa\cos(\kappa\eta_o)A_1\right] z_s \\ &+ \epsilon\left[-\kappa\cos(\kappa\eta_o)A_2 + \kappa^2\frac{\sin(\kappa\eta_o)A_1^2}{2} - \kappa\cos(\kappa\eta_o)A_1\right] z_s^2 \\ &- \epsilon\left[\kappa\cos(\kappa\eta_o)A_3 - \kappa^2\sin(\kappa\eta_o)A_1A_2 - \kappa^3\frac{\cos(\kappa\eta_o)}{6}A_1^3\right. \\ &\left. + \kappa\cos(\kappa\eta_o)A_2 - \kappa^2\frac{\sin(\kappa\eta_o)A_1^2}{2}\right] z_s^3 + O(z_s^4). \end{aligned} \tag{5.75}$$

Substituting everything inside equation (5.64) we obtain an expansion of the luminosity distance d_L in terms of the redshift z

$$\begin{aligned}\frac{d_L}{a_o} = &\left[\frac{1}{\mathcal{H}_o} + \epsilon\mathcal{X}\right] z_s + \left[-\frac{\mathcal{Q}_o}{2\mathcal{H}_o} + \epsilon\mathcal{Y}\right] z_s^2 \\ &+ \left[\frac{1}{\mathcal{H}_o}\left(\frac{3\mathcal{Q}_o^2 + 3\mathcal{Q}_o - \mathcal{J}_o}{6}\right) + \epsilon\mathcal{Z}\right] z_s^3 + O(z_s^4). \end{aligned} \tag{5.76}$$

Here

$$\mathcal{X} := 2\frac{\sin(\kappa\eta_o)}{\kappa} - 2\frac{\cos(\kappa\eta_o)}{\kappa} - \frac{\sin(\kappa\eta_o)}{\mathcal{H}_o} + \kappa\frac{\cos(\kappa\eta_o)}{\mathcal{H}_o^2}; \tag{5.77}$$

$$\mathcal{Y} := \kappa\frac{6\cos(\kappa\eta_o)}{\mathcal{H}_o^2} + \kappa\frac{3\cos(\kappa\eta_o)\mathcal{Q}_o}{2\mathcal{H}_o^2} + \frac{\sin(\kappa\eta_o)\mathcal{Q}_o}{2\mathcal{H}_o} + \kappa^2\frac{\sin(\kappa\eta_o)}{2\mathcal{H}_o^3}; \tag{5.78}$$

and

$$\begin{aligned}\mathcal{Z} := &\kappa\frac{2\cos(\kappa\eta_o)}{\mathcal{H}_o^2} - \kappa^2\frac{2\cos(\kappa\eta_o)}{\mathcal{H}_o^3} - \kappa^2\frac{\cos(\kappa\eta_o)\mathcal{Q}_o}{\mathcal{H}_o^3} - \kappa^3\frac{\cos(\kappa\eta_o)}{6\mathcal{H}_o^4} \\ &+ \kappa\frac{\cos(\kappa\eta_o)}{\mathcal{H}_o^2}\left(\frac{36 + 15\mathcal{Q}_o^2 + 36\mathcal{Q}_o - 4\mathcal{J}_o}{6}\right) - \frac{4\sin(\kappa\eta_o)}{\mathcal{H}_o} - \kappa^2\frac{\sin(\kappa\eta_o)}{3\mathcal{H}_o^3} \\ &- \frac{\sin(\kappa\eta_o)\mathcal{Q}_o}{2\mathcal{H}_o} - \kappa^2\frac{\sin(\kappa\eta_o)\mathcal{Q}_o}{\mathcal{H}_o^3} + \frac{2\sin(\kappa\eta_o)}{\mathcal{H}_o}\left(\frac{6 + 3\mathcal{Q}_o^2 + 6\mathcal{Q}_o - \mathcal{J}_o}{6}\right). \end{aligned} \tag{5.79}$$

This agrees to zeroth order in ϵ with equation (4.87).

From the above we see that in our toy model (a sinusoidally perturbed $k = 0$ FLRW universe) without peculiar velocities we have

$$d_L = \frac{a_o}{\mathcal{H}_o}(1 + z)P(z), \tag{5.80}$$

where $P(z)$ is the specific polynomial

$$P(z) = B_1\, z + B_2\, z^2 + B_3\, z^3 + O(z^4) \tag{5.81}$$

with

$$B_1 = 1 + \epsilon\mathcal{H}_o\mathcal{X}; \tag{5.82}$$

$$B_2 = -\frac{\mathcal{Q}_o + 2}{2} + \epsilon\mathcal{H}_o(\mathcal{Y} - \mathcal{X}); \tag{5.83}$$

$$B_3 = \frac{3\mathcal{Q}_o^2 + 6\mathcal{Q}_o - \mathcal{J}_o + 6}{6} + \epsilon\mathcal{H}_o(\mathcal{Z} - \mathcal{Y} + \mathcal{X}). \tag{5.84}$$

The only thing that has changed with respect to standard FLRW is the coefficients of the polynomial. These coefficients are still constant because the toy model is isotropic. In a more sophisticated anisotropic model the coefficients would be functions of the angular coordinates on the sky.

Toy model with peculiar velocities If one now adds peculiar velocities, then, as discussed several times already, the only change is that

$$d_L = \frac{a_o}{\mathcal{H}_o}(1+z)P(z_c), \tag{5.85}$$

where z_c is the cosmological contribution to the total redshift z. Now in terms of the redshift contributions due to peculiar velocity z_p, we again have

$$z_c = \frac{1+z}{1+z_D} - 1 \approx z - (1+z)z_D + O(z_D^2), \tag{5.86}$$

again implying

$$d_L = \frac{a_o}{\mathcal{H}_o}\left\{(1+z)P(z) - (1+z)^2P'(z)z_D + O(z_D^2)\right\}. \tag{5.87}$$

Within the context of this model universe, this gives an explicit formula for estimating the potential effect of peculiar velocities on the luminosity distance. Again evaluating explicitly the polynomial $P(z)$ to $O(z^3)$ allows us to express d_L to $O(z^2)$ and $O(z_D)$

$$\begin{aligned} d_L = \frac{a_o}{\mathcal{H}_o}\Bigg\{ & -\left(1 + \epsilon\mathcal{H}_o\mathcal{X}\right)z_D \\ & + \left[1 + \epsilon\mathcal{H}_o\mathcal{X} - \left(-\mathcal{Q}_o + \epsilon 2\mathcal{H}_o\mathcal{Y}\right)z_D\right]z \\ & + \left[-\frac{\mathcal{Q}_o}{2} + \epsilon\mathcal{H}_o\mathcal{Y} - \left(\frac{3\mathcal{Q}_o^2 + 2\mathcal{Q}_o - \mathcal{J}_o}{2} + \epsilon\mathcal{H}_o(\mathcal{Y} + 3\mathcal{Z})\right)z_D\right]z^2 \\ & + O(z^3) + O(z_D^2)\Bigg\}. \end{aligned} \tag{5.88}$$

We could proceed further for instance by assuming $\langle z_D\rangle = 0$, (effectively temporarily ignoring peculiar Doppler shifts), and fitting

$$\langle d_L\rangle = \frac{a_o}{\mathcal{H}_o}\left\{\left[1 + \epsilon\mathcal{H}_o\mathcal{X}\right]z + \left[-\frac{\mathcal{Q}_o}{2} + \epsilon\mathcal{H}_o\mathcal{Y}\right]z^2 + O(z^3)\right\}. \tag{5.89}$$

Then

$$d_L - \langle d_L \rangle = -\frac{a_o}{\mathcal{H}_o} z_D \Bigg\{ \Big(1 + \epsilon \mathcal{H}_o \mathcal{X} \Big) + \Big(-\mathcal{Q}_o + \epsilon 2 \mathcal{H}_o \mathcal{Y} \Big) z + \Big(\frac{3\mathcal{Q}_o^2 + 2\mathcal{Q}_o - \mathcal{J}_o}{2} + \epsilon \mathcal{H}_o (\mathcal{Y} + 3\mathcal{Z}) \Big) z^2 \Bigg\} + O(z^3) + O(z_D^2). \quad (5.90)$$

So even in this sinusoidally perturbed FLRW model we see how we can use cosmographic techniques to estimate the size of the peculiar Doppler shifts.

Summary In this chapter we considered two possible deviations from the assumption that light propagates in a perfectly homogenous and isotropic Universe. First, we considered CFLRW cosmological models which represent a genuine modification of the ΛCDM model, where the FLRW background is deformed by an arbitrary large conformal factor consistent with CMB physics. Second, we considered linearly perturbed FLRW models with 2 scalar mode perturbations (the Bardeen potentials), which can be applied in two contexts - first, to investigate the effect of inhomogeneities from cosmic structure on the propagation of light and therefore on the inference of the cosmographic parameters, and second, to modifications of the ΛCDM model where dark energy is no longer a cosmological constant, but rather a dynamic field on space and time. While we considered no formal model of dark energy, we demonstrated how the generalised cosmographic procedure works for a simple toy model where a single scalar mode varies sinusoidally with time. The same approach can be applied to more sophisticated and realistic cosmological models where of course the calculations would be much more convoluted.

6

Testing the isotropy of the Universe

There are several claims in the literature of some evidence for an anisotropic expansion of the present universe [26–28]. Such an anisotropy is easiest to associate with the properties of dark energy, which might be something more complicated than a cosmological constant, and hence might be inhomogeneous at later times leading to an anisotropic expansion. On the other hand, there are other studies which claim no deviation from the null hypothesis of a statistically isotropic universe [87, 88], and indeed CMB data gives strong support to the idea that the early universe is isotropic [89]. We shall consider here how to test for a possible anisotropy in the late universe. In particular, we shall construct a simple toy model where dark energy undergoes a phase transition at DM-DE equality and leads to isotropy-breaking. We want to contrast this model with ΛCDM and see whether there is any signal in the data for such a late-time deviation from isotropy. Unfortunately, the current data from Supernovae Type Ia is not enough to distinguish between them. We can however make forecasts with mock datasets and show that if some conditions turn out to be satisfied, future data might be able to differentiate between the two models and hence settle the issue about the isotropy of the universe.

6.1 A simple toy model of an anisotropic expansion

We consider a simple toy model where the universe expands at a different rate along one of the directions, which we label as the z-direction (this can be generalised to an arbitrary direction). We capture this by introducing in the spatially flat FLRW metric a time-dependent function $\epsilon(t)$ in front of the other two directions:

$$\begin{aligned} ds^2 &= -dt^2 + a^2(t)[\epsilon^2(t)(dx^2 + dy^2) + dz^2] \\ &= a^2(\eta)[-d\eta^2 + \epsilon^2(\eta)(dx^2 + dy^2) + dz^2] \\ &= a^2(\eta)d\tilde{s}^2, \end{aligned} \tag{6.1}$$

where

$$\begin{aligned} d\tilde{s}^2 &= -d\eta^2 + \epsilon^2(\eta)(dx^2 + dy^2) + dz^2 \\ &= \epsilon^2(\tau)[-d\tau^2 + dx^2 + dy^2] + dz^2 \\ &= \epsilon^2(\tau)d\hat{s}^2 + dz^2, \end{aligned} \tag{6.2}$$

and

$$\begin{aligned} d\hat{s}^2 &= -d\tau^2 + dx^2 + dy^2 \\ &= -d\tau^2 + dr^2. \end{aligned} \tag{6.3}$$

Thus we have three metrics ds^2, $d\tilde{s}^2$ and $d\hat{s}^2$, related by conformal transformations. Under a conformal transformation $ds^2 = f^2(x)ds'^2$, the luminosity distance transforms as in Eq. (4.61):

$$d_L = \frac{f_o^2}{f_s} d_L^{'}, \tag{6.4}$$

where f_0 is the value of the conformal factor at the observer, and f_s is the value of the conformal factor at the source.

The luminosity distance in $d+1$ dimensions is related to the luminosity distance in d dimensions by

$$d_L^{(d+1)} = \sqrt{d_L^{(d)2} + z^2}, \tag{6.5}$$

because there is no coefficient in front of dz^2 and hence as we go in the z-direction, we simply increment the Euclidean distance. Therefore, by using (6.4) and (6.24), we can derive the following expressions for the luminosity distance in each of the three metrics:

$$\hat{d}_L = r, \tag{6.6}$$

$$\tilde{d}_L = \sqrt{\frac{\epsilon_o^4}{\epsilon_s^2} r^2 + z^2}, \tag{6.7}$$

$$d_L = \frac{a_o^2}{a_s} \sqrt{\frac{\epsilon_o^4}{\epsilon_s^2} r^2 + z^2}, \tag{6.8}$$

where r is the radial distance in the xy-plane. If we set $z = R\cos\theta$ and $r = R\sin\theta$, Eq. (6.8) simplifies to

$$\begin{aligned} d_L &= \frac{a_o^2}{a_s} R \sqrt{\frac{\epsilon_o^4}{\epsilon_s^2} \sin^2\theta + \cos^2\theta} \\ &= d_{L,FRW} \sqrt{\frac{\epsilon_o^4}{\epsilon_s^2} \sin^2\theta + \cos^2\theta}, \end{aligned} \tag{6.9}$$

where $d_{L,FRW} = \frac{a_o^2}{a_s} R$ is the luminosity distance in a FLRW spacetime.

The function $\epsilon(t)$ should capture the behaviour that we described before – at early times the universe is homogeneous and isotropic and therefore $\epsilon \to 1$, while at late times the isotropy is broken and ϵ settles to some value different from unity. Since this isotropy-breaking is associated with dark energy, the transition occurs roughly at the time of dark matter – dark energy equality t_{eq} and has a typical time-scale equal to the time scale for the transition from matter domination to dark energy domination T_{eq}. Thus $\epsilon(t)$ is assumed to have the following form:

$$\epsilon(t) = -\Big(\frac{1}{2} - \frac{\epsilon_0}{2}\Big)\tanh\Big(\frac{t - t_{eq}}{T_{eq}}\Big) + \frac{1}{2} + \frac{\epsilon_0}{2}, \tag{6.10}$$

where $0 \leq \epsilon_0 \leq 1$ is a free parameter which gives the amount of isotropy breaking. Fig. 6.1 shows the evolution of $\epsilon(t)$ in terms of logarithmic time for $\epsilon_0 = 0.84$ (we will comment on the value chosen later). The constant ϵ_0 represents the maximal amount of isotropy breaking – it is equal to the value of ϵ at future timelike infinity. The constant T_{eq} is specified by the physics at matter-dark energy equality, as we will see later.

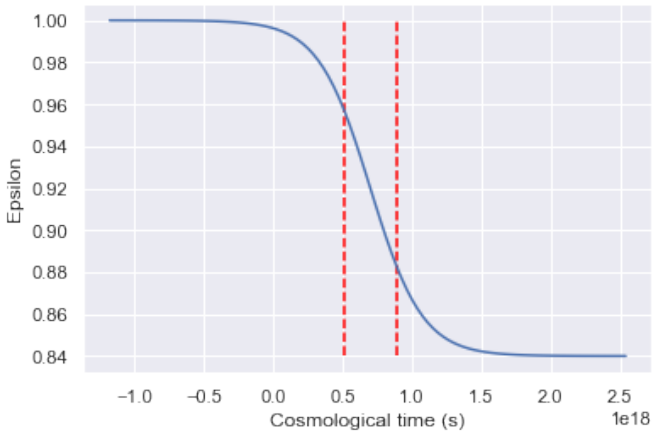

Figure 6.1: The function $\epsilon(t)$ with $\epsilon_0 = 0.84$. The time interval marked by the red dashed lines corresponds to T_{eq}.

6.2 A two-component universe

We want to write (6.9) as a function solely of the redshift z and the polar angle θ (it does not depend on the azimuthal angle ϕ because the isotropy is broken along a single dimension which we have identified with the z-direction). For that we need to write (6.10) as a function of z. For that we need to solve for the evolution of the scale factor $a(t)$ during the time when both matter and dark energy are important, i.e. we need to solve the Friedmann equations with both matter and dark energy sources (we

set $k = 0$)[1]:

$$\Big(\frac{\dot{a}}{a}\Big)^2 = \frac{8\pi G}{3}\rho, \tag{6.11}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P), \tag{6.12}$$

where $\rho = \rho_m + \rho_\Lambda$ with

$$\rho_m = \frac{1}{2}\rho_{eq}\Big(\frac{a_{eq}}{a}\Big)^3, \qquad \rho_\Lambda = \frac{1}{2}\rho_{eq}, \tag{6.13}$$

where ρ_{eq} is the total density at the time of the equality and a_{eq} is the scale factor at equality. While dark energy is not really a cosmological constant in the anisotropic model, it will be approximately so, and therefore it is fine to treat it as a constant as a first-order approximation. Since what we want to do is to solve for $t(z)$ in order to substitute it in Eq. (6.10), any deviation from FLRW and a cosmological constant would appear as second-order in ϵ. Thus we are allowed to estimate $t(z)$ by assuming that the metric is FLRW. Using (6.13), the first Friedmann equation (the constraint equation) becomes

$$\dot{a}^2 = \frac{4\pi G}{3}\rho_{eq}a_{eq}^3\frac{1}{a} + \frac{4\pi G}{3}\rho_{eq}a^2, \tag{6.14}$$

while the second Friedmann equation (the dynamical equation) is not independent anymore - it can be derived from the first by differentiating with respect to t.

The solution of (6.14) with the initial condition $a(t = 0) = 0$ is

$$a = a_{eq}\Big[\sinh\Big(\frac{3}{2}\sqrt{\frac{4\pi G}{3}\rho_{eq}}t\Big)\Big]^{2/3}, \tag{6.15}$$

which can be easily verified by differentiating with respect to time. Inverting that, we get

$$\begin{aligned} t(a) &= \frac{2}{3}\Big(\frac{4\pi G}{3}\rho_{eq}\Big)^{-1/2}\operatorname{arcsinh}\Big[\Big(\frac{a}{a_{eq}}\Big)^{3/2}\Big] \\ &= \frac{2}{3}\Omega_{\Lambda,0}^{-1/2}H_0^{-1}\ln\Big[\Big(\frac{a}{a_{eq}}\Big)^{3/2} + \sqrt{1 + \Big(\frac{a}{a_{eq}}\Big)^3}\Big], \end{aligned} \tag{6.16}$$

where to get to the second line we used that $\rho_{\Lambda,0} = \rho_{\Lambda,eq} = \frac{1}{2}\rho_{eq}$ and that $\operatorname{arcsinh}(x) = \ln(x + \sqrt{x^2 + 1})$. Using $a = (1 + z)^{-1}$ (we set $a_0 = 1$) this can be rewritten as

$$\begin{aligned} t(z) &= \frac{2}{3}\Big(\frac{4\pi G}{3}\rho_{eq}\Big)^{-1/2}\operatorname{arcsinh}\Big[\Big(\frac{1 + z_{eq}}{1 + z}\Big)^{3/2}\Big] \\ &= \frac{2}{3}\Omega_{\Lambda,0}^{-1/2}H_0^{-1}\ln\Big[\Big(\frac{1 + z_{eq}}{1 + z}\Big)^{3/2} + \sqrt{1 + \Big(\frac{1 + z_{eq}}{1 + z}\Big)^3}\Big]. \end{aligned} \tag{6.17}$$

[1]Note that the Friedmann equations would still hold approximately if we assume that the effect of ϵ is very small, which will certainly be true at late times.

The time of equality is given by

$$t_{eq} = \frac{2}{3}\Omega_{\Lambda,0}^{-1/2} H_0^{-1} \text{arcsinh}[1] \approx 7 \times 10^{17} s. \tag{6.18}$$

In addition, we can easily determine the redshift at equality z_{eq} and the scale factor at equality a_{eq}. Defining $\Omega_i \equiv \frac{\rho_i}{\rho_{cr}}$, we have

$$\Omega_m = \frac{\Omega_{m,0}(1+z)^3}{\Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}}, \tag{6.19}$$

$$\Omega_\Lambda = \frac{\Omega_{\Lambda,0}}{\Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}}. \tag{6.20}$$

Imposing $\Omega_{m,eq} = \Omega_{\Lambda,eq}$ leads to

$$z_{eq} = \left(\frac{\Omega_{\Lambda,0}}{\Omega_{m,0}}\right)^{1/3} - 1 \approx 0.3, \tag{6.21}$$

and therefore

$$a_{eq} = (1+z_{eq})^{-1} = \left(\frac{\Omega_{m,0}}{\Omega_{\Lambda,0}}\right)^{1/3}. \tag{6.22}$$

The time scale for the dark matter – dark energy transition T_{eq} is given by the inverse of the Hubble scale at the time of equality

$$T_{eq} = \left(\frac{a}{\dot{a}}\right)_{eq} = \frac{1}{\sqrt{2}}\Omega_{\Lambda,0}^{-1/2} H_0^{-1} \approx 3.8 \times 10^{17} s. \tag{6.23}$$

These values are calculated for the ΛCDM model which assumes a flat FLRW universe, so they will not hold exactly in our toy model, but they will hold approximately, as the deviation from FLRW is minimal and is controlled by the value of ϵ at any given time. We can substitute $t(z)$ from (6.17) into $\epsilon(t)$ from (6.10) to get $\epsilon(z)$. Fig. 6.2 shows the plot. For larger redshifts (earlier times) $\epsilon(t)$ get closer to the isotropic case of $\epsilon = 1$.

Substituting $\epsilon(z)$ in (6.9) leads to the luminosity distance function in the toy model $d_{L,toy}(z,\theta)$:

$$d_{L,toy}(z,\theta) = d_{L,FRW}(z)\sqrt{\frac{\epsilon_o^4}{\epsilon^2(z)}\sin^2\theta + \cos^2\theta}. \tag{6.24}$$

Eq.(6.24) shows that for $\theta = 0$ the two models are indistinguishable, whereas one gets maximal deviation for $\theta = \pi/2$. Fig.6.3 shows a plot of (6.24) for $\theta = \pi/2$ and for different values of ϵ_0 including for the ΛCDM model which corresponds to $\epsilon_0 = 1$.

6.3 The JLA dataset

The JLA (Joint Light-Curve Analysis) dataset comprises 740 spectroscopically confirmed Type Ia Supernovae drawn from SDSS, SNLS and several other experiments [90]. A Type Ia Supernova occurs when a white dwarf accretes enough mass from a nearby

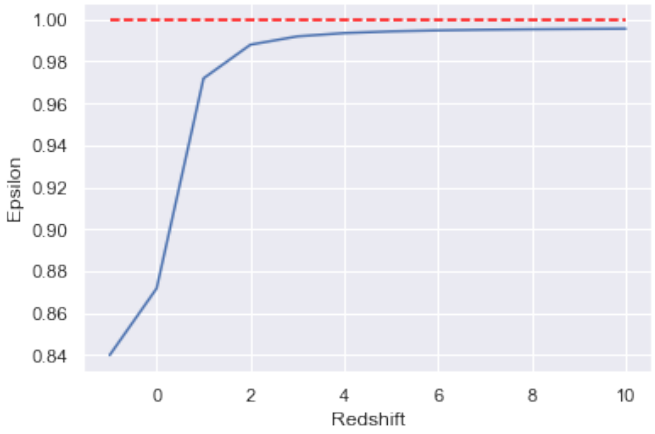

Figure 6.2: The function $\epsilon(z)$ with $\epsilon_0 = 0.84$. The red dashed line corresponds to $\epsilon(t) = 1$ which is the case of perfectly isotropic cosmology as in ΛCDM.

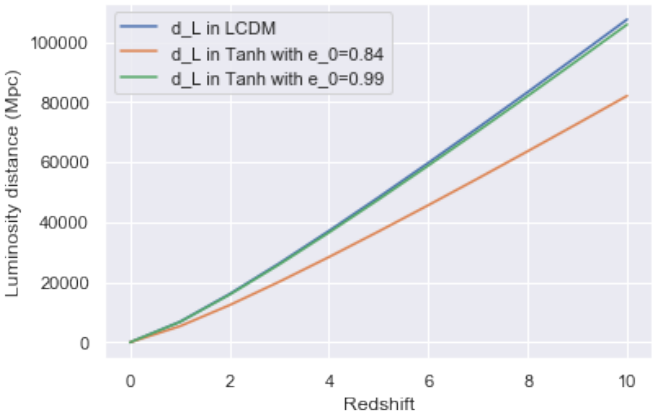

Figure 6.3: The luminosity distance as a function of redshift for ΛCDM model (blue), and Tanh model at $\theta = \pi/2$ for $\epsilon_0 = 0.2$ (orange) and $\epsilon_0 = 0.9$ (green).

object in order to ignite a thermonuclear explosion. These supernovae can be standardised and hence serve as distance indicators. In order to standardise them, one first needs to take into account that brighter supernovae shine for a longer time. The apparent brightness in the sky allows to measure the apparent magnitude, while the time scale for the decline of the brightness allows to estimate the absolute magnitude, and hence determine the luminosity distance of the object. In addition, one needs to perform a colour correction because brighter supernovae are bluer in colour. In this way for each standardised supernova we get a measurement of the luminosity distance and the redshift. The corresponding data points are shown in Fig.6.4 together with a theoretical $d_L - z$ curve derived from the ΛCDM model (we do not show the uncertainty of the luminosity distance). We see that the match is quite good and so JLA data is consistent with the

ΛCDM model and hence with an isotropic universe. However, the data is inconclusive because the number of supernovae is too low, they are not uniformly scattered in the sky, and most of the supernovae are located at low redshifts where any potential large-scale deviation from isotropy would not be apparent. As we will discuss in the Conclusion, future supernovae data from Euclid and LSST would involve more data points and at higher redshifts and hence would allow us to better test for a possible anisotropy of the universe. For now we will consider mock data sets and explore under what conditions they would be able to give us a signal of an anisotropy.

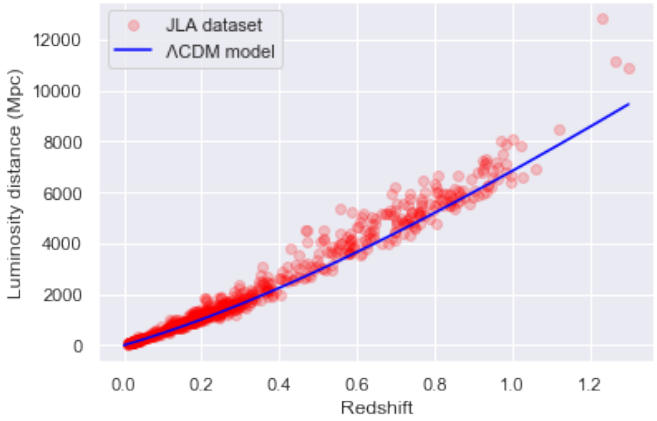

Figure 6.4: The luminosity distance – redshift curve for JLA data (red points) and the theoretically predicted one by the ΛCDM model (blue curve).

6.4 Statistical tests of the toy model

In order to test the anisotropic model and determine under what circumstances it would be distinguishable from ΛCDM, we first generate a mock catalog under the assumption that the toy model is correct. We fix the redshift to $z = 0.01$ for now and split the sky into 12288 pixels by setting NSIDE = 32 in the Healpix module. We select two orthogonal regions such that one of them is in the anisotropic direction and for each region we populate 30% of the pixels with supernovae. The value of d_L in each pixel is drawn from a Gaussian distribution with a mean $d_{L,toy}(z = 0.01)$ and a standard deviation $d_{L,toy}(z = 0.01)/20$. We also temporarily fix $\epsilon_0 = 0.84$ (this is the value of ϵ_0 most consistent with the amount of anisotropy claimed in [26]). Fig.6.5 shows an image of the sky, where the values of the luminosity distance in the two regions have been highlighted.

The setup that we consider is that the data is given by the mock data generated from the toy model and we want to check whether this data is compatible with the ΛCDM model. We offer two statistical tests to distinguish between the ΛCDM model and the late-time anisotropic model.

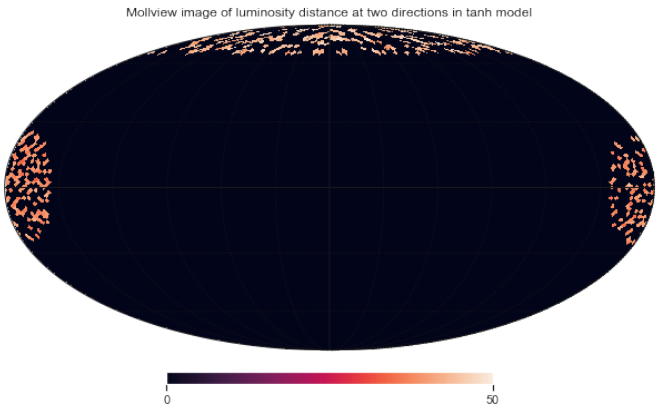

Figure 6.5: An image of the sky where two orthogonal regions are populated with supernovae. The brightness shows the value of the luminosity distance in Mpc.

Permutation test The null hypothesis is that the data in the two regions is drawn from an isotropic distribution in the sky (as in the ΛCDM model). We simulate the null hypothesis by drawing many artificial data sets under the assumption that the null hypothesis is true. For that, we combine all the values of the luminosity distance from the two patches into one array, permute the elements in the array and split the array into two new arrays corresponding to each of the two regions. Then for each artificial dataset, we calculate a test statistic which is given by the difference of the means of the luminosity distance in the two regions. We repeat that 10000 times and get a distribution of the test statistic which is shown on the histogram in Fig. 6.6. The red line corresponds to the value of the test statistic for the actual data (the one drawn by assuming that the late-time anisotropic model is correct). The p-value is defined as the fraction of the permutation samples for which the test statistic is at least as extreme as the value of the test statistic for the actual data. In this case: $p \approx 0$ (note that we can never claim that p is exactly 0 because drawing more permutation samples might eventually lead to a non-zero p), which implies that the null hypothesis is rejected: the data is not consistent with an isotropic universe. However, as we increase the value of ϵ_0, the toy model gets closer and closer to the ΛCDM model and at some threshold value, the red line in Fig. 6.6 will enter the 99% confidence interval and the two models would be indistinguishable at that confidence level.

Bootstrap replicate test Now we test a weaker null hypothesis: the luminosity distance distributions of the two patches are drawn from parent distributions with the same mean (but not necessarily the same parent distribution as in the previous case). In order to simulate the null hypothesis, we first calculate the mean of the full data comprising the two patches, then we shift the two arrays so that they have a mean equal to the common mean and then calculate the same test statistic as before. Repeating that 10000 times allows to calculate the p-value which is again $p \approx 0$. Therefore, the

null hypothesis is rejected (Fig.6.7).

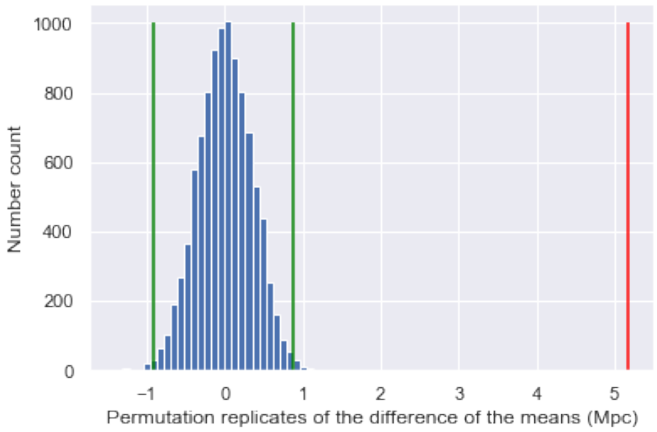

Figure 6.6: The distribution of the difference of the means for 10000 permutation replicates. The green lines show the 99% confidence interval. The red line shows the observed difference which is obviously incompatible with the null hypothesis.

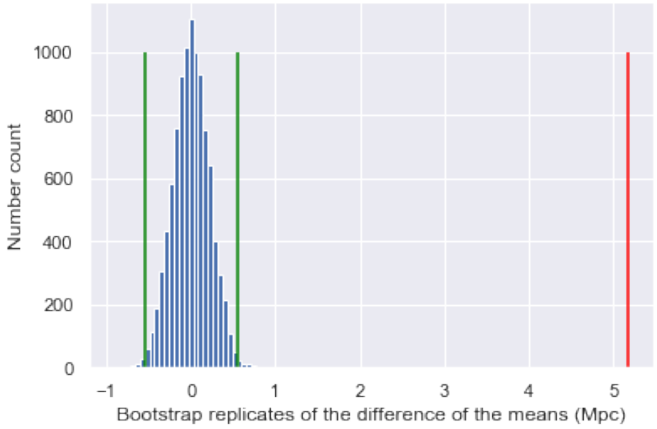

Figure 6.7: The distribution of the difference of the means for 10000 bootstrap replicates. The green lines show the 99% confidence interval. The red line shows the observed difference which is obviously incompatible with the null hypothesis.

6.5 Distinguishing the toy model from the ΛCDM with future data

The analysis performed so far shows that given enough supernovae (or other standard candles) in the sky, the two models – ΛCDM and the anisotropic toy model, can be

distinguished as long as ϵ_0 is sufficiently different from 1. The closer ϵ_0 is to 1, the more data we would need in order to distinguish them. This is something which is obvious even prior to any detailed analysis. A sensible question then is how much future data we would actually need in order to distinguish between the ΛCDM model and the anisotropic model for any non-trivial value of ϵ_0. We can begin to answer this question by parametrising any future data set by two free parameters: the fractional sky coverage parameter q and the Gaussian scattering parameter s. The former tells us the percentage of pixels in some region of the sky which is empty, i.e there are no supernovae in these pixels, and no value for the luminosity distance. Thus $1-q$ is proportional to the number of supernovae in this region of the sky. We assume that q is identical for both regions of the sky that we consider, and we set its default value to $q = 0.7$. That was the value of q that we used in the previous section.

The scattering parameter s represents the random error of the luminosity distance in a future dataset. This means that when we generate a mock dataset from the anisotropic toy model, we assume that for any given supernova, its luminosity distance is drawn from a Gaussian distribution with a mean equal to the theoretical prediction from the toy model, and a standard deviation equal to s times the mean. We set the default value of the scattering parameter to be $s = 0.05$ because that coincides with the average error in the luminosity distance from the JLA dataset [90]. That was the value of s that we used in the previous section.

Together with the free parameter of the toy model ϵ_0 (whose default value is 0.84 as before), we have in total 3 free parameters to keep track of. We run permutation tests for different values of the free parameters and keep track of the variation of the number of σ's between the two models, where σ is the standard deviation of the distribution of the test statistic for the different permutation samples. First, we vary only one of the free parameters, while keeping the other two at their default values. Next, we vary two of the free parameters, keeping the third one at its default value. Figs. 6.8–6.13 show the number of σ's between the two models as a function of the values of the free parameters in each of the six cases.

There are a few points that need to be stressed here. First, it is easier to distinguish between the two models for small values of each of the three parameters. For the anisotropy parameter, this is because smaller values of ϵ_0 imply a stronger anisotropy at late times and hence a greater theoretical difference between the two models. For the data parameters, this is because a smaller value of q implies more supernovae in the sky and hence a smaller sampling error, and a smaller value of s implies a smaller random error.

The second point is that the boundaries between the different regions on the two-dimensional plots are relatively smooth. The apparent jaggedness comes simply from the finite number of sample points that we drew for each parameter.

The third point is that the shapes of the curves in Figs.6.8–6.10 and the boundaries in Figs.6.11–6.13 remain approximately the same for different realisations of the mock data set and the permutation samples, and thus these curves capture general dependancies of the number of σ's in terms of the model– and data parameters.

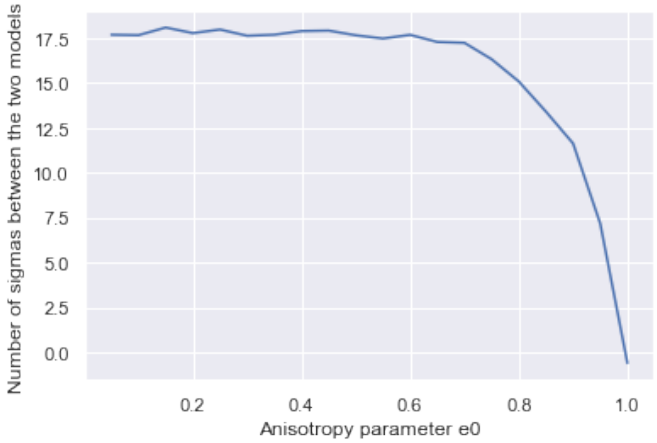

Figure 6.8: The number of σ's between the ΛCDM and the anisotropic toy model as a function of the model parameter ϵ_0 where the data parameters are kept at their default values: $q = 0.7$, $s = 0.05$.

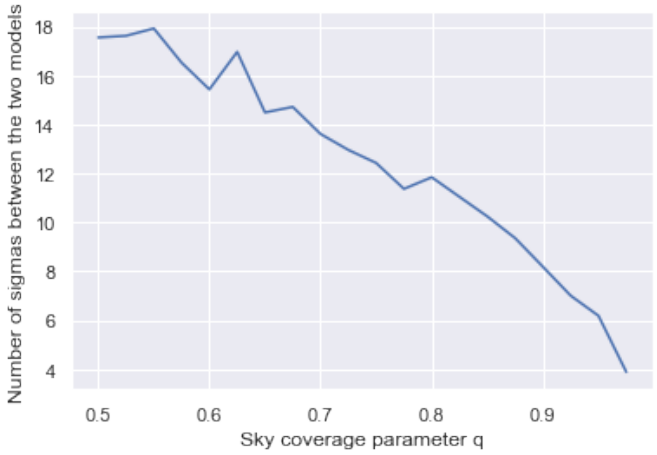

Figure 6.9: The number of σ's between the ΛCDM and the anisotropic toy model as a function of the data parameter q where the other data parameter and the model parameter are kept at their default values: $s = 0.05$, $\epsilon_0 = 0.8$.

Note that Fig.6.9 shows that if we keep ϵ_0 and s and their default values, we can distinguish between the two models even for a large value of q and thus for a small number of supernovae. However, these supernovae would need to be located precisely in the anisotropic direction, and since we do not know a priori what such a direction might be, we need a significant number of supernovae distributed over the whole sky.

In addition, this analysis assumes that we have removed all sources of systematic errors and that we have very good understanding of the astrophysics of supernovae.

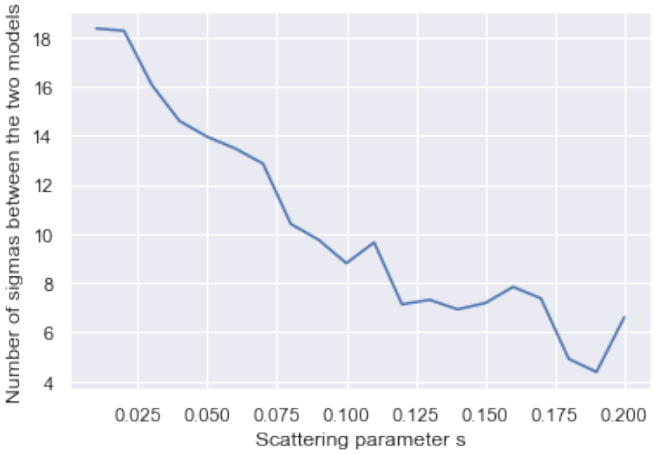

Figure 6.10: The number of σ's between the ΛCDM and the anisotropic toy model as a function of the data parameter s where the other data parameter and the model parameter are kept at their default values: $q = 0.7$, $\epsilon_0 = 0.8$.

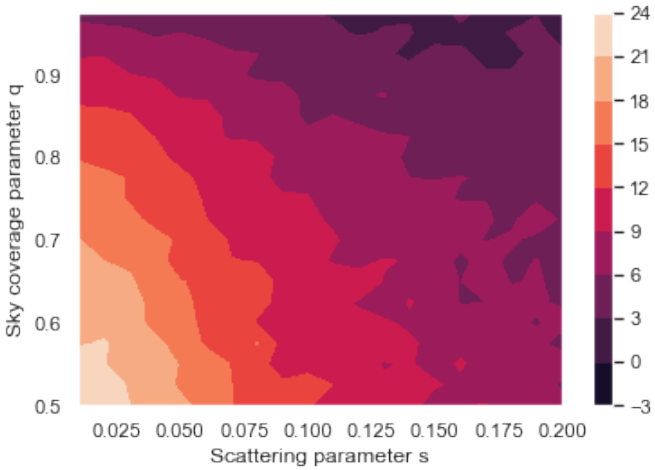

Figure 6.11: The number of σ's between the ΛCDM and the anisotropic toy model as a function of the data parameters s where the model parameter is kept at its default value $\epsilon_0 = 0.8$. The colour code shows the number of σ's.

In reality, this would not be the case, and we would need more supernovae in order to perform cross-checks and validate any particular claim of an anisotropy. In a similar way, Fig.6.10 shows that we can distinguish the two models even for a large random error in the luminosity distance measurements, but again by making idealised assumptions which will be unlikely to hold in real data.

So far we have kept the redshift fixed at $z = 0.01$. However, Fig.6.3 shows that the difference between the two models tends to become bigger as we increase the redshift.

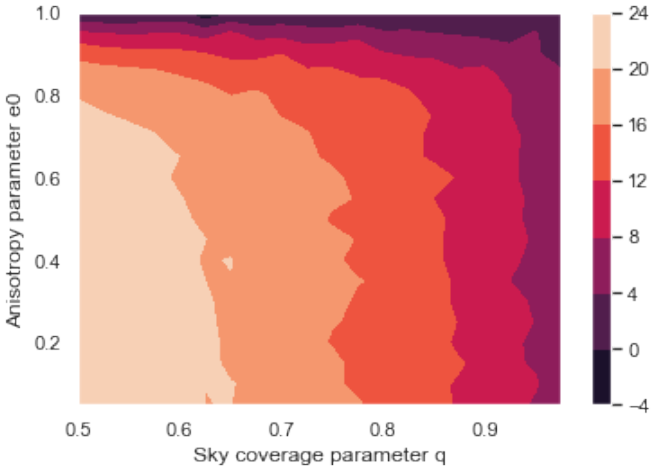

Figure 6.12: The number of σ's between the ΛCDM and the anisotropic toy model as a function of the data parameter q and the model parameter ϵ_0 where the other data parameter is kept at its default value $s = 0.05$. The colour code shows the number of σ's.

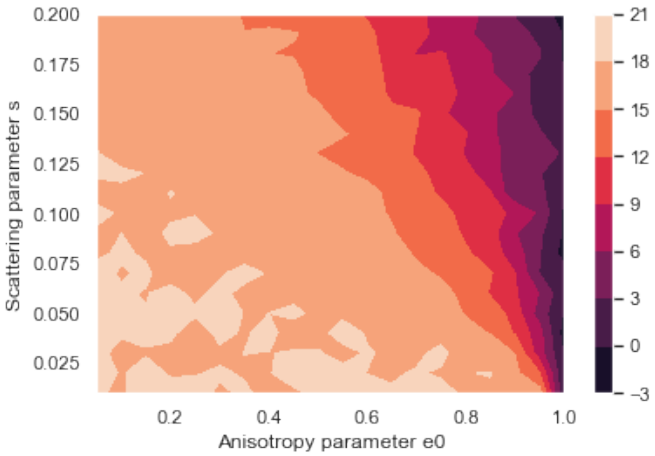

Figure 6.13: The number of σ's between the ΛCDM and the anisotropic toy model as a function of the data parameter s and the model parameter ϵ_0 where the other data parameter is kept at its default value $q = 0.7$. The colour code shows the number of σ's.

Thus having more supernovae at higher redshifts would help to distinguish between the two models. This relation is shown in Fig.6.14. Overall, we would be able to detect an anisotropy in the universe by the permutation test outlined above, if we had more supernovae and these supernovae were located at higher redshifts, and if they were distributed in the sky so that at least some of them lie in the anisotropic direction.

To summarise, while detecting a potential anisotropy of the universe with present

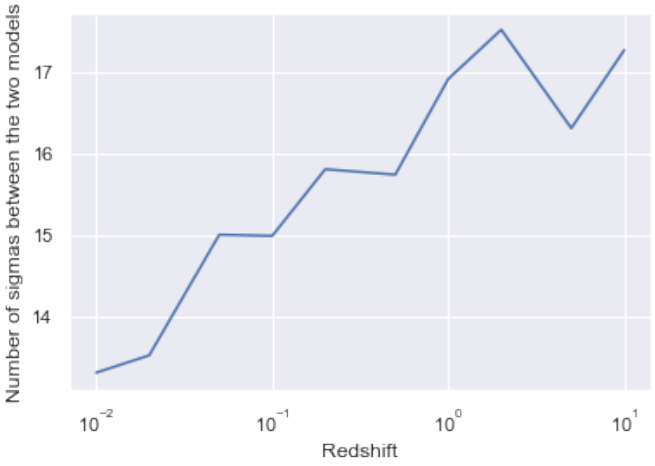

Figure 6.14: The number of σ's between the ΛCDM and the anisotropic toy model as a function of the redshift at which the supernovae are located where the model parameter ϵ_0 and the data parameters q and s are kept at their default values.

data might be difficult or even impossible, future data might allow us to do that as long as the anisotropy is big enough. In particular, by considering a simple extension of the ΛCDM model where the amount of anisotropy at late times is given by a single parameter ϵ_0, we have investigated what properties a future data set might need in order to allow us to distinguish this model from the standard one. We stress that there are several assumptions in this analysis: that the isotropy is broken along a single direction, that there are supernovae along this direction and a perpendicular direction, and that the supernovae have a well-measurable luminosity distance with an error proportional the value of the luminosity distance. We have parametrised the number of supernovae by a parameter q and the random error by a parameter s, and have shown in Figs. 6.8–6.13, the values of these parameters and the model parameter ϵ_0 needed to distinguish between the two models at any particular significance level.

7 Conclusion

The ΛCDM model is the crowning achievement in our attempts to understand the Universe. It provides a very robust cosmological framework, in that it is both very simple and elegant (being based on clear, justified assumptions with only six free parameters), and very well tested with different cosmological probes. It is remarkable that a biological species that has been doing science for only 10^{-8} of the age of the Universe has gone so far into understanding its structure and evolution.

The robustness of the standard cosmological model can be seen both as an advantage and a disadvantage. One the one hand, it implies that we have understood well the basic physics on which it is built and it gives a clear starting point for further model-building. Any alternative cosmological model would have to be based on the ΛCDM and would have to reduce to it in the regime in which it has been successfully tested. For example, the DM BEC model that we considered in Chapter 2 reduces to CDM on large scales, while modifying the CDM-like behaviour only on small (galactic) and intermediate (cluster) scales. Similarly, the conformal FLRW and linearly perturbed FLRW metrics that we considered in Chapter 6 represent simple modifications of the FLRW metric on which the ΛCDM model is built (at the background level).

On the other hand, the robustness of the standard cosmological model makes it extremely difficult to make further progress. The problem is that any extension of the ΛCDM would have to solve or at least improve upon some of its problems in order to be worth considering, and at the same time be minimal enough in order to preserve its successes. This would make it very difficult to experimentally distinguish such an extension from the standard case by using current data. We addressed this problem in Chapter 6, in the context of a possible deviation from the Cosmological Principle. Still, new cosmological probes such as binary neutron star mergers, more precise CMB measurements, the expected launches of the Euclid and LISA missions, and new measurements of standard candles such Type Ia supernovae and possibly GRBs, might give us enough leverage in order to distinguish between the leading proposed modifications of the ΛCDM and hence settle on a new extended cosmological model which would fit the data better than ΛCDM and would resolve most of its issues. As an example, one way

to put the BEC DM model to the test is to look for correlations between the arrival time difference between GWs and GRBs from binary neutron star mergers and the integrated gradient of the density along the path of the waves.

Of particular importance and interest are the LISA and Euclid missions. LISA (Laser Interferometer Space Antenna) is going to be the first space-based GW observatory [91]. While ground-based observatories are only able to observe GWs in the frequency range $10-10^3$ Hz, LISA will also be able to cover the interval $10^{-4}-1$ Hz. Therefore, it will be able to detect GWs emitted by more massive objects at earlier times. This will allow us to probe extreme gravitational physics such as in the very early universe and surrounding black holes, to make further cosmological tests of General Relativity, to probe the dark universe, and to probe compact binary star systems and merging black holes. A very important observable will be the gravitational-wave luminosity distance d_L^{gw} which is defined in the same way as the electromagnetic one, but in terms of GWs [59]. While it coincides with the standard luminosity distance in GR, the two distances are different in modified theories of gravity, and also in models where either DM or DE is non-minimally coupled to the metric. Thus it could give an additional independent test of the BEC DM model. Furthermore, since binary neutron star systems serve as standard sirens, one can also do cosmography with the GW luminosity distance. This would allow both for an independent measurement of H_0, and for measuring the acceleration of the universe in more detail. So it will be able to probe the expansion of the universe at earlier times and search for potential anisotropies at late times.

The Euclid mission involves the launch of a space-based telescope which will map the geometry of the Universe and will help us to understand better the nature of the dark components of the Universe [92]. It will measure the shapes and redshifts of galaxies and clusters up to $z \sim 2$ and since matter–dark energy equality occurs at $z_{eq} \sim 0.3$, it will cover the whole period of time when dark energy is important. Thus by using the cosmographic approach, it will be possible to probe better the expansion and acceleration of the universe at earlier times. In particular, Euclid will allow us to measure the jerk parameter j (Eqn. (4.81)) to a very high accuracy and hence to determine whether dark energy is actually constant or not. By looking at different directions in the sky, it will also allow us to tell whether there are any signs of anisotropy in the cosmological expansion. The telescope will also investigate the structure and distribution of dark matter in the universe - first by tracking the evolution and distribution of visible matter, and second by weak lensing. Dark matter at the foreground distorts the images of background galaxies and thus allows to form 3D maps of the dark matter distribution in the universe. Furthermore, Euclid will be able to detect about 140 high-quality superluminous supernovae (SLSNe-I) up to $z \sim 3.5$ [93]. They will allow for astrophysical tests, such as probing star-formation rates, as well as the interstellar and intergalactic mediums. In addition, a fraction of them could be calibrated and standartised, and hence can serve as standard candles. With their aid we would be able to constrain better the equation of state of dark energy $w(z)$ in the early (high-redshift) universe.

Another very important upcoming experiment is the Large Synoptic Survey Telescope (LSST), recently renamed the Rubin observatory, which will make detailed light

observations of the low-redshift universe [94]. In particular, it will be able to probe dark matter and dark energy by measuring the effects of weak and strong lensing of billions of galaxies, galaxy correlation functions and Type Ia supernovae up to $z \sim 1$. The many supernovae that are expected to be discovered – $O(10^5)$, will allow us to study in detail their astrophysical properties and evolution, to perform cross-calibrations and hence to reduce the systematic error in luminosity-distance measurements. They will also help us to further constrain $w(z)$ in the late (low-redshift) universe and to potentially distinguish between a cosmological constant, dynamical dark energy and modified gravity. Overall, precision measurements from the LSST will allow us to better constrain the composition of the universe, the parameters of the ΛCDM and possible deviations especially in the dark sector.

Beside the issues faced by the ΛCDM mentioned in the Introduction, there are further issues related to initial conditions and the earliest stages of the evolution of the Universe. It is well known that in order to have a well-defined thermodynamic arrow of time, the very early Universe has to be in a state of a very low entropy (this is referred to as the Past Hypothesis). However, we do not know yet how to define and calculate the entropy of a general-relativistic gravitational system. The low entropy of the early universe is most often associated with its homogeneity and the smoothness of the matter distribution, and this is something that is supposed to be explained by inflation. However, the initial conditions for inflation would also have to be fine-tuned in order to get the right amount of inflation and therefore inflation by itself cannot resolve the issue of initial conditions – a deeper explanation for the arrow of time is required.

Moreover, at very early times the temperature of the Universe is likely to reach the Planck scale, and the state of the Universe would then be described by an unknown quantum theory of gravity. While classical General Relativity predicts that timelike geodesics hit a singularity in the past, this singularity is very likely to disappear at the quantum level, and hence the very early Universe would be ruled by some exotic and yet unknown physics. This is another direction in which the ΛCDM should be extended and ideally embedded in a quantum cosmological model.

To conclude, this thesis attempted to make a modest contribution towards our understanding of what might lie beyond the ΛCDM and how it could be tested. We looked at a model where dark matter forms a macroscopic Bose-Einstein condensate and acquires a non-minimal coupling to the metric, and we constrained this model by using the GW170817 event. We also considered the possibility that dark energy might be something more than a cosmological constant, and motivated by this, we extended the luminosity distance – redshift formalism to conformally FLRW and linearly perturbed FLRW spacetimes, thus allowing us to generalise the cosmographic approach to an inhomogeneous universe. We also addressed the question of whether the universe might have a large-scale anisotropy and looked at how future data might shed some light on this question. Overall, the next decades are expected to bring some exciting new developments in the whole field of Cosmology which will allow us to test all different components of the ΛCDM model - General Relativity, dark matter, dark energy and the Cosmological Principle. This exemplifies the spirit of science in which we constantly update our

best models in order to improve our understanding of the Universe.

Appendices

A

Cosmological linear perturbation theory

The first step in Cosmological Perturbation Theory is to expand the metric into a background plus a small perturbation:

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} = a^2(\eta)(\gamma_{\mu\nu} + h_{\mu\nu}). \tag{A.1}$$

We are in the linear regime as long as $h_{\mu\nu} \ll \gamma_{\mu\nu}$. The stress-energy tensor is expanded in a similar way:

$$T_{\mu\nu} = \bar{T}_{\mu\nu} + \delta T_{\mu\nu}. \tag{A.2}$$

Substituting the metric and the stress-energy tensor in EFE and the equation for covariant conservation leads to dynamical equations at the background and perturbation levels

$$\bar{G}_{\mu\nu} = 8\pi G \bar{T}_{\mu\nu}, \qquad \nabla^\mu \bar{T}_{\mu\nu} = 0, \tag{A.3}$$

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}, \qquad \nabla^\mu (\delta T_{\mu\nu}) = 0. \tag{A.4}$$

It will be useful to classify perturbations according to how they transform under spatial rotations. A scalar perturbation s remains invariant under spatial rotations. A vector perturbation v^i can be decomposed into a scalar-type vector and vector-type vector in the following way:

$$v^i = \nabla^i s + v^i_*, \tag{A.5}$$

where s is the solution of the Poisson equation

$$\Box s = \vec{\nabla}.\vec{v}, \tag{A.6}$$

which guarantees that $\vec{v}_*$ is divergenceless:

$$\vec{\nabla}.\vec{v}_* = 0. \tag{A.7}$$

A tensor perturbation can be split into a scalar-type tensor, a vector-type tensor and a tensor-type tensor:

$$t^{ij} = \nabla^j \nabla^i s + \nabla^j v^i_* + t^{ij}_*, \tag{A.8}$$

where s satisfies the Poisson equation

$$\Box s = \mathrm{Tr}(t), \tag{A.9}$$

which guarantees that $\vec{v}_*$ is divergenceless and t_* is traceless:

$$\vec{\nabla}.\vec{v}_* = 0, \qquad \mathrm{Tr}(t_*) = 0. \tag{A.10}$$

In addition, we require that $\vec{v}_*$ satisfies

$$\nabla^2 v_*^i = \nabla_j t^{ij} - \nabla^2(\nabla^i s), \tag{A.11}$$

which implies that t_* is further constrained:

$$\nabla_j t^{ij} = 0, \tag{A.12}$$

i.e. t_* is both transverse (divergenceless) and traceless. From the dynamical equations it can be shown that this is a propagating mode and it represents gravitational waves.

The next step is to expand the different types of perturbations in a harmonic basis, i.e. in terms of the eigenfunctions of the Laplace operator $\gamma^{ij}\nabla_i\nabla_j$. Scalars can be expanded in terms of the functions $Y(\vec{k},\vec{x})$ satifying

$$(\Box_\gamma + k^2)Y(\vec{k},\vec{x}) = 0. \tag{A.13}$$

Scalar-type vectors are expanded in terms of

$$Y_i \equiv -\frac{1}{k}\nabla_i Y, \tag{A.14}$$

while scalar-type tensors are expanded in terms of

$$Y_{ij} \equiv \frac{1}{k^2}\nabla_j\nabla_i Y + \frac{1}{3}\gamma_{ij}Y. \tag{A.15}$$

Vector-type vectors are expanded in terms of the vector-functions $Y_i^{(\pm1)}(\vec{k},\vec{x})$:

$$(\Box_\gamma + k^2)Y_i^{(\pm1)}(\vec{k},\vec{x}) = 0, \tag{A.16}$$

while vector-type tensors are expanded in terms of

$$Y_{ij}^{(\pm1)} = -\frac{1}{2k}(\nabla_j Y_i^{(\pm1)} + \nabla_i Y_j^{(\pm1)}). \tag{A.17}$$

Finally, tensor-type tensors are expanded in terms of $Y_i^{(\pm2)}(\vec{k},\vec{x})$:

$$(\Box_\gamma + k^2)Y_i^{(\pm2)}(\vec{k},\vec{x}) = 0. \tag{A.18}$$

Now we proceed to expand the different perturbations of the metric in a Fourier basis. We need to take into account the scalar perturbation δg_{00}, the vector perturbation δg_{0i},

which consists of a scalar-type vector $\delta g_{0i}^{(0)}$ and vector-type vector $\delta g_{0i}^{(\pm1)}$, and the tensor perturbation δg_{ij}, which consists of a scalar-type tensor $\delta g_{ij}^{(0)}$, a vector-type tensor $\delta g_{ij}^{(\pm1)}$ and a tensor-type tensor $\delta g_{ij}^{(\pm2)}$:

$$\delta g_{00}(\eta,\vec{x}) = -2a^2(\eta)\int \frac{\mathrm{d}^3k}{(2\pi)^3} A(\eta,\vec{k})Y(\vec{k},\vec{x}) \tag{A.19}$$

$$\delta g_{0i}^{(0)}(\eta,\vec{x}) = -a^2(\eta)\int \frac{\mathrm{d}^3k}{(2\pi)^3} B(\eta,\vec{k})Y_i(\vec{k},\vec{x}) \tag{A.20}$$

$$\delta g_{ij}^{(0)}(\eta,\vec{x}) = 2a^2(\eta)\int \frac{\mathrm{d}^3k}{(2\pi)^3}\Big(H_L(\eta,\vec{k})Y(\vec{k},\vec{x})\gamma_{ij} + H_T(\eta,\vec{k})Y_{ij}(\vec{k},\vec{x})\Big) \tag{A.21}$$

$$\delta g_{0i}^{(\pm1)}(\eta,\vec{x}) = -a^2(\eta)\int \frac{\mathrm{d}^3k}{(2\pi)^3} B^{(\pm1)}(\eta,\vec{k})Y_i^{(\pm1)}(\vec{k},\vec{x}) \tag{A.22}$$

$$\delta g_{ij}^{(\pm1)}(\eta,\vec{x}) = 2a^2(\eta)\int \frac{\mathrm{d}^3k}{(2\pi)^3} H_T^{(\pm1)}(\eta,\vec{k})Y_{ij}^{(\pm1)}(\vec{k},\vec{x}) \tag{A.23}$$

$$\delta g_{ij}^{(\pm2)}(\eta,\vec{x}) = 2a^2(\eta)\int \frac{\mathrm{d}^3k}{(2\pi)^3} H_T^{(\pm2)}(\eta,\vec{k})Y_{ij}^{(\pm2)}(\vec{k},\vec{x}) \tag{A.24}$$

General Relativity and all other physical theories should be invariant under change of coordinates. This means that all equations should change covariantly and therefore should be constructed from tensors. For example, when going from a frame F to a frame $\tilde{F}$, the metric transforms as

$$g_{\mu\nu}(x) = \frac{\partial\tilde{x}^{\mu'}}{\partial x^{\mu}}\frac{\partial\tilde{x}^{\nu'}}{\partial x^{\nu}}\tilde{g}_{\mu'\nu'}(\tilde{x}). \tag{A.25}$$

We can apply this lesson to infinitesimal transformations:

$$\tilde{x}^{\mu} = x^{\mu} + \delta x^{\mu}(\eta,\vec{x}). \tag{A.26}$$

We refer to different frames infinitesimally shifted from the Hubble frame as different gauges and to transformations between such frames as gauge transformations. The infinitesimal shift δx^{μ} can itself be expanded in terms of the Fourier basis:

$$\delta x^0(\eta,\vec{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} T(\eta,\vec{k})Y(\vec{k},\vec{x}), \tag{A.27}$$

$$\delta x_i(\eta,\vec{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\Big(L(\eta,\vec{k})Y_i(\vec{k},\vec{x}) + L^{(\pm1)}(\eta,\vec{k})Y_i^{(\pm1)}(\vec{k},\vec{x})\Big). \tag{A.28}$$

Therefore by using the general rule (A.25), we can derive the transformation equations for the perturbations of the metric:

$$\tilde{A} = A - T' - \mathcal{H}T, \tag{A.29}$$

$$\tilde{B} = B + L' + kT, \tag{A.30}$$

$$\tilde{H}_L = H_L - \frac{k}{3}L - \mathcal{H}T, \tag{A.31}$$

$$\tilde{H}_T = H_T + kL, \tag{A.32}$$

$$\tilde{B}^{(\pm 1)} = B^{(\pm 1)} + L'^{(\pm 1)}, \tag{A.33}$$

$$\tilde{H}_L^{(\pm 1)} = H_L^{(\pm 1)} - \frac{k}{3}L^{(\pm 1)}, \tag{A.34}$$

$$\tilde{H}_T^{(\pm 1)} = H_T^{(\pm 1)} + kL^{(\pm 1)}, \tag{A.35}$$

$$\tilde{H}_T^{(\pm 2)} = H_T^{(\pm 2)}. \tag{A.36}$$

These equations show that it is always possible to go to some frame where a particular component of the perturbation vanishes. However, it is not possible to completely erase the perturbation by performing coordinate shifts. They also show that the different functions A, B, H_L, etc. are gauge-dependent. However, it is possible to construct gauge-invariant potentials by combining several such functions. These are known as the Bardeen potentials:

$$\Phi = H_L + \frac{1}{3}H_T + \frac{\mathcal{H}}{k}(B - \frac{1}{k}H_T'), \tag{A.37}$$

$$\Psi = A + \frac{\mathcal{H}}{k}(B - \frac{1}{k}H_T') + \frac{1}{k}(B' - \frac{1}{k}H_T''). \tag{A.38}$$

It is also usual in cosmology to fix a particular gauge, to derive all important results in that gauge and then to write them in a covariant fashion so that they hold in all gauges. A common choice of gauge is the Newtonian gauge, given by the requirement:

$$\tilde{B} = \tilde{H}_T = \tilde{H}_T^{(\pm 1)} = 0, \tag{A.39}$$

$$\tilde{B} = \tilde{H}_T = \tilde{H}_T^{(\pm 1)} = 0. \tag{A.40}$$

In the Newtonian gauge, the formulae for the Bardeen potentials simplify a lot:

$$\Phi = H_L, \qquad \Psi = A. \tag{A.41}$$

We can actually go back to position space and write the Bardeen potentials as

$$\Phi(\eta, \vec{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} H_L(\eta, \vec{k}) Y(\vec{k}, \vec{x}), \tag{A.42}$$

$$\Psi(\eta, \vec{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} A(\eta, \vec{k}) Y(\vec{k}, \vec{x}). \tag{A.43}$$

Then the metric can simply be written as

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\ &= (\bar{g}_{00} + \delta g_{00}) dt^2 + (\bar{g}_{ij} + \delta g_{ij}) dx^i dx^j \\ &= a^2(\eta)[-(1 + 2\Phi) dt^2 + (1 - 2\Psi)\gamma_{ij} dx^i dx^j]. \end{aligned} \tag{A.44}$$

We can repeat the same kind of analysis for the perturbations of the stress-energy tensor and hence from (A.4) derive the dynamical equations for the evolution of perturbations.

B

The Relation between $G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi$ and $\mathcal{L}_4$ of Horndeski

Starting from $\mathcal{L}_5$ of the Horndeski Lagrangian with $G_5(\phi, X) = L^2\phi$:

$$\mathcal{L}_5 = L^2\phi G_{\mu\nu}\nabla^\mu\nabla^\nu\phi, \tag{B.1}$$

integrating by parts, ignoring the boundary term and using the Bianchi identity: $\nabla^\mu G_{\mu\nu} = 0$, leads to

$$\mathcal{L}_5 = -L^2 G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi. \tag{B.2}$$

Thus the coupling $L^2 G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi$ is a subclass of the Horndeski action. Now we show that it is equivalent to $\mathcal{L}_4$ of Horndeski with $G_4(\phi, X) = L^2 X$:

$$\mathcal{L}_4 = L^2 XR - L^2\big[(\Box\phi)^2 - \nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi\big]. \tag{B.3}$$

Starting from $G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi$ we have

$$\begin{aligned} G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi &= R_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi - XR \\ &= \nabla_\rho\nabla_\nu\nabla^\rho\phi\nabla^\nu\phi - \nabla_\nu\nabla_\rho\nabla^\rho\phi\nabla^\nu\phi - XR \\ &= -\nabla_\nu\nabla_\rho\phi\nabla^\nu\nabla^\rho\phi + (\Box\phi)^2 - XR \\ &= -\frac{1}{L^2}\mathcal{L}_4, \end{aligned} \tag{B.4}$$

where in the first line we use the definition of the Einstein tensor, in the second line we use the definition of the Riemann tensor

$$R^\rho_{\mu\sigma\nu}V^\mu = \nabla_\sigma\nabla_\nu V^\rho - \nabla_\nu\nabla_\sigma V^\rho \tag{B.5}$$

with $V^\mu = \nabla^\mu\phi$, in the third line we integrate by parts ignoring the boundary terms, and in the fourth line we use (B.3). Notice the minus sign that we have picked in the integration by parts. What we have essentially shown is that

$$\mathcal{L} = \frac{1}{16\pi G}R + L^2 G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi \tag{B.6}$$

is equivalent to

$$\mathcal{L} = \frac{1}{16\pi G}R - L^2\Big\{XR - \big[(\Box\phi)^2 - \nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi\big]\Big\}. \tag{B.7}$$

C

Conformal transformations

If we have two conformally related metrics $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$:

$$\mathrm{d}\tilde{s}^2 = \tilde{g}_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = f^2(x) g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = f^2(x)\mathrm{d}s^2. \tag{C.1}$$

then the inverse metrics are related by:

$$\tilde{g}^{\mu\nu} = f^{-2}(x) g^{\mu\nu}. \tag{C.2}$$

The connections are related by:

$$\tilde{\Gamma}^\nu_{\mu\rho} = \Gamma^\nu_{\mu\rho} + \delta^\nu_\rho \frac{\partial_\mu f}{f} + \delta^\nu_\mu \frac{\partial_\rho f}{f} - g^{\nu\sigma} g_{\mu\rho} \frac{\partial_\sigma f}{f}, \tag{C.3}$$

and we have that for an arbitrary vector V^μ:

$$V^\mu \tilde{\nabla}_\mu V^\nu = V^\mu \nabla_\mu V^\nu + 2V^\nu V^\mu \partial_\mu \ln f - (g_{\mu\rho} V^\mu V^\rho) g^{\nu\sigma} \partial_\sigma \ln f. \tag{C.4}$$

This equation shows that if V^μ satisfies the geodesic equation for the metric $g_{\mu\nu}$, it does not in general satisfy the geodesic equation for the metric $\tilde{g}_{\mu\nu}$. But if V^μ is a null vector with affine parameter λ:

$$l^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}, \tag{C.5}$$

and

$$l^\mu \nabla_\mu l^\nu = 0, \tag{C.6}$$

then

$$l^\mu \tilde{\nabla}_\mu l^\nu = \alpha l^\nu, \tag{C.7}$$

where

$$\alpha := 2l^\mu \partial_\mu \ln f = 2\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln f. \tag{C.8}$$

Therefore, l^μ satisfies the geodesic equation in the conformal metric $g_{\mu\nu}$ but λ is no longer an affine parameter. However, we can change parametrisation to a new affine parameter $\tilde{\lambda}$. The tangent vector in terms of the new affine parameter is given by:

$$\tilde{l}^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tilde{\lambda}}, \tag{C.9}$$

and we have that:

$$l^\mu = \frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\lambda}\, \tilde{l}^\mu. \tag{C.10}$$

Substituting that in (C.7) and imposing that the affine geodesic equation holds for $\tilde{\lambda}$, we obtain a differential equation for $\tilde{\lambda}$:

$$\frac{\mathrm{d}^2\tilde{\lambda}}{\mathrm{d}\lambda^2} = \alpha \frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\lambda}. \tag{C.11}$$

Solving it, we obtain:

$$\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\lambda} = K \exp\left(\int \alpha \mathrm{d}\lambda\right). \tag{C.12}$$

Using (C.8) and setting $K = 1$ we get that:

$$\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\lambda} = f^2. \tag{C.13}$$

Hence,

$$\tilde{l}^\mu = \frac{1}{f^2}\, l^\mu. \tag{C.14}$$

If instead we have a timelike geodesic with tangent vector:

$$U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}, \tag{C.15}$$

normalized in the metric $g_{\mu\nu}$

$$g_{\mu\nu} U^\mu U^\nu = -1, \tag{C.16}$$

then it won't be normalized in the conformal metric $\tilde{g}_{\mu\nu}$:

$$\tilde{g}_{\mu\nu} U^\mu U^\nu = -f^2. \tag{C.17}$$

So we have to reparameterize:

$$\mathrm{d}\tilde{\tau} = f\, \mathrm{d}\tau; \qquad \tilde{U}^\mu = \frac{1}{f}\, U^\mu. \tag{C.18}$$

D

Demonstration that $\delta\theta_s \perp (k_s, U_s)$ and $\delta x_o \perp (k_o, U_o)$

Here we demonstrate that the vectors $\delta\theta_s^\mu$ and δx_o^μ indeed belong to two dimensional subspaces orthogonal to $k^\mu = \tilde{\omega}\ell^\mu$ and to U_s^μ, U_o^μ.

$\delta x_o \perp k_o$: Since all photons start at the same point in spacetime, they must have the same phase P defined as

$$\ell_\mu = \nabla_\mu P. \tag{D.1}$$

Since the phase does not change along a cross section of the congruence, we must have that

$$0 = \nabla_{\delta x} P = \delta x^\mu \nabla_\mu P = \delta x^\mu \ell_\mu, \tag{D.2}$$

which implies that $\delta x \perp k_o$.

$\delta\theta_s \perp k_s$: Define

$$v^\mu := \frac{DY^\mu}{d\lambda} = \ell^\rho \nabla_\rho Y^\mu, \tag{D.3}$$

so that $\delta\theta^\mu = v^\mu \delta y$. Then we have that

$$v^\mu \ell_\mu = \ell_\mu \ell^\rho \nabla_\rho Y^\mu = \ell^\rho \nabla_\rho (l_\mu Y^\mu) - Y^\mu \ell^\rho \nabla_\rho \ell_\mu = 0, \tag{D.4}$$

where the first term vanishes due to (D.2) and the second term vanishes due to the geodesic equation. This implies that $\delta\theta_s \perp k_s$.

$\delta\theta_s \perp U_s$: This follows from the fact that spacetime at the source S is locally Minkowski and the emission of light is isotropic in all directions.

$\delta x_o \perp U_o$: In order for this to hold we must choose a suitable parametrisation of the one-parameter family of null geodesics. Let's say that we start with parameters (λ, y)

such that $\delta x_o \cdot U_o \neq 0$. We can obtain new parameters $(\tilde{\lambda}, \tilde{y})$ by performing a general coordinate transformation on the 2-surface spanned by Y^μ and ℓ^μ

$$\lambda = g_1(\tilde{\lambda}, \tilde{y}); \tag{D.5}$$

$$y = g_2(\tilde{\lambda}, \tilde{y}). \tag{D.6}$$

However, we want this transformation to preserve the null geodesic curves and to preserve the affinity of the parameter λ. Thus we are left with

$$\lambda = \tilde{\lambda} + h(\tilde{y}); \tag{D.7}$$

$$y = g(\tilde{y}). \tag{D.8}$$

This implies that

$$\begin{aligned} \tilde{Y}^\mu &= \frac{\partial f^\mu}{\partial \tilde{y}} \\ &= \frac{\partial \lambda}{\partial \tilde{y}} \frac{\partial f^\mu}{\partial \lambda} + \frac{\partial y}{\partial \tilde{y}} \frac{\partial f^\mu}{\partial y} \\ &= \frac{\partial h}{\partial \tilde{y}} \ell^\mu + \frac{\partial y}{\partial \tilde{y}} Y^\mu, \end{aligned} \tag{D.9}$$

which in turn implies

$$\delta \tilde{x}^\mu := \tilde{Y}^\mu \delta \tilde{y} = \ell^\mu \delta h + \delta x^\mu. \tag{D.10}$$

Hence

$$\delta \tilde{x}^\mu_o U_{O\mu} = (l^\mu_O U_{O\mu}) \delta h + \delta x^\mu_o U_{O\mu}, \tag{D.11}$$

and this will be zero, provided we choose the function h such that

$$\delta h = -\frac{\delta x^\mu_o U_{O\mu}}{\ell^\mu_o U_{O\mu}}. \tag{D.12}$$

E

Calculating the Jacobi map and Jacobi determinant

We now show the full calculation of the Jacobi map and Jacobi determinant in the perturbed FLRW spacetime. We first work in the unphysical spacetime (5.15). The system of equations (4.38) reduces to

$$\frac{\mathrm{d}}{\mathrm{d}\hat{\lambda}}(\delta\hat{x}^{(1)\alpha}) = C^{(1)\alpha}_{\nu}(\hat{\lambda})\,\delta\hat{\bar{x}}^{\nu} + (\delta\hat{\theta}^{\alpha})^{(1)}, \tag{E.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}\hat{\lambda}}(\delta\hat{\theta}^{\alpha})^{(1)} = A^{(1)\alpha}_{\nu}(\hat{\lambda})\,\delta\hat{\bar{x}}^{\nu} + C^{(1)\alpha}_{\nu}(\hat{\lambda})\,(\overline{\delta\hat{\theta}^{\alpha}}). \tag{E.2}$$

The background equations are the same as those for Minkowski space. Therefore the background unprojected Jacobi map is given by

$$\hat{\bar{\mathcal{J}}}^{\alpha}_{\beta} = (\hat{\lambda}_o - \hat{\lambda}_s)\,\delta^{\alpha}_{\beta}, \tag{E.3}$$

while the first order correction to the unprojected Jacobi map is given by

$$\begin{aligned}\hat{\mathcal{J}}^{(1)\alpha}_{\beta} = &\int_{\hat{\lambda}_s}^{\hat{\lambda}_o} C^{(1)\alpha}_{\beta}(\hat{\lambda})(\hat{\lambda} - \hat{\lambda}_s)\mathrm{d}\hat{\lambda} \\ &+ \int_{\hat{\lambda}_s}^{\hat{\lambda}_o}\int_{\hat{\lambda}_s}^{\hat{\lambda}} \Big[A^{(1)\alpha}_{\beta}(\hat{\lambda}')(\hat{\lambda}' - \hat{\lambda}_s) + C^{(1)\alpha}_{\beta}(\hat{\lambda}')\Big]\mathrm{d}\hat{\lambda}'\mathrm{d}\hat{\lambda},\end{aligned} \tag{E.4}$$

where

$$C^{(1)\alpha}_{\beta} := -\Gamma^{(1)\alpha}_{\mu\beta}\,\hat{\bar{k}}^{\mu}, \qquad A^{(1)\alpha}_{\beta} := R^{(1)\alpha}_{\rho\mu\beta}\,\hat{\bar{k}}^{\rho}\hat{\bar{k}}^{\mu}, \tag{E.5}$$

and

$$\hat{\bar{k}}^{\mu} = \tilde{\omega}(1, \vec{n}). \tag{E.6}$$

Calculating these for the metric (5.15), we obtain:

$$C^{(1)0}_{0} = -(\dot{\xi} + \vec{\nabla}\xi.\vec{n}), \qquad C^{(1)j}_{0} = C^{(1)0}_{j} = -\xi_{,j}, \qquad C^{(1)j}_{k} = 0, \tag{E.7}$$

and

$$A_0^{(1)0} = \xi_{,ij} n^i n^j, \qquad A_0^{(1)k} = -A_k^{(1)0} = \xi_{,ki} n^i, \qquad A_l^{(1)k} = -\xi_{,kl}. \tag{E.8}$$

The photon direction vectors (4.34) can be split into background plus perturbation

$$n_s^\mu = \bar{n}_s^\mu + n_s^{(1)\mu}, \qquad n_o^\mu = \bar{n}_o^\mu + n_o^{(1)\mu}, \tag{E.9}$$

where

$$\bar{n}_s^\mu = \bar{n}_o^\mu = (0, \vec{n}), \tag{E.10}$$

and

$$n_s^{(1)\mu} = (0, \hat{\vec{\ell}}_s^{(1)}) \approx \vec{0}, \qquad n_o^{(1)\mu} = (0, \hat{\vec{\ell}}_o^{(1)}). \tag{E.11}$$

Also the projectors (4.33) can be split into background plus perturbation:

$$\bar{P}_{s\nu}^\mu = \delta_\nu^\mu + \hat{\bar{U}}_s^\mu \hat{\bar{U}}_{s\nu} - \bar{n}_s^\mu \bar{n}_{s\nu}, \qquad \bar{P}_{o\nu}^\mu = \delta_\nu^\mu + \hat{\bar{U}}_o^\mu \hat{\bar{U}}_{o\nu} - \bar{n}_o^\mu \bar{n}_{o\nu}, \tag{E.12}$$

and

$$P_{s\nu}^{(1)\mu} = \hat{\bar{U}}_s^\mu \hat{U}_{s\nu}^{(1)} + \hat{U}_s^{(1)\mu} \hat{\bar{U}}_{s\nu}, \qquad P_{o\nu}^{(1)\mu} = \hat{\bar{U}}_o^\mu \hat{U}_{o\nu}^{(1)} + \hat{U}_o^{(1)\mu} \hat{\bar{U}}_{o\nu} - \bar{n}_o^\mu n_{o\nu}^{(1)} - n_o^{(1)\mu} \bar{n}_{o\nu}. \tag{E.13}$$

In terms of the metric (5.15), the projectors are given by

$$\bar{P}_0^0 = \bar{P}_i^0 = \bar{P}_0^i = 0, \qquad \bar{P}_j^i = \delta_j^i - n^i n_j, \tag{E.14}$$

and

$$P_{s\nu}^{(1)\mu} = 0, \qquad P_{o0}^{(1)0} = P_{oi}^{(1)0} = P_{o0}^{(1)i} = 0, \qquad P_{oj}^{(1)i} = -n^i \hat{\ell}_{oj}^{(1)} - \hat{\ell}_o^{(1)i} n_j. \tag{E.15}$$

The projected Jacobi map is given by

$$\begin{aligned} \hat{J} &= P_o \hat{\mathcal{J}} P_s \\ &= \bar{P}_o \hat{\bar{\mathcal{J}}} \bar{P}_s + P_o^{(1)} \hat{\bar{\mathcal{J}}} \bar{P}_s + \bar{P}_o \hat{\mathcal{J}}^{(1)} \bar{P}_s \\ &= \hat{\bar{J}} + \hat{J}^{(1)}. \end{aligned} \tag{E.16}$$

The background projected Jacobi map is the same as in Minkowski space

$$\hat{\bar{J}}^0{}_0 = \hat{\bar{J}}^0{}_i = \hat{\bar{J}}^i{}_0 = 0, \tag{E.17}$$

$$\hat{\bar{J}}^i{}_j = (\hat{\lambda}_o - \hat{\lambda}_s)(\delta^i{}_j - n^i n_j). \tag{E.18}$$

After a long but straightforward calculation, the full projected Jacobi map can be shown to be equal to

$$\hat{J}^0{}_0 = \hat{J}^0{}_i = \hat{J}^i{}_0 = 0, \tag{E.19}$$

$$\hat{J}^i{}_j = (\hat{\lambda}_o - \hat{\lambda}_s)\Big[\delta^i{}_j - n^i n_j\Big]$$
$$- \int_{\hat{\lambda}_s}^{\hat{\lambda}_o}\int_{\hat{\lambda}_s}^{\hat{\lambda}}\Big[\xi_{,ij} - \xi_{,il}n^{,l}n_{,j} - n^i n_k \xi^{,k}_{,j} + n^i n_k \xi_{,kl} n^l n_j\Big](\hat{\lambda}' - \hat{\lambda}_s)\mathrm{d}\hat{\lambda}'\mathrm{d}\hat{\lambda}$$
$$+ (\hat{\lambda}_o - \hat{\lambda}_s)\Big[-n^i\hat{\ell}^{(1)}_{oj} + n^i n_j(\hat{\vec{\ell}}^{(1)}_o \cdot \vec{n})\Big]. \tag{E.20}$$

Solving the characteristic equation, we find the two non-vanishing eigenvalues of this 3×3 spatial matrix. Their product gives the determinant of the Jacobi map (to first order)

$$(\det \hat{J})^{\frac{1}{2}} = (\hat{\lambda}_o - \hat{\lambda}_s) - \frac{1}{2}\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}\int_{\hat{\lambda}_s}^{\hat{\lambda}}(\hat{\lambda}' - \hat{\lambda}_s)\xi_{,ij}(\delta_{ij} - n_i n_j)\mathrm{d}\hat{\lambda}'\mathrm{d}\hat{\lambda}. \tag{E.21}$$

We can rewrite the double integral as a single integral by using the identity

$$\int_{\eta_s}^{\eta_o}\int_{\eta_s}^{\eta} g(\eta')\mathrm{d}\eta'\mathrm{d}\eta = \int_{\eta_s}^{\eta_o}(\eta_o - \eta)g(\eta)\mathrm{d}\eta. \tag{E.22}$$

We then obtain

$$(\det \hat{J})^{\frac{1}{2}} = (\hat{\lambda}_o - \hat{\lambda}_s)\left\{1 - \frac{1}{2}\frac{1}{\hat{\lambda}_o - \hat{\lambda}_s}\int_{\hat{\lambda}_s}^{\hat{\lambda}_o}(\hat{\lambda}_o - \hat{\lambda})(\nabla^2\xi - n^i n^j \xi_{,ij})(\hat{\lambda} - \hat{\lambda}_s)\mathrm{d}\hat{\lambda}\right\}. \tag{E.23}$$

Since $(\hat{\lambda}_o - \hat{\lambda}_s)$ is positive and ξ is by assumption extremely small we must have that $(\det \hat{J})^{\frac{1}{2}} = |\det \hat{J}|^{\frac{1}{2}}$. If one desires, one can now easily obtain the Jacobi map and Jacobi determinant in the full perturbed FLRW spacetime by using the relations

$$J = f_o\,\hat{J}, \qquad \text{and} \qquad \det J = f_o^2\,\det \hat{J}. \tag{E.24}$$

Bibliography

[1] G. F. R. Ellis, R. Maartens, and M. A. H. MacCallum, *Relativistic Cosmology*. England: Cambridge University Press, 2012.

[2] S. M. Carroll, *Spacetime and Geometry*. Cambridge University Press, 2019.

[3] M. Visser, "Cosmography: Cosmology without the Einstein equations," *Gen. Rel. Grav.*, vol. 37, pp. 1541–1548, 2005.

[4] H. Kodama and M. Sasaki, "Cosmological Perturbation Theory," *Prog. Theor. Phys. Suppl.*, vol. 78, pp. 1–166, 1984.

[5] N. Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters," 7 2018.

[6] A. G. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astron. J.*, vol. 116, pp. 1009–1038, 1998.

[7] B.-L. Young, "A survey of dark matter and related topics in cosmology," *Front. Phys. 12(2), 121201*, 2016.

[8] P. Salucci, "The distribution of dark matter in galaxies," *Astron. Astrophys. Rev.*, vol. 27, no. 1, p. 2, 2019.

[9] S. M. Carroll, "Dark energy and the preposterous universe," 7 2001.

[10] E. Bianchi and C. Rovelli, "Why all these prejudices against a constant?," 2 2010.

[11] D. H. Weinberg, J. S. Bullock, F. Governato, R. Kuzio de Naray, and A. H. G. Peter, "Cold dark matter: controversies on small scales," *Proc. Nat. Acad. Sci.*, vol. 112, pp. 12249–12255, 2015.

[12] J. L. Bernal, L. Verde, and A. G. Riess, "The trouble with H_0," *JCAP*, vol. 10, p. 019, 2016.

[13] R. R. Caldwell, R. Dave, and P. J. Steinhardt, "Cosmological imprint of an energy component with general equation of state," *Phys. Rev. Lett.*, vol. 80, pp. 1582–1585, 1998.

[14] M. Chevallier and D. Polarski, "Accelerating universes with scaling dark matter," *Int. J. Mod. Phys. D*, vol. 10, pp. 213–224, 2001.

[15] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, "Modified Gravity and Cosmology," *Phys. Rept.*, vol. 513, pp. 1–189, 2012.

[16] A. De Felice and S. Tsujikawa, "Cosmology of a covariant Galileon field," *Phys. Rev. Lett.*, vol. 105, p. 111301, 2010.

[17] L. Hui, J. P. Ostriker, S. Tremaine, and E. Witten, "Ultralight scalars as cosmological dark matter," *Phys. Rev. D*, vol. 95, no. 4, p. 043541, 2017.

[18] T. Kobayashi, R. Murgia, A. De Simone, V. Irsic, and M. Viel, "Lyman-α constraints on ultralight scalar dark matter: Implications for the early and late universe," *Phys. Rev. D*, vol. 96, no. 12, p. 123514, 2017.

[19] N. Bar, D. Blas, K. Blum, and S. Sibiryakov, "Galactic rotation curves versus ultralight dark matter: Implications of the soliton-host halo relation," *Phys. Rev. D*, vol. 98, no. 8, p. 083027, 2018.

[20] C. Boehmer and T. Harko, "Can dark matter be a Bose-Einstein condensate?," *JCAP*, vol. 06, p. 025, 2007.

[21] T. Harko, "Bose-Einstein condensation of dark matter solves the core/cusp problem," *JCAP*, vol. 05, p. 022, 2011.

[22] M. Milgrom, "A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis," *Astrophysical Journal*, vol. 270, pp. 365–370, 1983.

[23] J.-P. Bruneton, S. Liberati, L. Sindoni, and B. Famaey, "Reconciling MOND and dark matter?," *JCAP*, vol. 03, p. 021, 2009.

[24] D. Bettoni, S. Liberati, and L. Sindoni, "Extended LCDM: generalized non-minimal coupling for dark matter fluids," *JCAP*, vol. 11, p. 007, 2011.

[25] J. D. Bekenstein, "The Relation between physical and gravitational geometry," *Phys. Rev. D*, vol. 48, pp. 3641–3647, 1993.

[26] K. Migkas, G. Schellenberger, T. Reiprich, F. Pacaud, M. Ramos-Ceja, and L. Lovisari, "Probing cosmic isotropy with a new X-ray galaxy cluster sample through the $L_{\rm X} - T$ scaling relation," *Astron. Astrophys.*, vol. 636, p. A15, 2020.

[27] J. Colin, R. Mohayaee, M. Rameez, and S. Sarkar, "Evidence for anisotropy of cosmic acceleration," *Astron. Astrophys.*, vol. 631, p. L13, 2019.

[28] S. Appleby, A. Shafieloo, and A. Johnson, "PROBING BULK FLOW WITH NEARBY SNe ia DATA," *The Astrophysical Journal*, vol. 801, p. 76, mar 2015.

[29] B. Abbott *et al.*, "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral," *Phys. Rev. Lett.*, vol. 119, no. 16, p. 161101, 2017.

[30] T. Baker, E. Bellini, P. Ferreira, M. Lagos, J. Noller, and I. Sawicki, "Strong constraints on cosmological gravity from GW170817 and GRB 170817A," *Phys. Rev. Lett.*, vol. 119, no. 25, p. 251301, 2017.

[31] M. Sasaki, "The Magnitude - Redshift relation in a perturbed Friedmann universe," *Mon. Not. Roy. Astron. Soc.*, vol. 228, pp. 653–669, 1987.

[32] C. Bonvin, R. Durrer, and M. A. Gasparini, "Fluctuations of the luminosity distance," *Phys. Rev. D*, vol. 73, p. 023523, Jan 2006.

[33] E. Barausse, S. Matarrese, and A. Riotto, "The Effect of inhomogeneities on the luminosity distance-redshift relation: Is dark energy necessary in a perturbed Universe?," *Phys. Rev.*, vol. D71, p. 063537, 2005.

[34] O. Umeh, C. Clarkson, and R. Maartens, "Nonlinear relativistic corrections to cosmological distances, redshift and gravitational lensing magnification: I. Key results," *Class. Quant. Grav.*, vol. 31, p. 202001, 2014.

[35] J. Yoo and F. Scaccabarozzi, "Unified Treatment of the Luminosity Distance in Cosmology," *JCAP*, vol. 1609, no. 09, p. 046, 2016.

[36] G. Fanizza, M. Gasperini, G. Marozzi, and G. Veneziano, "An exact Jacobi map in the geodesic light-cone gauge," *JCAP*, vol. 1311, p. 019, 2013.

[37] I. Ben-Dayan, G. Marozzi, F. Nugier, and G. Veneziano, "The second-order luminosity-redshift relation in a generic inhomogeneous cosmology," *JCAP*, vol. 1211, p. 045, 2012.

[38] E. L. Lokas and G. A. Mamon, "Properties of spherical galaxies and clusters with an nfw density profile," *Mon. Not. Roy. Astron. Soc.*, vol. 321, p. 155, 2001.

[39] P. Salucci and A. Burkert, "Dark matter scaling relations," *Astrophys. J. Lett.*, vol. 537, pp. L9–L12, 2000.

[40] A. W. Graham, D. Merritt, B. Moore, J. Diemand, and B. Terzic, "Empirical models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models," *Astron. J.*, vol. 132, pp. 2685–2700, 2006.

[41] C. J. Pethick and H. Smith, *Bose–Einstein Condensation in Dilute Gases*. Cambridge University Press, 2 ed., 2008.

[42] D. Bettoni, M. Colombo, and S. Liberati, "Dark matter as a Bose-Einstein Condensate: the relativistic non-minimally coupled case," *JCAP*, vol. 02, p. 004, 2014.

[43] C. Barcelo, S. Liberati, and M. Visser, "Analogue gravity," *Living Rev. Rel.*, vol. 8, p. 12, 2005.

[44] D. Ivanov and S. Liberati, "Testing Non-minimally Coupled BEC Dark Matter with Gravitational Waves," *JCAP*, vol. 07, p. 065, 2020.

[45] T. Kobayashi, M. Yamaguchi, and J. Yokoyama, "Generalized G-inflation: Inflation with the most general second-order field equations," *Prog. Theor. Phys.*, vol. 126, pp. 511–529, 2011.

[46] D. Bettoni, J. M. Ezquiaga, K. Hinterbichler, and M. Zumalacárregui, "Speed of Gravitational Waves and the Fate of Scalar-Tensor Gravity," *Phys. Rev. D*, vol. 95, no. 8, p. 084029, 2017.

[47] E. E. Flanagan and S. A. Hughes, "The Basics of gravitational wave theory," *New J. Phys.*, vol. 7, p. 204, 2005.

[48] S. Liberati, S. Sonego, and M. Visser, "Scharnhorst effect at oblique incidence," *Phys. Rev. D*, vol. 63, p. 085003, 2001.

[49] S. Liberati, S. Sonego, and M. Visser, "Faster than c signals, special relativity, and causality," *Annals Phys.*, vol. 298, pp. 167–185, 2002.

[50] H. Mo, F. C. van den Bosch, and S. White, *Galaxy Formation and Evolution*. 2010.

[51] A. Lapi, P. Salucci, and L. Danese, "Precision Scaling Relations for Disk Galaxies in the Local Universe," *The Astrophysical Journal*, vol. 859, p. 2, May 2018.

[52] K.-H. Huang, S. M. Fall, H. C. Ferguson, A. van der Wel, N. Grogin, A. Koekemoer, S.-K. Lee, P. G. Perez-Gonzalez, and S. Wuyts, "Relations between the Sizes of Galaxies and Their Dark Matter Halos at Redshifts 0 < z < 3," *The Astrophysical Journal*, vol. 838, p. 6, Mar. 2017.

[53] S. Boran, S. Desai, E. Kahya, and R. Woodard, "GW170817 Falsifies Dark Matter Emulators," *Phys. Rev. D*, vol. 97, no. 4, p. 041501, 2018.

[54] E. D. Schiappacasse and M. P. Hertzberg, "Analysis of Dark Matter Axion Clumps with Spherical Symmetry," *JCAP*, vol. 01, p. 037, 2018. [Erratum: JCAP 03, E01 (2018)].

[55] O. Erken, P. Sikivie, H. Tam, and Q. Yang, "Cosmic axion thermalization," *Phys. Rev. D*, vol. 85, p. 063520, 2012.

[56] D. Ivanov, S. Liberati, M. Viel, and M. Visser, "Non-perturbative results for the luminosity and area distances," *JCAP*, vol. 06, p. 040, 2018.

[57] D. Ivanov, S. Liberati, M. Viel, and M. Visser, "Perturbative treatment of the luminosity distance," *Phys. Rev. D*, vol. 98, no. 6, p. 063505, 2018.

[58] M. Visser, "Jerk and the cosmological equation of state," *Class. Quant. Grav.*, vol. 21, pp. 2603–2616, 2004.

[59] E. Belgacem *et al.*, "Testing modified gravity at cosmological distances with LISA standard sirens," *JCAP*, vol. 07, p. 024, 2019.

[60] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, "Modified gravitational-wave propagation and standard sirens," *Phys. Rev. D*, vol. 98, no. 2, p. 023510, 2018.

[61] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, "Gravitational-wave luminosity distance in modified gravity theories," *Phys. Rev. D*, vol. 97, no. 10, p. 104066, 2018.

[62] B. A. Bassett and M. Kunz, "Cosmic distance-duality as a probe of exotic physics and acceleration," *Phys. Rev. D*, vol. 69, p. 101305, 2004.

[63] M. Kunz and B. A. Bassett, "A tale of two distances," in *39th Rencontres de Moriond Workshop on Exploring the Universe: Contents and Structures of the Universe*, 6 2004.

[64] R. Nair, S. Jhingan, and D. Jain, "Testing the consistency between cosmological measurements of distance and age," *Phys. Lett. B*, vol. 745, pp. 64–68, 2015.

[65] M. Visser, "Conformally Friedmann-Lemaitre-Robertson-Walker cosmologies," *Class. Quant. Grav.*, vol. 32, no. 13, p. 135007, 2015.

[66] D. W. Hogg, "Distance measures in cosmology," 5 1999.

[67] P. H. F. Reimberg and L. R. Abramo, "The Jacobi map for gravitational lensing: the role of the exponential map," *Class. Quant. Grav.*, vol. 30, p. 065020, 2013.

[68] O. Umeh, C. Clarkson, and R. Maartens, "Nonlinear relativistic corrections to cosmological distances, redshift and gravitational lensing magnification. II - Derivation," *Class. Quant. Grav.*, vol. 31, p. 205001, 2014.

[69] E. Di Dio, F. Montanari, A. Raccanelli, R. Durrer, M. Kamionkowski, and J. Lesgourgues, "Curvature constraints from Large Scale Structure," *JCAP*, vol. 1606, no. 06, p. 013, 2016.

[70] J. H. Van Vleck, "The correspondence principle in the statistical interpretation of quantum mechanics," *Proceedings of the National Academy of Sciences*, vol. 14, no. 2, pp. 178–188, 1928.

[71] M. Visser, "van Vleck determinants: Traversable wormhole space-time," *Phys. Rev.*, vol. D49, pp. 3963–3980, 1994.

[72] M. Visser, *Lorentzian wormholes: From Einstein to Hawking.* AIP Press, now Springer, 1995.

[73] M. Visser, "van Vleck determinants: Geodesic focusing and defocusing in Lorentzian space-times," *Phys. Rev.*, vol. D47, pp. 2395–2402, 1993.

[74] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time.* Cambridge Monographs on Mathematical Physics, Cambridge University Press, 1973.

[75] R. M. Wald, *General Relativity.* Chicago, USA: Chicago Univ. Pr., 1984.

[76] G. Gubitosi, F. Piazza, and F. Vernizzi, "The Effective Field Theory of Dark Energy," *JCAP*, vol. 1302, p. 032, 2013. [JCAP1302,032(2013)].

[77] N. Zwane, N. Afshordi, and R. D. Sorkin, "Cosmological Tests of Everpresent Λ," 2017.

[78] C. Armendariz-Picon, V. F. Mukhanov, and P. J. Steinhardt, "A Dynamical solution to the problem of a small cosmological constant and late time cosmic acceleration," *Phys. Rev. Lett.*, vol. 85, pp. 4438–4441, 2000.

[79] M. Li, "A Model of holographic dark energy," *Phys. Lett.*, vol. B603, p. 1, 2004.

[80] T. Buchert, "Dark Energy from Structure: A Status Report," *Gen. Rel. Grav.*, vol. 40, pp. 467–527, 2008.

[81] S. Rasanen, "Accelerated expansion from structure formation," *JCAP*, vol. 11, p. 003, 2006.

[82] S. Rasanen, "Light propagation in statistically homogeneous and isotropic universes with general matter content," *JCAP*, vol. 03, p. 018, 2010.

[83] C. Clarkson, G. Ellis, J. Larena, and O. Umeh, "Does the growth of structure affect our dynamical models of the universe? The averaging, backreaction and fitting problems in cosmology," *Rept. Prog. Phys.*, vol. 74, p. 112901, 2011.

[84] N. Meures and M. Bruni, "Redshift and distances in a ΛCDM cosmology with non-linear inhomogeneities," *Monthly Notices of the Royal Astronomical Society*, vol. 419, pp. 1937–1950, Jan. 2012.

[85] K. Enqvist, M. Mattsson, and G. Rigopoulos, "Supernovae data and perturbative deviation from homogeneity," *JCAP*, vol. 2009, p. 022, Sept. 2009.

[86] R. K. Sachs and A. M. Wolfe, "Perturbations of a cosmological model and angular variations of the microwave background," *Astrophys. J.*, vol. 147, pp. 73–90, 1967. [Gen. Rel. Grav.39,1929(2007)].

[87] Z. Sun and F. Wang, "Testing the anisotropy of cosmic acceleration from Pantheon supernovae sample," *Mon. Not. Roy. Astron. Soc.*, vol. 478, no. 4, pp. 5153–5158, 2018.

[88] U. Andrade, C. A. Bengaly, B. Santos, and J. S. Alcaniz, "A Model-independent Test of Cosmic Isotropy with Low-z Pantheon Supernovae," *Astrophys. J.*, vol. 865, no. 2, p. 119, 2018.

[89] Y. Akrami *et al.*, "Planck 2018 results. VII. Isotropy and Statistics of the CMB," 6 2019.

[90] M. Betoule *et al.*, "Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples," *Astron. Astrophys.*, vol. 568, p. A22, 2014.

[91] P. e. a. Amaro-Seoane, "Laser Interferometer Space Antenna," 2017.

[92] L. Amendola *et al.*, "Cosmology and fundamental physics with the Euclid satellite," *Living Rev. Rel.*, vol. 21, no. 1, p. 2, 2018.

[93] C. Inserra *et al.*, "Euclid: Superluminous supernovae in the Deep Survey," *Astron. Astrophys.*, vol. 609, p. A83, 2018.

[94] v. Z. Ivezić *et al.*, "LSST: from Science Drivers to Reference Design and Anticipated Data Products," *Astrophys. J.*, vol. 873, no. 2, p. 111, 2019.