University of Cape Town

Master's Thesis
National Astrophysics and Space Science programme

A $f(R)$esh Take on Gravity

Jess Worsley

supervised by
Prof. Peter Dunsby

2024

Acknowledgements

I would like to express my sincere gratitude to the donors of the Washkansky, Manuel, and Luby Scholarship Award, the National Astrophysics and Space Science Programme (NASSP), and the University of Cape Town (UCT) for their financial support, without which this work would not have been possible.

I am especially grateful to my supervisor, Prof. Peter Dunsby, for his invaluable guidance, patience, and encouragement throughout this research. I began my master's with limited knowledge of the field, and his insights and support have been instrumental in shaping my understanding of cosmology and modified gravity.

I extend my heartfelt thanks to my family for their unwavering support and belief in me, even when I doubted myself, and for proofreading 150 pages of mathematics. I am also deeply appreciative of my friends and colleagues in the UCT Cosmology and Gravity group for their camaraderie, insightful discussions, and shared moments of frustration and celebration. Their support has made this journey all the more rewarding.

Abstract

This thesis explores perturbations in $f(R)$ modified gravity theories, with a focus on understanding the growth of large-scale structures in the Universe. Modified gravity models, particularly $f(R)$ theories, offer an alternative explanation for cosmic acceleration without invoking dark energy. The work is divided into four key areas: foundational methods in cosmology, structure formation, perturbation theory, and the application of these methods in modified gravity models. By analysing the evolution of cosmological perturbations in these models, this research aims to provide new insights into their viability and consistency with observational data. Particular attention is given to the development of a bottom-up approach, allowing the reconstruction of gravity theories from observational constraints and bypassing the limitations of traditional model-specific methods.

Contents

Conventions

Metric signature	$[-,+,+,+]$
Natural units[1]	$\hbar = c = 8\pi G = 1$
Latin indices $(1,2,3)$	Spatial components
Greek indices $(0,1,2,3)$	Spacetime components
∇_μ	Covariant derivative
$\widetilde{\nabla}_\mu \equiv h^\nu{}_\mu \nabla_\nu$	Orthogonally projected covariant derivative
$\Box \equiv g^{\mu\nu}\nabla_\mu\nabla_\nu$	d'Alembertian operator
$\nabla_\mu f = \partial_\mu f$	Covariant derivative of scalar function
$\nabla_\nu X_\mu = \partial_\nu X_\mu + \Gamma^\sigma{}_{\mu\nu} X_\sigma$	Covariant derivative of covariant tensor
$\nabla_\nu X^\mu = \partial_\nu X^\mu + \Gamma^\mu{}_{\nu\sigma} X^\sigma$	Covariant derivative of contravariant tensor
$X_{(\mu\nu)} = \frac{1}{2}(X_{\mu\nu} + X_{\nu\mu})$	Symmetricity
$X_{[\mu\nu]} = \frac{1}{2}(X_{\mu\nu} - X_{\nu\mu})$	Anti-symmetricity
$X_{\langle\mu\nu\rangle} = X_{(\mu\nu)} - \frac{1}{4} g_{\mu\nu} X^\sigma{}_\sigma$	Trace-free symmetricity

Abbreviations

2dFGRS	Two-degree-Field Galaxy Redshift Survey
2MASS	Two Micron All-Sky Survey
AGN	Active Galactic Nucleus
APM	Automated Plate Measurement
BAO	Baryon Acoustic Oscillations
CDM	Cold Dark Matter
CMB	Cosmic Microwave Background
DESI	Dark Energy Spectroscopic Instrument
FLRW	Friedmann-Lemaître-Robertson-Walker
GR	General Relativity
HDM	Hot Dark Matter
JWST	James Webb Space Telescope
LISA	Laser Interferometer Space Antenna
NRAO	National Radio Astronomy Observatory
NVSS	NRAO VLA Sky Survey
QS	Quasi-static
SDSS	Sloan Digital Sky Survey
SNIa	Supernova type Ia
TRGB	Tip of the red giant branch
VLA	Very Large Array

[1] $8\pi G = 1$ is used from ch. 2 onwards; $c = 1$ from ch. 3 onwards.

Introduction

Of all the fundamental forces we know, gravity is the weakest [1]. And yet, it rules the Universe. With a light touch, it compelled a uniform cosmos to diversify into an intricate network spanning billions of light years. It spins galaxy clusters – each 10^{15} solar masses – into a colossal web; tethers together the 100 billion stars in the Milky Way; and engulfs light itself at the centre of the galaxy. 30,000 light years away from this black hole, in the Orion Arm, gravity governs the orbits of planets around a small, unassuming star; and the orbit of a moon around a small, unassuming planet. It is here that for thousands of years, humans have wondered "why do things fall?".

The Greek philosopher Aristotle thought it quite simple: all objects have a natural tendency to move towards the centre of the Universe [2]. Since this point, of course, coincided with the centre of the Earth, all objects must fall towards the ground. It is certainly not what we would today call "rigorous". It was not until the 16^{th} century that Galileo would attempt a truly scientific approach, leading him to realise that the velocity of an object in free fall is independent of its mass [3]. Isaac Newton expanded on these ideas in his studies on the trajectories of moving objects, which led him to the first theory on artificial satellites. By considering the moon as continuously falling towards Earth, but never colliding, he was able to calculate the acceleration due to Earth's gravity on its satellite. Comparing this to the free fall of objects on the ground, he noticed a simple relation: the force of attraction between two objects is proportional to the inverse square of the distance between them. But why did objects appear attracted towards the earth, but not each other? Newton knew then that the masses of the objects also influenced the gravitational attraction between them, or more formally: two objects attract each other with a force proportional to the product of their masses. More massive objects like the Earth therefore experience less acceleration towards a falling object, despite both experiencing the same gravitational force. Combined, these relations yielded the famous equation that we still use today:

$$F = \frac{GM_1M_2}{R^2}\,. \tag{1}$$

Newton had to lay the foundations for a new branch of mathematics before he could even publish the complete mathematical formulation of his theory in 1687: *Philosophiae Naturalis Principia Mathematica* [4]. Known now as calculus, these methods became fundamental to scientists across all fields.

The success of Newton's gravitational laws ushered in a new era in physics, but it was not until Einstein's paper in 1915 that the underlying nature of the gravitational interaction was coherently described: the General Theory of Relativity [5]. This supplanted the old Newtonian point of view that large masses produce forces that cause objects to move along curved trajectories. From the Einsteinian perspective, the space around large masses becomes curved itself, and objects move along the 'straightest' line in curved space, a geodesic. This interpretation of gravity as a curvature of spacetime – rather than a force produced by mass – leads to slightly different results than those predicted by Newtonian gravity. Observations soon showed that

Einsteinian gravity solved the mystery of Mercury's orbital precession[2] [6], and the increased deflection angle of the sun – which was famously and somewhat fortuitously confirmed during the 1919 solar eclipse. Later, general relativity was used to confirm gravitational redshift [7, 8], cosmological expansion [9], and time dilation [10]. Most recently in 2015, direct observations verified the existence of gravitational waves from a black hole merger event [11]. These ripples in spacetime were predicted by Einstein almost a hundred years earlier, definitively proving that gravity is not an instantaneous action as Newton thought.

While Einstein initially adopted a static universe model, Alexander Friedmann showed in 1922 that Einstein's field equations allowed for dynamic universes with different geometries, depending on the curvature [12, 13]. Hubble corroborated this theory in 1929 when he discovered that galaxies are drifting apart as the Universe expands [9]. Another astonishing discovery in 1998 by Perlmutter, Riess and Schmidt confirmed that the Einstein static model was inconsistent with observations [14], supporting the theory that the Universe is undergoing *accelerated* expansion. The cosmological constant Λ in the Einstein field equations – initially placed to preserve a static universe – was now responsible for a groundbreaking phenomenon which its creator had not predicted.

In the last few decades, advances in the study of cosmology and quantum theory have uncovered a number of obstacles to general relativity (GR). The presence of the Big Bang singularity at the beginning of the Universe [15], as well as the flatness, horizon, and monopole problems [16, 17, 18], signals that the standard cosmological model – based on GR and the Standard Model of particle physics [19] – is insufficient for describing the Universe at extreme scenarios, such as during the Planck era or at the centres of black holes. Nor can it account for the rapid exponential expansion in the early universe. As a classical theory, it is inapplicable at quantum scales [20]. This is particularly troublesome since it is believed that primordial quantum fluctuations are responsible for sowing the seeds of density perturbations and thus large-scale structure [21]. As such, one of the primary objectives of many cosmologists is a unified theory of quantum gravity.

Meanwhile, headway in astronomical observations unveiled several other glaring issues: galactic rotation curves do not display the velocity falloff predicted by GR, suggesting that galaxies contain more matter than observed; gravitational lensing effects from distant galaxy clusters suggest more mass than we observe in those clusters; and large structures in the early universe formed more quickly than expected for baryonic matter, which is slower to collapse due to interactions via radiation pressure and viscosity. These all point to a problem of 'missing mass'. In 1933, Fritz Zwicky observed that the galaxies in the Coma cluster were moving too fast for the mass of the cluster, indicating that there was some unseen matter present with enough gravity to stop the galaxies from spinning out. He called this unseen mass *dark matter*[3] [22]. In the 1970s, Vera Rubin and Kent Ford observed the same phenomenon in some sixty galaxies, from which they calculated that galaxies must contain approximately ten times more mass than observed [23]. We now know that approximately 85% of the matter in the observable universe is virtually undetectable through direct means [24, 25].

Some of this mass is dark matter, but in a universe comprised of only baryonic and dark matter, the gravitational pull of the matter should decelerate the expansion of the Universe over time. Since this is not the case, we also require some as-yet undiscovered form of energy that coun-

[2] Astronomers noticed that Mercury's orbit exhibited an unexplained precession of its perihelion. Newtonian mechanics accounted for most of the precession due to gravitational interactions with other planets, but an additional unexplained 43 arcseconds per century remained.

[3] Technically he called it *dunkle Materie*.

teracts gravity and drives the accelerated expansion. Additionally, observations of the cosmic microwave background (CMB) indicate that the Universe is geometrically flat, i.e. that its total energy density is equal to the critical density [24]. Ordinary and dark matter only make up 32% of the required density. The mysterious missing energy component is – in keeping with tradition – known as *dark energy*. In our current model, according to Planck data, the approximate mass-energy content of the Universe is: 5% ordinary matter, 27% dark matter, and 68% dark energy. The dark sector therefore comprises 95% of the total mass-energy density of the Universe, yet its nature remains unknown.

Despite this, general relativity remains a compelling theory, and the search for dark matter and dark energy continues. Meanwhile, while some theorists pursue a definitive quantum theory of gravity, others have posited various alternative theories of gravity that attempt to formulate a semi-classical theory of relativity that replicates the success – and resolves the issues – of GR; these are known as *extended theories of gravity* (ETGs) [26], and involve corrections or additions to Einstein's theory by introducing higher-order curvature invariants and scalar fields to the dynamics arising from some effective quantum gravity action. ETGs also attempt to fully incorporate Mach's principle into GR, which allows explicitly anti-Machian solutions such as Gödel's universe. Mach's principle implies that the gravitational coupling at a spacetime point is dependent on the surrounding matter and thus a function of spacetime location, i.e., a scalar field [27]. Brans-Dicke theory, the first comprehensive alternative to GR, better aligned with Mach's principle by allowing a gravitational scalar field to couple non-minimally to the geometry [28]. This is also achieved by schemes such as superstring, supergravity, or Grand Unified Theories (GUTs). One of the great benefits of these extended gravity modifications is that they exhibit inflationary behaviour, allowing for inflationary epochs – related to the dynamics of a specific scalar field – not predicted in the standard model.

The majority of the vast amount of data currently points to a cosmological model known as ΛCDM, which consists of a cosmological constant Λ combined with cold dark matter – dark matter with a vanishing equation of state – and baryonic matter. Despite being well supported by data, ΛCDM fails to explain why the inferred value of Λ is 120 orders of magnitude lower than the value predicted by particle physics; this is the *coincidence problem*. Alternative models known as *quintessence* attempt to circumvent this by replacing the cosmological constant with a slow-rolling scalar field ϕ, but have not yet found success. In these schemes, the dark energy and matter densities evolve differently, yet align in the current epoch, leading to another coincidence problem. The origin of the quintessence scalar field presents another puzzle. Unified Dark Energy (UDE) models propose single fluids (e.g. Chaplygin gas [29]) with properties of both dark matter and dark energy to address the coincidence problem [30]. A single cosmic fluid characterised by an equation of state that causes it to act like dark matter at high densities and dark energy at low densities neatly solves the coincidence problem, at least phenomenologically.

Quantum gravity theories, on the other hand, have taken on the daunting task of creating a unified framework for all fundamental forces, by describing the gravitational fields using the tools of quantum mechanics and representing them as linear operators on a Hilbert space. This comes with its own set of challenges. Einstein's theory does not lend itself to a quantum formulation based on a fixed Minkowski background. In GR, spacetime itself is a dynamic variable, meaning its geometry cannot be predetermined. It is necessary to first solve the equations of motion and construct a spacetime before even attempting to find solutions for certain scenarios [26, 31]. The complexity intensifies at the quantum level, where the uncertainty principle implies that particles do not follow well-defined paths in non-relativistic quantum mechanics. One can only calculate the probability of detecting a particle at a specific location and time. Similarly,

in quantum gravity, the evolution of an initial state fails to yield a specific spacetime [32]. This raises profound questions about fundamental concepts like causality, time, and black holes in a framework where spacetime itself is not well defined.

The final approach – the subject of this thesis – suggests modifying gravitational laws rather than introducing exotic cosmic fluids. These ETGs adjust the Einstein field equations by maximally coupling the geometry to some scalar field, or introducing higher-order derivatives of the metric. These are generally known as scalar-tensor theories and higher-order theories, respectively. Mathematically, such theories can often be reduced to forms resembling GR with additional scalar fields using techniques like Legendre or conformal transformations [33, 26]. The mathematical complexity of these modified gravity theories, particularly that of higher-order non-linear equations, complicates studies and limits their predictive power. Unlike GR, which has been rigorously tested in various regimes, modified gravity models lack similar empirical support, and often encounter issues of fine-tuning. Regardless, modified gravity theories are able to mimic the behaviour of dark matter and dark energy, explain the early inflationary epoch, and pass astrophysical viability tests [34, 35].

The particular modified gravity theory under consideration in this thesis is $f(R)$ gravity, in which the Ricci scalar in the Einstein-Hilbert action is replaced by a function of itself. Specifically, we will be working towards formulating a dynamical, model-independent, cosmography-based system for $f(R)$. This allows us to solve for any general function of R instead of first guessing its functional form, where previous studies required a specific $f(R)$ model. By expressing the scale factor and its temporal derivatives using observable cosmographic parameters, we can bypass previous limitations to closing the dynamical system [36]. We then study the formation of large-scale structure in $f(R)$ using covariant, gauge-invariant perturbation theory, in which we expand our first-order quantities using harmonic decomposition. This enables us to derive differential equations for the scalar quantities relevant to structure formation, as well as the growth function. Finally, we apply our dynamical system to ΛCDM cosmology by mimicking the expansion history at high redshifts, allowing us to plot the evolution of the growth function and growth factor at various Fourier wavenumbers. The growth function is a critical tool in cosmology because it describes the evolution of matter density perturbations over time, serving as a bridge between theory and observation [37, 38]. With it, we can study the observable consequences of perturbations in modified gravity models on cosmic structure formation, and how specific modifications to gravity influence linear and nonlinear perturbations.

The thesis is laid out as follows: In the first chapter, we introduce the necessary mathematical tools and techniques for studying relativity. We begin with classical, Newtonian spacetime before moving on to GR, in which we go over some basic differential geometry and the equations of motion. This allows us to derive the Einstein field equations and examine the origin of the cosmological constant. In chapter 2 we delve into cosmology, deriving the Robertson-Walker metric and Friedmann equations and examining some well-known cosmological models. We then get to inflation and its implications for the current standard model of the Universe. This includes the history of structure formation and the role of dark energy and dark matter in the Universe's evolution. In chapter 3 we examine this more rigorously. We discuss the gauge problem and various methods in perturbation theory, particularly the gauge-invariant approach. Using this method we derive the linearised propagation equations, to which we apply a harmonic decomposition. This leads us to Jeans' theory of instability and gravitational oscillations, and the study of fluctuations using the power spectrum. In the fourth chapter, we introduce $f(R)$ gravity, for which we formulate a model-independent dynamical system. We then use the tools from chapter 3 to study the structure growth of a general $f(R)$ theory and compare its evolution

to that of ΛCDM, utilising the exact solutions as well as the quasi-static and semi-quasi-static approximations, before concluding in the final chapter.

Chapter 1

Theoretical Foundations

This chapter lays the foundations for general relativity. We begin by reviewing the classical spacetime of Gallileo and Newton, the tenets of special relativity and how we can make use of the Lorentz transformations to preserve the laws of physics. In the next section we lay the groundwork for differential geometry, covering the covariant derivative, Christoffel symbols and geodesic, which allow us to define the Ricci, Riemann and Einstein tensors. This is followed by a discussion of the geodesic in curved spacetime and the equivalence principle before a summary of the Einstein field equations. The chapter culminates in an analysis of the equation of motion, energy conservation and the equation of state.

1.1 Classical spacetime

In classical Newtonian physics, space is described by three-dimensional, Euclidian geometry. Time acts as an extrinsic parameter, independent of the motion of any observer. Both space and time in this regime are absolute, and events adhere to a causal structure where simultaneous events all exist on a three-dimensional Euclidian space. Because the space is Euclidian, the distance between two neighbouring points is simply

$$\mathrm{d}\ell^2 = (\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2 = \sum_{ij} \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \equiv \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \,, \tag{1.1}$$

according to Pythagoras' theorem, where δ_{ij} is the Kronecker delta equal to 1 if $i = j$ and 0 otherwise, and the trajectory of any body has the parametric form $x^i(t)$. It is not necessary to use Cartesian coordinates as we have above; spherical or cylindrical coordinates may be more suited to a specific problem. It is therefore helpful to establish a method of expressing distances in any coordinate system. Suppose we have a coordinate system (y^i) which is related to the Cartesian system (x^j) via $x^j(y^i)$. The distance between two points in any coordinate system is expressed as

$$\mathrm{d}\ell^2 = g_{ij}(y^k)\mathrm{d}y^i\mathrm{d}y^j, \quad g_{ij}(y^k) = \frac{\partial x^m}{\partial y^i}\frac{\partial x^n}{\partial y^j}\delta_{mn} \,, \tag{1.2}$$

where g_{ij} is the metric of the new coordinate space. In Cartesian coordinates, g_{ij} simplifies to δ_{ij}. The spacetime metric is therefore essential in the study of relativity, allowing us to express observer-dependent coordinates as an invariant line element.

The principle of special relativity [39] asserts that all laws of physics maintain the same form for two observers moving at a constant velocity relative to each other. If the dynamical equations of a system remain invariant under such a change of coordinates, they must exhibit a corresponding symmetry. Newtonian mechanics adheres to this principle through Galilean relativity, where the coordinates are related by $\mathbf{x}'(t) = \mathbf{x}(t) - \mathbf{v}t$ and time is the same in both reference frames.

However, Maxwell's laws of electromagnetism do not retain their form under Galilean transformations. To account for this, the Maxwell equations would have to be valid only in the absolute reference frame. This meant that the speed of light in another inertial frame could potentially be measured as $c \pm v$. This notion of absolute time in Newtonian theory was challenged when light was shown to propagate at the same speed in all inertial reference frames. Einstein therefore excluded this idea in favour of reaffirming the tenets of special relativity we know today [40]:

Relativity: Nature obeys the same laws in all inertial, non-accelerating reference frames.

Universal light speed: The speed of light relative to any non-accelerating observer is c, regardless of the speed of the light's source relative to the observer.

Since these should also apply to Maxwell's equations, a new framework is required for changing the inertial frame: the Lorentz transformations. These are a set of linear transformations which preserve the spacetime interval between two events. If an initial observer notes an event at (t, x, y, z), a second observer in an inertial reference frame travelling at velocity v along the x-axis with respect to the first observer can identify the same event as (t', x', y', z') using

$$\begin{aligned} ct' &= \frac{ct - (v/c)x}{\sqrt{1 - v^2/c^2}}\,, \\ x' &= \frac{x - (v/c)ct}{\sqrt{1 - v^2/c^2}}\,, \\ y' &= y\,, \\ z' &= z\,. \end{aligned} \tag{1.3}$$

Note that for $v \ll c$, the Lorentz transformations[1] reduce to purely Galilean transformations while still maintaining a constant speed of light. We have therefore altered the way we measure space and time to ensure that all observers measure the speed of light to be exactly the same.

In special relativity, the **Minkowski metric**

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.4}$$

is used, which takes the same form everywhere in spacetime. The line element between two events in this case is

$$\mathrm{d}s^2 = -(\mathrm{d}x^0)^2 + (\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2 \equiv \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu\,, \tag{1.5}$$

where the time coordinate is $x^0 = ct$ and the space coordinates are x^1, x^2 and x^3. This united spacetime, like that of the Newtonian scheme, is described by a fixed metric that remains globally valid in special relativity. However, experiments such as the Pound-Rebka experiment [8], which confirmed gravitational time dilation, demonstrate that this structure does not hold universally in the presence of a gravitational field, as described by general relativity.

[1] A rotation-free Lorentz transformation is known a Lorentz *boost*, making (1.3) more specifically a Lorentz boost in the x-direction.

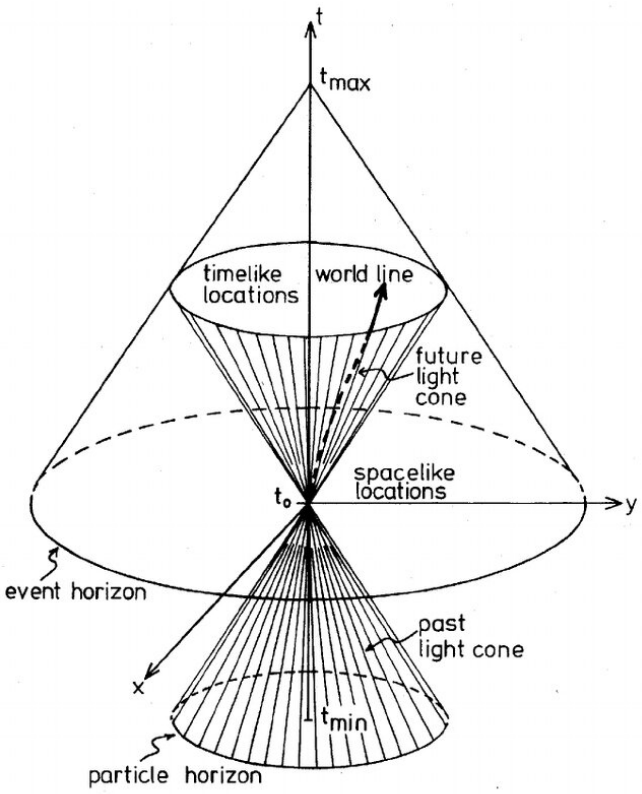


Figure 1.1: A light cone in x, y, t-space, showing that events at the origin ($x = y = 0$) at present time (t_0) follow world lines within the future light cone. These events are time-like with respect to the origin. No signals can be sent to or received from space-like locations. The region in the past from which signals can be received at the origin is bounded by the particle horizon, while the event horizon limits the space causally related to the present origin at a future time $t_{\max}$. Image from Roos (2008) [41].

The set of points such that $\mathrm{d}s^2 = 0$ represents the light cone – so called because light spreading out from an event $P(t_0) = P_0$ forms a three-dimensional cone in four-dimensional spacetime – is known as the future light cone. Similarly, the set of events from which a pulse of light is able to reach P_0 is known as the past line cone. These can be seen above and below t_0 in figure 1.1, aptly forming an hourglass shape. These hourglasses exist for every event at t_0, and divide spacetime into three distinct regions. The absolute future, inside the future light cone, contains all events that can possibly be affected by the event at t_0 [41]. The absolute past, inside the past light cone, contains all the events from which signals travelling at or below the speed of light can affect the event at t_0. Since the velocity of any body cannot exceed the speed of light, only pairs of points (P_0, P_1) that satisfy $\mathrm{d}s^2 < 0$ can be connected, i.e., P_1 must be in P_0's light cone for P_0's trajectory to reach it. This interval is said to be time-like. Events outside of the hourglass – a region known as elsewhere – satisfy $\mathrm{d}s^2 > 0$. This interval is said to be space-like, and a point elsewhere cannot reach P_0 via any physical trajectory [42].

Time-like intervals are related to the *proper time* τ via

$$\mathrm{d}s^2 = -c^2 \mathrm{d}\tau^2 \,. \tag{1.6}$$

This measures the time elapsed between two events as seen by an observer following a straight path between the events [43]. If t is the time measured by a clock at a fixed position in space, τ is the time measured by a clock that actually passes through both events [40]. The proper time between two events evaluated in an inertial frame is therefore the same for both a moving

observer and an observer at rest, making proper time an invariant quantity.

The trajectory of any massive particles can be represented parametrically as $x^\mu(\lambda)$, where λ is an arbitrary parameter that increases towards infinity. As a function of proper time, this becomes

$$\tau = \int \sqrt{-\eta_{\mu\nu} U^\mu U^\nu} \mathrm{d}\lambda \,, \tag{1.7}$$

where $U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$ is the tangent vector to the worldline (found in figure 1.1). The new tangent vector is given by $u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ and satisfies $u^\mu u_\mu = -c^2$. If we consider an observer O' moving at velocity $\boldsymbol{v}$ with respect to inertial reference frame S containing observer O, the proper times measured by each observer are related by

$$\mathrm{d}\tau_{O'} = \sqrt{1 - \frac{v^2}{c^2}} \, \mathrm{d}\tau_O \,. \tag{1.8}$$

This demonstrates the phenomenon of time dilation: the duration of the journey measured by O is different to that measured O', and the extent of the difference depends on their trajectories. We can similarly apply the Lorentz transformations to 'proper length' to demonstrate length contraction.

1.2 General relativity

Special relativity is undoubtedly useful for studying space and time in inertial reference frames using comparatively simple calculations. However, we need to use general relativity if we want to incorporate the effects of gravity and spacetime curvature in our calculations. Einstein's Theory of General Relativity (GR) [5, 6] is generally considered to be the most accurate description of spacetime geometry. One of the core concepts on which Einstein based his theory is that in Newtonian gravity all objects fall identically in an external gravitational field regardless of mass or composition [42]. The universality of free fall is a result of the weak equivalence principle – introduced in §1.4 – that inertial mass is equal to gravitational mass. It is referred to as 'weak' as it only accounts for bodies influenced by gravity. Einstein realised that he would need to extend this principle to include the other laws of physics.

If we consider, as Einstein did, an observer in a free-falling elevator releasing an object at rest relative to the elevator, they will observe it floating weightlessly. Similarly, they will observe projectiles travel in a straight line at constant velocity instead of in a downward arc. This is because the object and the elevator have the same acceleration relative to Earth due to the weak equivalence principle, i.e., the object's acceleration relative to the elevator is zero. These observations are valid only for a uniform gravitational field. This is not the case for Earth; the gravitational field acts radially towards its centre. Nevertheless, if we consider this case over a short time in a small, non-rotating space, the elevator mimics an inertial frame of reference. This leads to one of the postulates of the Einstein equivalence principle: that for a free-falling, non-rotating laboratory occupying a small region of spacetime, the laws of physics adhere to those of special relativity [44]. Gravity and acceleration are therefore locally indistinguishable, and what we perceive as gravity is actually an effect of being in an accelerating reference frame [45]. Einstein concluded that we can eliminate the concept of a gravitational force related to the choice of a non-inertial reference frame by applying curvature to the space [42]. Gravity is then just a manifestation of that curvature, which is induced by the presence of matter [44]. We can show this by starting again with the line element connecting space and time:

$$\mathrm{d}s^2 = g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu \,, \tag{1.9}$$

where the metric tensor $g_{\mu\nu}$ does not necessarily describe flat spacetime, unlike $\eta_{\mu\nu}$ in (1.5). In general relativity, the metric needs to incorporate the effects of gravity. It is therefore dependent on the location in spacetime and the associated distribution of matter and energy at that point in the Universe. This implies that in a region of space devoid of mass and energy, straight lines describe the shortest distance between points and the metric is just Minkowski. The introduction of mass into this region warps the background spacetime such that free particles follow curved trajectories, giving the impression that these particles are being 'pulled' by gravity. General relativity therefore does away with the concept of gravity as a *force* [45]. We can instead characterise it as a metric theory, since the spacetime geometry is described by a metric (1.9); free bodies conform to geodesics (1.33); and the laws of physics default to special relativity under certain conditions (detailed above). To summarise, Einstein's equivalence principle is based on the following concepts [42]:

- The weak equivalence principle or the universality of free fall, according to which test body trajectories are independent of composition;
- Local position invariance, such that non-gravitational experiments yield results independent of their occurrence in spacetime; and
- Local Lorentz invariance, by which the results of non-gravitational experiments in a free-falling laboratory are independent of any motion of the laboratory.

These ideas form the foundation of general relativity and allow us to build further on Einstein's theory.

1.2.1 Differential geometry

The Euclidian space in Newtonian mechanics is unable to account for spatial curvature. In order to fully generalise Newtonian theory to four dimensions, we need to review tensor geometry and covariance. From above, we know that the metric allows us to measure distances and angles on a surface, or manifold. The metric can also be used to connect the vector space V to its dual V^*[2] via the isomorphism ψ. This means that every vector field X has a corresponding one-form $g(X,\cdot)$. Under ψ, (1.1) becomes

$$X^i\partial_j = X \to g(X,\cdot) = g_{ij}X^i\mathrm{d}x^j\,. \tag{1.10}$$

If we define the covariant components of the vector field as $X_j \equiv g_{ij}X^i$, this can be written as $X^j\partial_j \to X_j\mathrm{d}x^j$. This is known as lowering the index with the metric. It is also possible to raise indices with the metric, since $X^i = g_{ij}X_j$. These raising and lowering operations can be extended to tensors. For example,

$$T^{ij}{}_k = g^{il}g^{jm}T_{lmk} \quad \text{and} \quad T_{ijk} = g_{im}g_{jn}g_{kp}T^{mnp}\,. \tag{1.11}$$

The inverse metric $g^{\mu\nu}$ is related to the metric tensor via the Kronecker delta δ,

$$g^{\mu\nu}g_{\nu\rho} = \delta^\mu_\rho\,, \tag{1.12}$$

which is also yielded when raising an index on the metric:

$$g^{\mu\nu}g_{\mu\nu} = \delta^\mu_\mu = 4\,, \tag{1.13}$$

confirming that we are indeed working in 4-dimensional spacetime for $\mu,\nu = 0,1,2,3$.

[2]The set of linear scalar functions $\ell : V \to X$.

Tensors transform under a change of coordinates $x^\mu \to x^{\mu'}$ according to the tensor transformation law:

$$T^{\mu'}{}_{\nu'\rho'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial x^\rho}{\partial x^{\rho'}}T^\mu{}_{\nu\rho}\,, \tag{1.14}$$

where the unprimed indices on the right-hand side are dummy indices that are summed over. The metric tensor transforms analogously to the Minkowski case:

$$g_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}}g_{\mu\nu}\,. \tag{1.15}$$

The partial derivative ∂_μ is a non-tensorial object we will use a great deal. When acting on a scalar it returns a simple (0,1) tensor, but on a vector V^μ we obtain something a little more complicated:

$$\begin{aligned}\partial_\mu V^\nu \to \partial_{\mu'}V^{\nu'} &= \left(\frac{\partial x^\mu}{\partial x^{\mu'}}\partial_\mu\right)\left(\frac{\partial x^{\nu'}}{\partial x^\nu}V^\nu\right)\\ &= \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^{\nu'}}{\partial x^\nu}\left(\partial_\mu V^\nu\right) + \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial^2 x^{\nu'}}{\partial x^\nu \partial x^\mu}V^\mu.\end{aligned} \tag{1.16}$$

It is therefore not guaranteed to transform like a tensor. To remedy this, we define the **covariant derivative** ∇_μ [3] as a partial derivative plus a correction:

$$\boxed{\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu{}_{\mu\lambda}V^\lambda\,.} \tag{1.17}$$

The symbol $\Gamma^\nu_{\mu\lambda}$ is known as a connection coefficient. It represents a bundle of numbers with an appropriate non-tensorial transformation law designed to cancel out the non-tensorial term in (1.16). The connection coefficient used here in particular is the **Christoffel symbol**, defined by

$$\boxed{\Gamma_{\lambda\mu\nu} = \frac{1}{2}\left(\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu}\right) \quad \text{and} \quad \Gamma^\alpha_{\mu\nu} = g^{\lambda\alpha}\Gamma_{\lambda\mu\nu}\,.} \tag{1.18}$$

Therefore, for $\nabla_\mu V^\nu$ to transform like a tensor, we need to have

$$\Gamma^{\nu'}{}_{\mu'\lambda'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\lambda}{\partial x^{\lambda'}}\frac{\partial x^{\nu'}}{\partial x^\nu}\Gamma^\nu{}_{\mu\lambda} - \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\lambda}{\partial x^{\lambda'}}\frac{\partial^2 x^{\nu'}}{\partial x^\mu \partial x^\lambda}\,, \tag{1.19}$$

clearly indicating that Christoffel symbols do not transform as tensors [43]. This also works for tensors with lower indices; we only need to introduce a minus sign and change the dummy index to be summed over:

$$\nabla_\mu V_\nu = \partial_\mu V_\nu - \Gamma^\lambda{}_{\mu\nu}V_\lambda\,. \tag{1.20}$$

The general expression for the covariant derivative of an arbitrary (m,n)-tensor is

$$\begin{aligned}\nabla_\alpha T^{\mu_1\ldots\mu_m}{}_{\nu_1\ldots\nu_n} &= \partial_\alpha T^{\mu_1\ldots\mu_m}{}_{\nu_1\ldots\nu_n} + \Gamma^{\mu_1}_{\alpha\lambda_1}T^{\lambda_1\mu_2\ldots\mu_m}{}_{\nu_1\ldots\nu_n} + \ldots + \Gamma^{\mu_m}_{\alpha\lambda_n}T^{\mu_1\ldots\mu_{m-1}\lambda_m}{}_{\nu_1\ldots\nu_n}\\ &\quad - \Gamma^{\lambda_1}_{\alpha\nu_1}T^{\mu_1\ldots\mu_m}{}_{\lambda\nu_2\ldots\nu_n} - \ldots - \Gamma^{\lambda_m}_{\alpha\nu_n}T^{\mu_1\ldots\mu_m}{}_{\nu_1\ldots\nu_{n-1}\lambda_n}\,.\end{aligned} \tag{1.21}$$

A consequence of this result is that $\nabla_\alpha g_{\mu\nu} = \partial_\alpha g_{\mu\nu} - \Gamma^\sigma_{\alpha\mu}g_{\sigma\nu} - \Gamma^\sigma_{\alpha\nu}g_{\sigma\mu}$. From the definition in (1.18) we can gather that since the metric is symmetric, $\Gamma_{\lambda\mu\nu} = \Gamma_{\lambda\nu\mu}$ and $\Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu}$. We can also derive the expression for the partial derivative of the metric,

$$\partial_\alpha g_{\mu\nu} = g_{\lambda\mu}\Gamma^\lambda_{\alpha\nu} + g_{\lambda\nu}\Gamma^\lambda_{\alpha\mu}\,. \tag{1.22}$$

[3] The covariant derivative is also commonly denoted by a subscript '$;\mu$', i.e., $\nabla_\mu T^\nu = T^\nu{}_{;\mu}$, and the ordinary partial derivative by a subscript '$,\mu$'.

In four dimensions there are $4 \times 10 = 40$ Christoffel symbols, equal to the exact number of partial derivatives in $\partial_\alpha g_{\mu\nu}$. Combining all of these results yields

$$\nabla_\alpha g_{\mu\nu} = 0\,, \tag{1.23}$$

meaning that we can associate a specific covariant derivative to each metric, such that the above is satisfied [42].

Like the ordinary derivative, the covariant derivative can be used to define mathematical operators such as curl,

$$\nabla_\mu T^\mu = \partial_\mu T^\mu + \Gamma^\mu_{\mu\lambda} T^\lambda = \partial_\mu T^\mu + T^\lambda \partial_\lambda \ln\sqrt{-g}\,, \tag{1.24}$$

divergence of a vector field,

$$\nabla_\mu T^\mu = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} T^\mu\right)\,, \tag{1.25}$$

and the divergence of an antisymmetric tensor,

$$\nabla_\mu F^{\mu\nu} = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} F^{\mu\nu}\right)\,, \tag{1.26}$$

using the fact that the contraction of a symmetric and antisymmetric tensor vanishes. Finally we can define the d'Alembertian operator, $\Box = \nabla_\mu \nabla^\mu$. When applied to a scalar, it yields

$$\Box\,\phi = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} g^{\mu\nu} \partial_\nu \phi\right)\,, \tag{1.27}$$

since $\nabla\phi = \partial_\mu \phi$.

In addition to the covariant derivative, the **Lie derivative** – denoted $\pounds$ – can also be used to measure a change of a tensor field from point to point on a manifold. It allows us to examine how a tensor changes as we move along the integral curves of the vector field. The Lie derivative of the tensor along the vector field is then defined [43]. Given a vector field ξ^μ, we want to define how the coordinates of vector T^μ change for an observer moving along the flow curves of ξ^μ from point P with coordinates x^α to point Q with coordinates $x^\alpha + \varepsilon\xi^\alpha$, $\varepsilon \ll 1$. We also assume the observer transports its system of coordinates with itself. The coordinates of the vector T^μ at Q are given by

$$T'^\mu(Q) = \frac{\partial x'^\mu}{\partial x^\nu} T^\nu(x^\alpha + \varepsilon\xi^\alpha) = (\delta^\mu_\nu - \varepsilon\partial_\nu \xi^\mu)[T^\nu(P) + \varepsilon\xi^\alpha \partial_\alpha T^\nu(P)]\,. \tag{1.28}$$

The Lie derivative is obtained by comparing this with T^μ at P

$$\pounds_\xi T^\mu = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}[T'^\mu(Q) - T^\mu(P)]\,, \tag{1.29}$$

giving

$$\pounds_\xi T^\mu = \xi^\nu \partial_\nu T^\mu - T^\nu \partial_\nu \xi^\mu\,. \tag{1.30}$$

This can be generalised to any tensor as

$$\boxed{\pounds_\xi T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q} = \xi^\alpha \partial_\alpha T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q} - \sum_{i=1}^{p} T^{\mu_1\ldots\alpha\ldots\mu_p}_{\nu_1\ldots\nu_q} \partial_\alpha \xi^{\mu_i} + \sum_{j=1}^{q} T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\beta\ldots\nu_q} \partial_{\nu_j} \xi^\beta\,,} \tag{1.31}$$

where $T'^{\mu}_{\nu}(x') = T^{\mu}_{\nu}(x')$ if $\pounds_{\xi} T^{\mu} = 0$[4][42]. The Lie derivative essentially applies the conventional definition of an ordinary derivative to the tensor's component functions. It is a more fundamental concept than the covariant derivative, as it does not depend on the choice of a connection – though it does require the presence of a vector field [43]. We will need to use this in §3.1.

With the covariant derivative in hand, we can begin to define the theory of parallel transport. A vector undergoes parallel transport if it satisfies $\mathrm{D}T^{\mu} \equiv \mathrm{d}x^{\nu}\nabla_{\nu}T^{\mu} = 0$. A vector travelling along a curve $x^{\mu}(\tau)$[5] can therefore be defined by

$$\frac{\mathrm{D}T^{\mu}}{\mathrm{D}\tau} = u^{\nu}\nabla_{\nu}T^{\mu} = \frac{\mathrm{d}T^{\mu}}{\mathrm{d}\tau} + \Gamma^{\mu}_{\nu\alpha}T^{\alpha}u^{\nu} = 0\,, \tag{1.32}$$

where $u^{\mu} = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}\tau}$ is the tangent vector. In the special case of auto-parallel transport, the curve parallel transports its own tangent vector, i.e., $\frac{\mathrm{D}u^{\mu}}{\mathrm{D}\tau} = u^{\nu}\nabla_{\nu}u^{\mu} = 0$. We can rewrite this as

$$\boxed{u^{\nu}\partial_{\nu}u^{\mu} + \Gamma^{\mu}_{\alpha\nu}u^{\alpha}u^{\nu} = 0 \implies \frac{\mathrm{d}^2x^{\mu}}{\mathrm{d}\tau^2} + \Gamma^{\mu}_{\alpha\nu}\frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\tau} = 0\,.} \tag{1.33}$$

This formula is simply the **geodesic equation** in curved spacetime[6], so we can conclude that geodesics are auto-parallel. Parallel transport requires a well-defined connection; in flat space the Christoffel connections are implicit, but in curved space, the result of parallel transporting a vector depends on the path taken. Different connections therefore yield different results. If the connection is metric-compatible, the metric is always parallel transported, such that $\mathrm{D}g_{\mu\nu} = \mathrm{d}x^{\sigma}\nabla_{\sigma}g_{\mu\nu} = 0$. From this, it follows that the inner product of two parallel-transported vectors is preserved [42].

Let us examine the geodesic (1.33) in more detail. The geodesic is the equivalent of a straight line in curved space, representing the shortest distance between two points. More formally, a geodesic curve extremises the spacetime interval. This means that the infinitesimal distance along a trajectory $x^{\mu}(\lambda)$ is

$$\mathrm{d}s = \sqrt{\left|g_{\mu\nu}\frac{\mathrm{d}x^{\mu}}{\mathrm{d}\lambda}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\lambda}\right|}\,\mathrm{d}\lambda\,, \tag{1.34}$$

with length $L = \int \mathrm{d}s$. These equations are useful because they represent the trajectories of unaccelerated test particles. In this way, the geodesic equation is analogous to Newton's law $\mathbf{F} = m\mathbf{a}$ in curved space for the case $\mathbf{F} = 0$. One important thing to note here is that we can choose any parameter $\lambda(\tau)$ such that the geodesic curve is described by $x^{\mu}(\lambda)$ and (1.33) remains invariant. Here we used a general parameter λ, but when deriving the extremal of the spacetime interval, we can use the specific parameterisation of proper time. If the geodesic we are working with describes 'light-like' trajectories, i.e. $\mathrm{d}s^2 = 0$ (see figure 1.1), then we cannot parametrise it using time because light particles have no reference frame. We can therefore generalise (1.33) to

$$\frac{\mathrm{d}^2x^{\mu}}{\mathrm{d}\lambda^2} + \Gamma^{\mu}_{\alpha\nu}\frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\lambda}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\lambda} = 0\,. \tag{1.35}$$

The geodesic equation remains invariant if and only if it has a linear relation to λ, i.e.

$$\lambda(\tau) = a\tau + b\,, \tag{1.36}$$

[4] An object with zero Lie derivative is said to be *Lie dragged*.

[5] Recall that proper time τ is used if the curve is timelike; otherwise an arbitrary parameter along the curve is used.

[6] In Minkowski spacetime the proper time τ would be replaced by coordinate time t.

where a and b are integration constants. This condition is called the affine transformation; it guarantees that any two parameters related in this way will leave the geodesic equation invariant. The parameter λ is therefore referred to as the affine parameter [45]. Any parameter related to proper time via (1.36) is called an affine parameter and is equally valid for parameterising a geodesic. We can of course use any other parametrisation, though this will not satisfy (1.33). Instead, we would need to solve an equation of the form [43]

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\sigma^2} + \Gamma^\mu_{\alpha\nu}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma} = f(\sigma)\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma} \tag{1.37}$$

for some parameter $\sigma(\lambda)$ and function $f(\sigma)$. If this equation is satisfied along a curve, it is always possible to find an affine parameter which satisfies (1.33). For our purposes, since we only consider timelike geodesics, we can use proper time as the affine parameter. Substituting the four-velocity for the tangent vector then allows us to write the **geodesic equation** as

$$\boxed{\frac{\mathrm{d}U^\mu}{\mathrm{d}\tau} + \Gamma^\mu{}_{\rho\sigma}U^\rho U^\sigma = 0\,.} \tag{1.38}$$

Test bodies will always move along these curves in general relativity, provided that they are not under the influence of any other forces [43]. The particle's motion along a geodesic is essentially equivalent to its free-fall acceleration. A particle influenced by external forces is forced to *not* follow the geodesic, and its acceleration will *not* vanish as it does in (1.38) [45]. We can redefine the four-acceleration using a combination of the chain rule and the substitution $\partial \to \nabla$ to obtain

$$A^\mu = \frac{\mathrm{d}U^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial U^\mu}{\partial x^\nu} = U^\nu \partial_\nu U^\mu = U^\nu \nabla_\nu U^\mu\,, \tag{1.39}$$

which allows us to write the geodesic equation – a vanishing acceleration – as

$$U^\nu \nabla_\nu U^\mu = 0\,. \tag{1.40}$$

We can use this to briefly define a concept we will need in §2.1: the Killing vector. Taking an arbitrary four-vector ξ^μ we can construct the derivative

$$U^\nu \nabla_\nu (\xi_\mu U^\mu) = U^\nu U^\mu \nabla_\nu \xi_\mu + \xi_\mu U^\nu \nabla_\nu U^\mu, \tag{1.41}$$

where the last term vanishes due to (1.40). If $\xi_\mu U^\mu$ is conserved along the path of the particle, the left-hand side must go to zero. For this to happen, ξ_μ must satisfy

$$U^\nu U^\mu \nabla_\nu \xi_\mu = 0\,, \tag{1.42}$$

and since this expression is symmetric under $\mu \leftrightarrow \nu$, we can write

$$U^\nu U^\mu \nabla_\mu \xi_\nu = U^\mu U^\nu \nabla_\nu \xi_\mu = 0 \implies U^\mu U^\nu (\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu) = 0, \tag{1.43}$$

yielding **Killing's equation**:

$$\boxed{\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0\,.} \tag{1.44}$$

Its solutions ξ^μ define symmetries of the metric, i.e. quantities $\xi_\mu U^\mu$ that are conserved for free particles. The metric therefore remains unchanging along the direction of the Killing vector. The technical term for this symmetry is isometry; if some coordinate x^α in the metric $g^{\mu\nu}$ undergoing an infinitesimal change

$$x^\alpha \to x^\alpha + \epsilon \xi^\alpha \quad , \quad \epsilon \ll 1 \tag{1.45}$$

leaves the metric invariant, i.e. $g_{\mu\nu}(x^\alpha + \epsilon\xi^\alpha) = g_{\mu\nu}(x^\alpha)$, this is called an isometry. Therefore, if all components of the metric are independent of x^α we have an isometry and a Killing vector with a single component along x^α [45]. One obvious example is Minkowski space (1.5), which is a maximally symmetric space and thus possesses the maximum number of isometries, corresponding to ten Killing vectors. The fact that the metric coefficients $\eta_{\mu\nu}$ are independent of all x^μ shows that the metric is invariant under (1.45). Minkowski space has 10 Killing vectors in total, representing translation, rotation, and Lorentz boosts. In $n \geq 2$ dimensions, there may be more Killing vectors than dimensions, as a linear combination of Killing vectors with constant coefficients still constitutes a Killing vector [43]. We'll see this again in §2.1, but for now, we can turn our attention back to the metric in curved space and the challenge of constructing a metric that describes the Universe as a whole.

With all the prerequisite concepts completed, we can begin to derive some important tensors. We start by returning to parallel transport and considering a vector on a sphere as in figure 1.2. Maintaining a constant direction results in the vector no longer being tangent to the surface, but rather projecting outward. Conversely, keeping the vector flat against the surface voids our global reference direction. Thus, the concept of maintaining the 'same direction' is not well-defined on a curved surface. Looking at figure 1.2, we can see that despite each segment being parallel to its neighbour, the vector field has rotated 90°in travelling from $A \to B \to C \to A$. More formally, we observe that a vector V^μ undergoing parallel transport in curved space, as opposed to Euclidian space, does not return to its initial state.

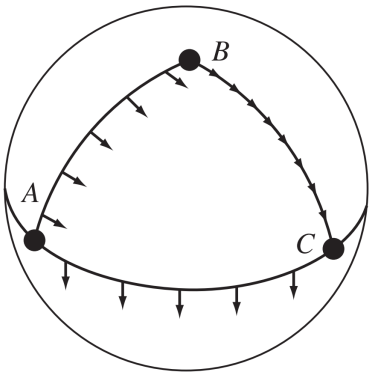


Figure 1.2: Parallel transport around a spherical triangle. Drawing vectors on a sphere requires maintaining tangency to the surface. In this closed loop $ABCA$, vectors rotate 90°despite being drawn parallel due to the sphere's curvature, illustrating its effect on space. Image from Schutz (2009) [40].

One way to analyse this would be to hold the covariant derivative at zero while transporting a vector around an infinitesimal loop. However, it is quicker and easier to make use of the commutator of the covariant derivative. The covariant derivative of a tensor in a specific direction measures the extent to which the tensor changes relative to what it would be if it was parallel transported – for which it would have a covariant derivative of zero. The commutator of two covariant derivatives therefore measures the difference between parallel transporting the tensor first one way and then the other, and vice versa. For a vector field V^ρ, the commutator of its covariant derivatives is

$$\begin{aligned}[\nabla_\mu, \nabla_\nu]V^\rho &= \nabla_\mu\nabla_\nu V^\rho - \nabla_\nu\nabla_\mu V^\rho \\ &= \left(\partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}\right) V^\sigma \\ &= R^\rho_{\sigma\mu\nu}V^\sigma, \end{aligned} \tag{1.46}$$

where we have defined the **Riemann tensor**, $R^{\rho}{}_{\sigma\mu\nu}$, as

$$\boxed{R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\sigma}, \quad R_{\lambda\sigma\mu\nu} = g_{\lambda\rho}R^{\rho}{}_{\sigma\mu\nu}\,.} \tag{1.47}$$

Notice that despite being constructed from non-tensorial components, the Riemann tensor still transforms as a tensor due to their deliberate arrangement. From the above equation we can conclude that the Riemann tensor is antisymmetric in the permutation of the first and second pair of indices, as well as symmetric in the permutation of the two pairs, i.e.

$$R_{\rho\mu\nu\omega} = -R_{\mu\rho\nu\omega} = -R_{\rho\mu\sigma\nu}, \quad R_{\rho\mu\nu\sigma} = R_{\nu\sigma\rho\mu}\,. \tag{1.48}$$

While we would expect the Riemann tensor to have $N^4 = 4^4 = 256$ components, these equations imply that it has the same number of independent components as a 6×6 symmetric matrix: 21 [42]. We can also derive

$$3R_{\lambda[\sigma\mu\nu]} = R_{\lambda\sigma\mu\nu} + R_{\lambda\nu\sigma\mu} + R_{\lambda\mu\nu\sigma} = 0\,. \tag{1.49}$$

This adds an extra constraint to the components, bringing the total down to $N^2(N^2-1)/12 = 20$ independent components [44]. The derivatives of the Riemann also exhibit cyclic symmetry such that

$$\boxed{\nabla_{[\rho}R_{\mu\nu]\lambda\sigma} = \nabla_\rho R_{\lambda\sigma\mu\nu} + \nabla_\nu R_{\lambda\sigma\rho\mu} + \nabla_\mu R_{\lambda\sigma\nu\rho} = 0\,,} \tag{1.50}$$

called the **Bianchi identity**. The last and most intuitive property is that all components of the Riemann tensor vanish if and only if the space is flat [43]. There are two contractions of the Riemann tensor – the Ricci tensor and the Ricci scalar – that will prove very useful when studying the action in §1.2.2. The **Ricci tensor** is obtained from the Riemann tensor by contracting the first and third indices and given by

$$\boxed{R_{\alpha\beta} = R^{\sigma}{}_{\alpha\sigma\beta}\,,} \tag{1.51}$$

and the **Ricci scalar** by the trace of the Ricci tensor with respect to the metric:

$$\boxed{R = R^{\sigma}{}_{\sigma} = g^{\mu\nu}R_{\mu\nu}\,.} \tag{1.52}$$

The symmetries of the Riemann tensor imply an additional symmetry of the Ricci tensor, i.e. $R_{\mu\nu} = R_{\nu\mu}$. With the Ricci tensor in hand, we can finally define the **Einstein tensor**:

$$\boxed{G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\,.} \tag{1.53}$$

The Bianchi identity (1.50) implies that the divergence of this tensor also vanishes, i.e $\nabla^\mu G_{\mu\nu} = 0$. This is known as the contracted Bianchi identity. We will now use everything we have derived to obtain the field equations.

1.2.2 Equations of motion

In the same way that the Maxwell equations are the field equations for electromagnetism, the Einstein equations are the field equations that describe quantitatively how the curvature of spacetime at any time is related to the matter distribution at that time. The Maxwell and

Einstein equations are second-order partial differential equations for the components of the electromagnetic field tensor $F_{\mu\nu}$ and the metric spacetime coefficients $g_{\mu\nu}$, respectively.

We begin by identifying a tensor that describes the matter distribution at certain points in spacetime. We first consider a neutral, non-interacting particle with rest mass m_0. A time-dependent fluid of these particles is commonly referred to as *dust*, with matter density ρ and four-velocity u^μ as measured in some inertial frame. In some arbitrary coordinate system x^μ, the contravariant components of the **energy-momentum tensor**, or stress tensor, are

$$\boxed{T^{\mu\nu} = \rho u^\mu u^\nu \,.} \tag{1.54}$$

For a local Cartesian inertial frame, the *spatial* components of $T^{\mu\nu}$ represent [44]

T^{00} : energy density of the particles;

T^{0i} : energy flux $\times\, c^{-1}$ in the i-direction;

T^{i0} : momentum density $\times\, c$ in the i-direction;

T^{ij} : flux of the ith component of momentum in the j-direction.

For a real fluid, we also need to take into account that (i) besides the bulk velocity, each particle has some random thermal velocity and (ii) various forces may exist between particles that contribute to the potential energy (T^{00}). Energy may be transmitted via heat conduction, contributing to T^{0i}, and therefore carry momentum, contributing to T^{i0}. Random thermal motions create momentum flow, so the isotropic pressure in the i-directon is given by T^{ii} and the viscous stresses by T^{ij} $(i \neq j)$.

For a perfect fluid, we neglect the forces between particles, as well as heat conduction and viscosity. The components are simply

$$[T^{\mu\nu}] = \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} . \tag{1.55}$$

This can be generalised for an arbitrary coordinate system to give a fully covariant expression for the **energy-momentum tensor for a perfect fluid**:

$$\boxed{T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right) u^\mu u^\nu - p g^{\mu\nu} \,.} \tag{1.56}$$

It is easy to see that $T^{\mu\nu}$ is symmetric ($T^{\mu\nu} = T^{\nu\mu}$) and that the momentum density is equal to the energy flux ($T^{\mu 0} = T^{0\mu}$). The latter is due to mass-energy equivalence; the energy flux is equal to the mass density multiplied by the speed at which it flows, i.e. the momentum density. We can also see that for a pressureless fluid like dust, $p \to 0$ and every component of the energy-momentum tensor is zero besides T^{00} [44, 40], reducing the stress tensor for dust to simply equation 1.54. For our purposes it is important to remember that the energy-momentum tensor is conserved such that

$$\nabla_\mu T^{\mu\nu} = 0 \,. \tag{1.57}$$

Before moving on to the action, it would be beneficial to briefly mention energy conditions. These impose coordinate-invariant restrictions on the energy-momentum tensor. Let's take the

perfect fluid from above as an example. We can construct scalars from the tensor by contracting it with timelike (t^μ) or null (ℓ^μ) vectors. The *weak energy condition* demands that $T_{\mu\nu}t^\mu t^\nu \geq 0$ for all timelike vectors. Since we can express a timelike vector as a linear combination $t^\mu = u^\mu + \ell^\mu$, the weak energy condition is satisfied if $T_{\mu\nu}u^\mu u^\nu \geq 0$ and $T_{\mu\nu}\ell^\mu \ell^\nu \geq 0$. We therefore have

$$T_{\mu\nu}u^\mu u^\nu = \rho \quad , \quad T_{\mu\nu}\ell^\mu\ell^\nu = \left(\rho + \frac{p}{c^2}\right)(u_\mu \ell^\mu)^2 \,, \tag{1.58}$$

which implies $\rho \geq 0$ and $\rho + p \geq 0$ for a perfect fluid. We additionally have a *strong energy condition*, which requires that

$$T_{\mu\nu}t^\mu t^\nu \geq \frac{1}{2}T^\lambda{}_\lambda\, t^\sigma t_\sigma \tag{1.59}$$

for all timelike vectors t^μ, implying that $\rho + p \geq 0$ and $\rho + 3p \geq 0$. While we do not make use of the energy conditions going forward, we will be expressing the pressure and density in terms of the **equation-of-state parameter**,

$$\boxed{w \equiv \frac{p}{\rho c^2}} \,. \tag{1.60}$$

In general this parameter is constant, but more exotic theories of gravity express it as a function of time. We will discuss the equation of state for different cosmological models in the next chapter, and move now to the Einstein field equations.

1.3 Einstein field equations

The most popular method of deriving the Einstein field equations is applying the principle of least action to the Einstein-Hilbert equation. Also known as Hamilton's principle, the principle of least action allows us to determine the dynamical equations of a particle or field [46]. The motion of any particle or the behaviour of any physical field has an associated integral known as the action S. Taking the vanishing variation, i.e. $\delta S = 0$ yields the equations of motion. An integral over time gives the well-known Lagrangian function L, while an integral over spacetime gives the Lagrangian density function $\mathcal{L}$.

Lagrangian mechanics relies on two fundamental concepts that are widely applicable across physics: the Lagrangian itself, a function that characterises a particle's motion; and the action, which defines a path in spacetime. Every trajectory in spacetime has a unique action associated with it. The Lagrangian can be expressed as a function of the kinetic and potential energies $\mathcal{L}(T, U)$, and the action as the integral of the Lagrangian over all of spacetime $\int \mathcal{L}\, \mathrm{d}^4x$. Through observations, we know that physical objects and fields will always behave in such a way that the action is minimised, i.e. stationary. This is the principle of least action.

For this derivation, we require that 1) the action must remain scalar while yielding tensorial field equations, 2) the action must account for both curvature and matter, and 3) the resulting field equations must reduce to **Poisson's equation** in the Newtonian weak-field limit,

$$\boxed{\nabla^2\Phi = 4\pi G\rho \,.} \tag{1.61}$$

To begin, we need to construct an action function; demanding that this functional remains stationary will allow us to obtain the equations of motion. Since our goal is to relate curvature to matter, we can start simply with

$$S = S_{\text{curvature}} + S_{\text{matter}} \,. \tag{1.62}$$

While the action itself is scalar, it must yield tensors to match the form of the field equations. This can be achieved by constructing the action from index-free tensor quantities. The matter part S_m is case-specific, so we keep it as general as possible. The curvature part S_c is formed from an integral of the Lagrangian $\mathcal{L}_c$ over four-dimensional spacetime. Typically when integrating in a curved coordinate system, we include a Jacobian factor inside the integral to make the volume elements behave like tensors. In this case, we use $\sqrt{-g}$, where the minus sign arises from the fact that the determinant of the metric is always negative [47]. A scalar action means that our Lagrangian also needs to be a scalar – one that describes curvature. We derived a very simple curvature scalar earlier in (1.52): the Ricci scalar [48]. We therefore obtain

$$S_{\mathrm{c}} = \int \mathrm{d}^4x\sqrt{-g}R\,. \tag{1.63}$$

We will of course need some proportionality constant outside the integral to relate the Ricci scalar to the Lagrangian, which for now we will set to $(2\kappa)^{-1}$, where κ is some coupling constant. Including this and the matter Lagrangian yields the **Einstein-Hilbert action** [49],

$$\boxed{S_{\mathrm{EH}}[g_{\alpha\beta}] = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2\kappa}R + \mathcal{L}_m(g_{\mu\nu}, \Psi_m)\right],} \tag{1.64}$$

where g is the determinant of the metric $g_{\mu\nu}$, and $\mathcal{L}_m$ is the matter Lagrangian, which is dependent on the metric and the matter fields Ψ_m. By the principle of least action, the variation in the action functional must be zero. Since both the Ricci scalar and the action are functions of the metric tensor, we can vary the action with respect to the inverse metric and write the functional derivative as

$$\delta \to \delta g_{\mu\nu}\frac{\delta}{\delta g_{\mu\nu}}\,. \tag{1.65}$$

We begin by setting the variation of the action to zero and distributing the functional derivative:

$$\begin{aligned}
0 &= \delta S \\
&= \delta \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2\kappa}R + \mathcal{L}_m(g_{\mu\nu}, \Psi_m)\right] \\
&= \int \mathrm{d}^4x \left[\frac{1}{2\kappa}\frac{\delta(\sqrt{-g}R)}{\delta g^{\mu\nu}} + \frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu} \\
&= \int \mathrm{d}^4x \left[\frac{1}{2\kappa}\left(\sqrt{-g}\frac{\delta R}{\delta g^{\mu\nu}} + R\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}}\right) + \frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu} \\
&= \int \mathrm{d}^4x \left[\frac{1}{2\kappa}\left(\frac{\delta R}{\delta g^{\mu\nu}} + \frac{R}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}}\right) + \frac{1}{\sqrt{-g}}\frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\right]\sqrt{-g}\delta g^{\mu\nu}\,.
\end{aligned}$$

Since this equation should hold for all $\delta g^{\mu\nu}$, the two terms in the square brackets should sum to zero, thus

$$\frac{\delta R}{\delta g^{\mu\nu}} + \frac{R}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}} = -\frac{2\kappa}{\sqrt{-g}}\frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}$$

represents our equation of motion. In the case of scalar fields, the energy-momentum tensor can be written as [48]

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\,, \tag{1.66}$$

leaving us with

$$\frac{\delta R}{\delta g^{\mu\nu}} + \frac{R}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}} = \kappa T_{\mu\nu}\,, \tag{1.67}$$

so we need to calculate the variations of both the metric and the Ricci scalar. We can obtain the variations of the metric and inverse metric with Jacobi's formula $\delta g = \delta \det(g_{\mu\nu}) = g g^{\mu\nu}\delta g_{\mu\nu}$, giving us

$$\begin{aligned}
\delta\sqrt{-g} &= -\frac{1}{2}\frac{\delta g}{\sqrt{-g}} \\
&= -\frac{1}{2}\frac{g g^{\mu\nu}\delta g_{\mu\nu}}{\sqrt{-g}} \\
&= \frac{1}{2}\sqrt{-g} g^{\mu\nu}\delta g_{\mu\nu} \\
&= -\frac{1}{2}\sqrt{-g} g_{\mu\nu}\delta g^{\mu\nu} \\
\implies \frac{\delta\sqrt{-g}}{\sqrt{-g}\delta g^{\mu\nu}} &= -\frac{1}{2} g_{\mu\nu} \,.
\end{aligned} \tag{1.68}$$

For the Ricci scalar, we begin with the Riemann tensor. We know that the Riemann curvature (1.47) depends only on the Levi-Civita connection $\Gamma^{\lambda}_{\mu\nu}$, so the variation of the Riemann tensor is

$$\delta R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu(\delta\Gamma^{\rho}_{\nu\sigma}) - \partial_\nu(\delta\Gamma^{\rho}_{\mu\sigma}) + \delta\Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma} + \Gamma^{\rho}_{\mu\lambda}\delta\Gamma^{\lambda}_{\nu\sigma} - \delta\Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\sigma} - \Gamma^{\rho}_{\nu\lambda}\delta\Gamma^{\lambda}_{\mu\sigma} \,. \tag{1.69}$$

The $\delta\Gamma^{\rho}_{\nu\sigma}$ term is the difference of two connections, so we can treat it as a tensor and thus calculate its covariant derivative,

$$\nabla_\mu\left(\delta\Gamma^{\rho}_{\nu_\sigma}\right) = \partial_\mu\left(\delta\Gamma^{\rho}_{\nu_\sigma}\right) + \Gamma^{\rho}_{\mu\lambda}\delta\Gamma^{\lambda}_{\nu_\sigma} - \Gamma^{\lambda}_{\mu\nu}\delta\Gamma^{\rho}_{\lambda\sigma} - \Gamma^{\lambda}_{\mu\sigma}\delta\Gamma^{\rho}_{\nu\lambda} \,. \tag{1.70}$$

Substituting this into (1.69) cancels most of the terms, giving

$$\delta R^{\rho}{}_{\sigma\mu\nu} = \nabla_\mu\left(\delta\Gamma^{\rho}_{\nu_\sigma}\right) - \nabla_\nu\left(\delta\Gamma^{\rho}_{\mu\sigma}\right) \,. \tag{1.71}$$

It now stands to reason that by contracting the indices in the variation of the Riemann tensor, we obtain the variation in the Ricci tensor:

$$\boxed{\delta R_{\sigma\nu} \equiv \delta R^{\rho}_{\sigma\rho\nu} = \nabla_\rho\left(\delta\Gamma^{\rho}_{\nu\sigma}\right) - \nabla_\nu\left(\delta\Gamma^{\rho}_{\rho\sigma}\right) \,.} \tag{1.72}$$

This is known as the **Palatini identity**. Since the Ricci scalar is $R = g^{\mu\nu}R_{\mu\nu}$, its variation with respect to the inverse metric $g^{\mu\nu}$ is given by

$$\begin{aligned}
\delta R &= R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu} \\
&= R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\left[\nabla_\rho\left(\delta\Gamma^{\rho}_{\nu\mu}\right) - \nabla_\nu\left(\delta\Gamma^{\rho}_{\rho\mu}\right)\right] \\
&= R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\rho\left[g^{\mu\nu}\delta\Gamma^{\rho}_{\nu\mu} - g^{\mu\nu}\delta\Gamma^{\rho}_{\rho\mu}\right] \,.
\end{aligned} \tag{1.73}$$

Using metric compatibility[7], i.e. $\nabla_\sigma g_{\mu\nu} = 0$, we can distribute the metric to give

$$\begin{aligned}
\delta R &= R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\rho\delta\Gamma^{\rho}_{\nu\mu}g^{\mu\nu} - \nabla_\nu\delta\Gamma^{\rho}_{\rho\mu}g^{\mu\nu} \\
&= R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\rho\delta\Gamma^{\rho}_{\nu\mu}g^{\mu\nu} - \nabla_\rho\delta\Gamma^{\sigma}_{\sigma\mu}g^{\mu\rho} \\
&= R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\rho\left(\delta\Gamma^{\rho}_{\nu\mu}g^{\mu\nu} - \delta\Gamma^{\sigma}_{\sigma\mu}g^{\mu\rho}\right),
\end{aligned} \tag{1.74}$$

where in the last term of the second line we have changed the dummy indices $(\rho, \nu) \to (\sigma, \rho)$ to isolate the covariant derivative [43]. The brackets on the final line now enclose a one-index

[7] This allows us to approximate small enough regions of spacetime as flat, even though spacetime itself is curved. This relates to our discussion of the equivalence principle in §1.2 .

tensor object which we will call V^ρ. We know that δR will end up being multiplied by $\sqrt{-g}$, making the last term in (1.74) $\sqrt{-g}\nabla_\rho V^\rho = \nabla_\rho(\sqrt{-g}V^\rho) = \partial_\rho(\sqrt{-g}V^\rho)$. By Stokes' theorem, this yields only a boundary term when integrated. If we assume that the spacetime manifold is open and has a boundary $\partial\mathcal{M}$, then $\delta g_{\mu\nu}|_{\partial\mathcal{M}} = 0$. This means that the variation of the field goes to zero at infinity and is therefore zero at all points on the boundary of spacetime. Therefore $\int \nabla_\rho V^\rho \sqrt{-g}\, \mathrm{d}^4x = 0$ and we can drop the final term in (1.74). Our result is then

$$\delta R = R_{\mu\nu}\delta g^{\mu\nu} \implies \frac{\delta R}{\delta g^{\mu\nu}} = R_{\mu\nu}\,. \tag{1.75}$$

Substituting these results (1.68) and (1.75) into (1.67) gives us the **Einstein field equations**:

$$\boxed{R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \kappa T_{\mu\nu}\,.} \tag{1.76}$$

We can find the trace of this equation by contracting both sides with the inverse metric to get

$$R = -\kappa g^{\mu\nu}T_{\mu\nu} \equiv -\kappa T\,. \tag{1.77}$$

If we substitute this for R in (1.76), we obtain

$$R_{\mu\nu} = \kappa\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right). \tag{1.78}$$

If we want to describe a weak gravitational field, we can express the metric as a small deviation from flat space,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}\,, \tag{1.79}$$

where $|h_{\mu\nu}| \ll 1$ are the small weak-field deviations. If we take the Newtonian limit and also assume the test particle's velocity is approximately zero, and the metric and its derivatives are static, this implies that we only need the 00 component of (1.78) and therefore $T = g^{00}T_{00} = (-1 - h^{00})$. The 00 component of (1.78) is then

$$\begin{aligned} R_{00} &= \kappa\rho c^2\left[1 - \frac{1}{2}(-1 + h_{00})\big(-1 - h^{00}\big)\right] \\ &= \frac{1}{2}\kappa\rho c^2\,, \end{aligned}$$

where we used $(-1 + h_{00})(-1 - h^{00}) \approx 1$. In the low-curvature limit, $h_{00} = -\frac{2}{c^2}\Phi$[8]; hence $R_{00} = \frac{1}{c^2}\nabla^2\Phi$ and

$$\nabla^2\Phi = \frac{1}{2}\kappa\rho c^4 = 4\pi G\rho\,. \tag{1.80}$$

We therefore obtain an expression for κ:

$$\kappa = \frac{8\pi G}{c^4} \sim 2.07665 \times 10^{-43}\,\mathrm{N}^{-1}\,, \tag{1.81}$$

such that the field equations (1.76) reduce to Poisson's equations (1.61) in the weak-field limit. This value is known as the Einstein gravitational constant.

When Einstein published this theory in 1916, he took the position that the Universe was static,

[8] This can be obtained by applying the weak-field approximation to the geodesic equation (1.33).

i.e. not expanding or contracting. To guarantee this, he added a cosmological constant Λ to keep the size scale of the Universe fixed. If we include Λ in the Lagrangian, we obtain the action

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2\kappa}(R - 2\Lambda) + \mathcal{L}_m\right], \tag{1.82}$$

where we can see that the factor of 2 was added to simplify the final result. If we apply the same method as above, it is easy to derive the **Einstein field equations with a cosmological constant**:

$$\boxed{R_{\mu\nu} - \frac{1}{2}Rg_{\mu u} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}\,.} \tag{1.83}$$

Einstein's doubt in his original result was a significant missed opportunity to postulate the idea of an expanding universe. He referred to the cosmological constant as "the greatest blunder of my life". It was over a decade later that Edwin's Hubble's observations would hint at an expanding universe. However, Λ would end up playing a crucial role in cosmology – though one that Einstein did not predict – manifesting as the source of the mysterious, antigravitational energy that we now call *dark energy*. In the next chapter, we will examine the motivations and implications of adding a cosmological constant, and how we can use (1.83) to define different universe models. To avoid confusion with the curvature in later chapters, we take the convention of making the Einstein gravitational constant κ equal to unity from chapter 2 onwards.

1.4 Alternative theories

Isaac Newton introduced his inverse square law in 1665 that aligned terrestrial and celestial gravity in a single theory [50]. He based his theory on two key concepts: absolute space; and the equivalence of inertial mass and gravitational mass. The former – that space is a fixed, rigid structure – proved to be an inapt description. Dicke [28] asserted that the gravitational constant should be dependent on the mass distribution in the Universe. The latter evolved into what we now know as the weak equivalence principle. Newton's theory of gravity was only truly challenged more than two centuries later when Einstein completed Special Relativity (SR) in 1905. His theory, able to explain several phenomena in non-gravitational physics, seemed incompatible with Newtonian gravity. He continued to expand his theory and completed General Relativity in 1915. Acceptance of GR by the scientific community was quickly followed by attempts to modify it. This began with Nordström, who posited a conformally flat scalar theory of gravity; and Weyl, who in 1919 proposed including higher-order curvature invariants in the Einstein–Hilbert action. Dirac advanced this work in 1937 by suggesting a varying gravitational constant, linking cosmological and fundamental constants. Jordan later expanded on this by introducing a scalar field to represent gravitational coupling, ultimately leading to the Brans–Dicke theory in 1961, which combines a scalar field and the metric tensor as the basis for gravitational interactions [31]. This theory is often considered the prototype of alternatives to GR.

Recent cosmological observations have further challenged the sufficiency of Einstein's equations, particularly at large scales, where dark energy and dark matter are invoked. An alternative perspective suggests modifying the gravitational side of the field equations rather than introducing exotic matter-energy components: modified gravity. These theories typically involve adding new fields, incorporating higher-order derivatives of curvature invariants, or increasing spacetime dimensions [26]. The final chapter of this thesis focuses on $f(R)$ theories, which introduce a general function of the Ricci scalar into the gravitational action. Originally considered for inflation, these theories gained prominence after the discovery of cosmic late-time acceleration

(see §2.2).

Any relativistic theory of gravity has to satisfy certain requirements that align with our current understanding of the Universe. The most obvious of these is that it supports astrophysical observations such as the orbits of planets and self-gravitating structures. To achieve this, it must reproduce the Newtonian dynamics in the weak-energy limit and pass the classical Solar System tests [51]. Secondly, the theory should reproduce galactic dynamics, taking into account observed baryonic matter and radiation constituents. Finally, it must account for large-scale structure in the Universe – such as the clustering of galaxies – and cosmological dynamics, by reproducing the expansion rate, Hubble parameter, and other cosmological parameters. There are also numerous observational tests such as standard sirens [52], binary pulsars [53], weak lensing [54], etc., reviews of which can be found in [31, 55].

Einstein also based his theory of GR on the following principles:

Relativity,: asserting that there is no preferred initial frame;

Equivalence,: requiring inertial effects to be locally indistinguishable from gravitational effects[9];

General Covariance,: requiring the field equations to remain invariant under the action of the group of space-time diffeomorphisms[10];

Causality,: asserting that an outcome cannot occur as a result of an event outside its past light cone; similarly, a cause cannot have an effect outside its future light cone.

In addition, GR relies on the assumption that space and time form an integrated spacetime that defaults to Minkowski spacetime in the absence of gravitational forces. Einstein also utilised earlier ideas put forward by Riemann, namely that the Universe is a curved manifold with variable curvature. This led him to postulate that gravitational forces are expressed by the curvature of a metric tensor field on a four-dimensional spacetime manifold, which is itself determined locally by the distribution of sources. Modified gravity theories should also adhere to these basic tenets.

The modified gravity framework is highly appealing for addressing the late-time acceleration of the Universe and dark energy, offering a natural gravitational alternative to dark energy where cosmic acceleration arises from the Universe's expansion, with sub-dominant terms (e.g., $1/R$) becoming significant at small curvatures [57]. It provides a seamless unification of early-time inflation and late-time acceleration by leveraging different gravitational terms at varying curvature scales, with some models supported by string/M-theory predictions [58]. Additionally, modified gravity offers a unified explanation for dark energy and dark matter, accounting for phenomena such as galaxy rotation curves without requiring exotic matter. This approach also naturally models transitions between non-phantom and phantom phases, avoiding the need for exotic scalar fields or fluids with $w < -1$, and often predicts transient phantom phases that circumvent catastrophic outcomes like a Big Rip [59]. Furthermore, modified gravity describes the Universe's transition from deceleration to acceleration, explaining dark energy dominance through the dynamics of expansion and resolving the coincidence problem. Its relevance extends to high-energy physics, potentially addressing issues like the hierarchy problem and facilitating the unification of gravity with grand unified theories (GUTs) [26]. Despite strict constraints from solar system tests, many viable models compete with General Relativity, making modified gravity a promising avenue for cosmological research.

[9] Analogous to Newton's concept of inertial and gravitational mass equivalence.

[10] An isomorphism of smooth manifolds involving an active coordinate transformation $\varphi : M \to M$ [56].

1.5 Summary

This chapter covered the theoretical foundations of general relativity, starting with the modification of classical spacetime to accommodate curvature and clarifying important principles like equivalence and Lorentz invariance. We then moved to differential geometry, where we used the concept of parallel transport to derive the Ricci scalar and tensor, the Riemann tensor, the Christoffel symbols, the covariant derivative and the geodesic equation. We then examined the energy-momentum tensor, energy conditions, the Lagrangian, and the Einstein-Hilbert action, culminating in the derivation of the Einstein field equations. We concluded by discussing the viability of alternative theories to GR. We now have a robust foundation for delving deeper into the intricacies of spacetime and gravity, and the next chapter will explore the application of these concepts to cosmology.

Chapter 2

Cosmology

The Copernican revolution not only replaced the geocentric model with the heliocentric one, but also informed the basis for a new philosophy about the Universe: humans are not uniquely placed in the cosmos; our observations from Earth are no more privileged than from any other point in the Universe [60]. This is known as the *Copernican Principle*. This is difficult to test because all observations are confined to our past light cone, limiting us to a three-dimensional light-like slice. As a result, it is challenging to identify a unique four-dimensional spacetime to test the Copernican principle. Uzan et al. [61] proposed combining observations of the luminosity distance and redshift drift to test this concept at dark energy-dominated epochs. In an expanding spacetime, the redshift of light from a distant source changes slightly over time due to the evolving expansion rate of the Universe [62, 63]. We express this as

$$\dot{z} = (1+z)H_0 - H(z)\,, \tag{2.1}$$

where H is the Hubble parameter (see (2.54)). The redshift drift method compares the redshift of the same source measured at two different times, so the data effectively span two infinitely close past light-cones instead of just one. This gives a more dynamic perspective on spacetime structure and evolution, extending our understanding beyond a static snapshot of the Universe [64]. Detecting this redshift drift requires extremely precise measurements over extended periods. Current estimates suggest that a time span of about 40 years of data collection would be necessary to observe this effect with planned instruments [65].

In stark contrast, pre-modern cosmologists had little in the way of empirical data to inform their models on the distribution of matter in the Universe. To aid their progress, they based their models on a concept that describes the Copernican principle in more explicit terms, the *Cosmological Principle*. Like many early theories, it is a generalisation that simplifies calculations by reducing the number of degrees of freedom to consider [66]. The Cosmological Principle asserts that, in the context of large-scale structure, the Universe is both *homogeneous* and *isotropic*. The former demands that all points on a spacelike hypersurface are identical, while the latter demands that the views in every direction on the hypersurface are equivalent [44]. This is obviously not the case on small scales; the voids of space are nothing like the centre of the Sun, and the views from Earth are very different depending on whether we are looking to the centre of the Milky Way or its edge [66]. The properties of homogeneity and isotropy only come into play when looking at sufficiently large volumes of the Universe, and there is observational evidence [67, 68, 69] that, on average, this is indeed the case.

The cosmic microwave background (CMB), discovered by chance in 1965 by Penzias and Wilson [70, 71], indicates thermal radiation pervading the entire universe at an almost constant temperature of 2.726 K[1] [72]. CMB photons are the most abundant particles in the Universe,

[1] Reported as 3.5 ± 1.0 K in the original letters.

originating from the period of recombination approximately 380 000 years after the Big Bang (see figure 2.4). They are so remarkably isotropic that $\Delta T/T < 10^{-5}$, making the CMB the leading evidence for an isotropic universe. The most accurate blackbody spectrum we have today is that of the CMB, which earned Mather and Smoot the Nobel Prize in 2006 [73, 74]. In addition to the CMB, both the X-ray background [75] imaged by Chandra [76], and the sky galaxy distribution – on cosmic scales – from the 2MASS [77], APM [78] and NVVS [79] surveys were found to be almost perfectly isotropic. The final evidence arises from the universal Hubble expansion, which was also found to be identical in every direction.

To test for homogeneity in the Universe, the simplest method is to map the matter distribution and identify if there is structure and, if so, the largest of these structures. This would imply a scale beyond which the Universe converges to homogeneity. The cosmic web, obtained with a multitude of redshift surveys such as SDSS [80] and 2dFGRS [81], showed that galaxies are organised into cosmic megastructures, e.g. filaments, superclusters, and voids [82, 83, 84, 85]. Long before this, Hubble [86] used galaxy number counts to demonstrate that the distribution of galaxies is homogeneous on large cosmological scales. Beyond this however, we have not yet found evidence for inhomogeneities on scales beyond $\sim 150\ h^{-1}$ Mpc [87].

Note that homogeneity does not imply isotropy, but isotropy around every point with analyticity does imply homogeneity [88]. Given we only see the Universe from one vantage point, we cannot be certain if our assumption of isotropy is universally valid. If the Universe's properties are described by analytic functions and isotropy is observed around every point, this analyticity can lead to homogeneity. This is because the smoothness of analytic functions combined with isotropy can enforce homogeneity. Therefore, if we were to experience isotropy at two locations, we could conclude that the Universe is both isotropic and homogeneous.

2.1 Friedmann-Robertson-Walker geometry

Fortunately, assuming the Universe is generally homogeneous and isotropic makes modelling it much easier. At large distances, we find that the Universe can even be approximated as a perfect fluid with uniform density, where galaxy clusters represent its fluid elements [45]. Assuming spatial homogeneity and isotropy allows us to represent the Universe as a sequence of maximally symmetric 3-spaces. Since these spaces are themselves isotropic and homogeneous, they have a constant 3-curvature – or Riemann curvature – which is either zero, positive, or negative. A Riemannian manifold has zero curvature if and only if it is flat. A zero curvature therefore represents the familiar three-dimensional Euclidean space $\mathbb{R}^3$ with line element (1.1). In two dimensions this would exist as an infinite plane. Positively and negatively curved space can be represented as a hypersphere S^3 in four-dimensional Euclidean space $\mathbb{R}^4$ or a hyperboloid H^3 in a four-dimensional Lorentzian space $\mathbb{R}^{1,3}$, respectively. In two dimensions, a surface of constant positive curvature is simply a sphere, while one of constant negative curvature is a hyperbolic paraboloid. If we consider two (initially) parallel lines on these surfaces, it is easy to visualise that in flat space, both will remain parallel; on a sphere, they will always cross; and on a hyperbolic paraboloid, they will always diverge. Homogeneity and isotropy are a little trickier to prove on these surfaces but are indeed satisfied.

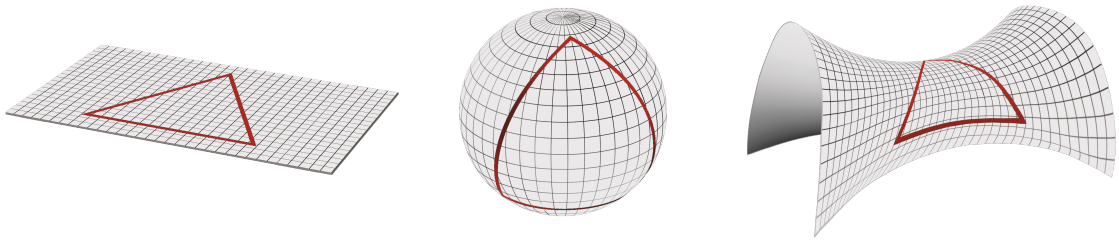


Figure 2.1: The three possible geometries in a Friedmann universe: Euclidian showing zero curvature i.e. flat (left), spherical showing positive curvature i.e. closed (centre), and hyperbolic showing negative curvature i.e. open (right). The red shapes demonstrate how the sum of the angles of a triangle differs in each case. Image from NASA.

2.1.1 The Robertson-Walker metric

Examining a homogeneous and isotropic space naturally leads us to the subject of *maximally symmetric spaces*. These are n-dimensional manifolds with $\frac{1}{2}n(n+1)$ Killing vectors. One example is n-dimensional Euclidean space $\mathbf{R}^n$; the isometries of this space are translations and rotations in n dimensions. The translations move points along n axes, giving n possible translations, while the rotations rotate one of n axes to or from $n-1$ other axes, giving $\frac{1}{2}n(n-1)$ total independent rotations. This gives

$$n + \frac{1}{2}n(n-1) = \frac{1}{2}n(n+1) \tag{2.2}$$

independent symmetries of $\mathbf{R}^n$. A space with the maximum number of Killing fields is a maximally symmetric space. Since the combination of homogeneity and isotropy implies a maximally symmetric space, it must also imply that the space has the maximum number of Killing vectors [43]. We can show that maximally symmetric spaces are defined as having a curvature tensor of the form

$$R_{ijkl} = \widetilde{K}(g_{ik}g_{jl} - g_{il}g_{jk})\,, \tag{2.3}$$

with a constant Ricci scalar R [44], where

$$\widetilde{K} \equiv \frac{R}{n(n-1)} \tag{2.4}$$

is the normalised Ricci curvature. The Ricci tensor is given by

$$\begin{aligned} R_{jk} = g^{il}R_{ijkl} &= \widetilde{K}g^{il}\left(g_{ik}g_{jl} - g_{il}g_{jk}\right) \\ &= \widetilde{K}\left(\delta^l_k g_{jl} - \delta^l_l g_{jk}\right) \\ &= \widetilde{K}\left(g_{jk} - 3g_{jk}\right) \\ &= -2\widetilde{K}g_{jk}\,. \end{aligned} \tag{2.5}$$

Maximally symmetric spaces are therefore spaces of constant curvature, and we can use this value $\widetilde{K}$ to define different maximally symmetric spacetimes. In this case, a maximally symmetric spacetime is one in which the spatial hypersurfaces at constant time are three-dimensional spaces of constant curvature, and the geometry evolves in time according to a scale factor. For $n = 4$, there are three such spacetimes [42]. A vanishing curvature ($\widetilde{K} = 0$) corresponds to Minkowski space with the familiar metric (1.5). A maximally symmetric spacetime with positive curvature is called *de Sitter* (dS) space, while a negative curvature corresponds to *anti-de Sitter* (AdS) space [43]. To construct metrics for these kinds of spaces, we start by rewriting the line element (1.9) as

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + g_{ij}\mathrm{d}x^i\mathrm{d}x^j\,. \tag{2.6}$$

To incorporate homogeneity and isotropy, we consider the squared spatial separation of two nearby galaxies at coordinates $(x^1,\, x^2,\, x^3)$ and $(x^1 + \Delta x^1,\, x^2 + \Delta x^2,\, x^3 + \Delta x^3)$ on the same hypersurface at time t:

$$\mathrm{d}\sigma^2 = g_{ij}\Delta x^i \Delta x^j \,, \tag{2.7}$$

where σ encompasses a three-dimensional space of uniform curvature. Isotropy requires that the triangle formed by these galaxies at a later time must be similar to the original triangle at t. Furthermore, homogeneity requires that the factor to which the triangle is scaled must be independent of the triangle's position in space. There must therefore be a common factor that ensures the ratios of small distances remain equal. This allows us to write the metric as

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + S^2(t)h_{ij}\mathrm{d}x^i\mathrm{d}x^j \,, \tag{2.8}$$

where $S(t)$ is the time-dependent scaling factor. We know that the spatial part of the general stationary isotropic metric is of the form

$$\mathrm{d}\sigma^2 = B(r)\mathrm{d}r^2 - r^2\mathrm{d}\boldsymbol{\Omega}^2 \,, \tag{2.9}$$

where $\mathrm{d}\boldsymbol{\Omega}$ is the solid angle such that $\mathrm{d}\boldsymbol{\Omega}^2 \equiv \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2$ and $B(r)$ is some arbitrary function of r. The only non-zero coefficients of the connection in spherical coordinates are

$$\Gamma^r{}_{rr} = \frac{1}{2B(r)}\frac{\mathrm{d}B(r)}{\mathrm{d}r} \,, \qquad \Gamma^r{}_{\theta\theta} = -\frac{r}{B(r)} \,, \qquad \Gamma^r{}_{\phi\phi} = -\frac{r\sin^2\theta}{B(r)} \,,$$
$$\Gamma^\theta{}_{r\theta} = \Gamma^\phi{}_{r\phi} = \frac{1}{r} \,, \qquad \Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta \,, \qquad \Gamma^\phi{}_{\phi\theta} = \cot\theta \,.$$

Using this with the combination of (1.47) and (1.51) to obtain the Ricci tensor in terms of the connection coefficients, we find that the non-zero components of the Ricci tensor are

$$R_{rr} = -\frac{1}{rB}\frac{\mathrm{d}B}{\mathrm{d}r} \,,$$
$$R_{\theta\theta} = \frac{1}{B} - 1 - \frac{r}{2B^2}\frac{\mathrm{d}B}{\mathrm{d}r} \,,$$
$$R_{\phi\phi} = R_{\theta\theta}\sin^2\theta \,.$$

Remember that the condition in (2.5) is necessary for the space to be maximally symmetric, so we require that

$$\frac{1}{rB}\frac{\mathrm{d}B}{\mathrm{d}r} = 2\widetilde{K}B(r) \,, \tag{2.10}$$
$$1 + \frac{r}{2B^2}\frac{\mathrm{d}B}{\mathrm{d}r} - \frac{1}{B} = 2\widetilde{K}r^2 \,. \tag{2.11}$$

We can integrate (2.10) to obtain

$$B(r) = \frac{1}{A - \widetilde{K}r^2} \,, \tag{2.12}$$

where A is a constant of integration. Substituting this into (2.11) gives

$$1 - A + \widetilde{K}r^2 = \widetilde{K}r^2 \implies A = 1 \,. \tag{2.13}$$

The line element for the maximally symmetric 3-space is therefore

$$\mathrm{d}\sigma^2 = \frac{\mathrm{d}r^2}{1 - \widetilde{K}r^2} + r^2\mathrm{d}\boldsymbol{\Omega} \,, \tag{2.14}$$

with the value of $\widetilde{K}$ describing the curvature of space. The origin of the radial coordinate r is entirely arbitrary, allowing us to choose any point in space as the origin. We also note that for $r \ll \widetilde{K}$, (2.14) reduces to three-dimensional Euclidean space (1.1). Combining expressions (2.8) and (2.14) then gives us

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + S^2(t)\left[\frac{\mathrm{d}r^2}{1-\widetilde{K}r^2} + r^2\mathrm{d}\boldsymbol{\Omega}\right]. \tag{2.15}$$

The magnitude of $\tilde{K}$ here can take on any arbitrary value so ideally, we would like to absorb this into the radial coordinate and scale factor. Assuming that $\widetilde{K} \neq 0$, we can define $K = \widetilde{K}/|\widetilde{K}|$ such that K only takes on values of ± 1. If we let $\bar{r} \equiv \sqrt{|\widetilde{K}|}r$ then

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + \frac{S^2(t)}{|\widetilde{K}|}\left[\frac{\mathrm{d}\bar{r}^2}{1-\widetilde{K}\bar{r}^2} + \bar{r}^2\mathrm{d}\boldsymbol{\Omega}\right],$$

and we can define a rescaled scale function $R(t)$ as

$$R(t) = \begin{cases} \dfrac{S(t)}{\sqrt{|\widetilde{K}|}} & \text{if} \quad \widetilde{K} \neq 0\,, \\ S(t) & \text{if} \quad \widetilde{K} = 0\,. \end{cases} \tag{2.16}$$

This leads to the Friemann-Lemaître-Robertson-Walker (FLRW) metric[2]:

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + R^2(t)\left[\frac{\mathrm{d}\bar{r}^2}{1-K\bar{r}^2} + \bar{r}^2\mathrm{d}\boldsymbol{\Omega}\right], \quad K = \begin{cases} -1 & \text{open} \\ 0 & \text{flat} \\ +1 & \text{closed} \end{cases}. \tag{2.17}$$

In this case, K can only belong to the set $\{-1, 0, +1\}$, corresponding to constant negative curvature, zero curvature, and constant positive curvature, respectively. This leaves $\bar{r}$ unitless and $R(t)$ with units of length. We can therefore think of $R(t)$ as the size scale of the Universe. A constant $R(t)$ represents a static universe, while $\dot{R}(t) > 0$ and $\dot{R}(t) < 0$ represent an expanding and contracting universe, respectively. The value of the parameter K provides three distinct geometries. For $K = 0$, the metric is spatially flat and the distance between points is governed by the evolution of the scale factor. Notice that for $K = +1$, the metric has a coordinate singularity at $\bar{r} = 1$ that requires a change of coordinates. This can be achieved by defining a new radial coordinate χ with

$$\mathrm{d}\chi = \frac{\mathrm{d}\bar{r}}{\sqrt{1-K\bar{r}^2}}\,, \tag{2.18}$$

which can be integrated to give

$$\bar{r} = S_k(\chi), \quad S_k = \begin{cases} \sin\chi\,, & K = +1 \\ \chi\,, & K = 0 \\ \sinh\chi\,, & K = -1. \end{cases} \tag{2.19}$$

This allows us to simplify the metric (2.17) to

$$\mathrm{d}\sigma^2 = \mathrm{d}\chi^2 + S_k(\chi)\mathrm{d}\boldsymbol{\Omega}^2\,, \tag{2.20}$$

[2]Also known as the Friedmann-Robertson-Walker metric or Robertson-Walker metric.

and examine each curvature case more easily. For $K = 0$, the metric is simply flat Euclidean space:

$$\begin{aligned} \mathrm{d}\sigma^2 &= \mathrm{d}\chi^2 + \chi^2 \mathrm{d}\mathbf{\Omega}^2 && (2.21)\\ &= \mathrm{d}\bar{r}^2 + \bar{r}^2 \mathrm{d}\theta^2 + \bar{r}^2 \sin^2\theta \mathrm{d}\phi^2 && (2.22)\\ &= \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \,. && (2.23) \end{aligned}$$

For $K = +1$ and $K = -1$, we obtain the metrics for a 3-sphere and a hyperboloid, respectively. Looking at figure 2.1, we can see that this change of coordinates makes the physical interpretation more intelligible.

Before moving on to the Friedmann equations, it will be advantageous to redefine the FLRW metric. If we introduce a dimensionless scale factor

$$a(t) = \frac{R(t)}{R_0}, \tag{2.24}$$

we can obtain an alternative form of the **FLRW metric**:

$$\boxed{\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)\left[\frac{\mathrm{d}r^2}{1-\bar{K}r^2} + r^2\mathrm{d}\mathbf{\Omega}\right],} \tag{2.25}$$

where the radial component $r = R_0\bar{r}$ has units of length and the constant $\bar{K} = K/R_0^2$ represents the curvature of space when $a(t) = 1$ in units of length^{-2}. We generally set $a(t) = 1$ in the current cosmological era, making $\mathrm{d}\sigma$ equal to the *comoving distance*[3]

$$\chi_{\mathrm{H}}(t) = \frac{d_{\mathrm{H}}(t)}{R(t)} = \frac{c}{H(t)R(t)} = \frac{c}{\dot{R}(t)}\,, \tag{2.26}$$

where we have introduced the **Hubble distance**

$$\boxed{d_{\mathrm{H}}(t) = cH^{-1}(t)\,.} \tag{2.27}$$

The comoving distance is then clearly the χ-coordinate corresponding to this Hubble distance. This will be described in more detail in §2.2.1. We're almost done with the metric; we just need to evolve it to its final form to match the conventions we'll be using for the rest of the thesis. We would like a dimensionless scale factor *and* a curvature that takes on the values $\{-1, 0, +1\}$, so we need to reintroduce a length^2 factor somewhere. With a dimensionless radial component (dropping the bar for simplicity), we can bring back the scale constant R_0 to keep the units consistent:

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)R_0^2\left[\frac{\mathrm{d}r^2}{1-Kr^2} + r^2\mathrm{d}\mathbf{\Omega}\right]. \tag{2.28}$$

The advantage of the dimensionless scale factor is that it can be related very simply to the redshift via

$$\boxed{z = \frac{a_0}{a(t)} - 1 \implies a(t) = \frac{1}{1+z}\,,} \tag{2.29}$$

where $a_0 = a(t_0)$ and t_0 is the present age of the Universe, $13.799 \pm 0.021 \times 10^9$ yr [24]. In the second expression above, we have adopted the common convention of normalising the present-day scale factor such that $a(t_0) = 1$, which will be used henceforth. The benefit of using cosmic

[3] Also known as the comoving Hubble distance.

time t here is that it deals with familiar quantities and allows us to discern the distances that light has travelled. The main drawback is that the speed of light does not take on the value of the local speed of light. This is because distances between points expand in an expanding universe, so the observed rate at which light travels across these expanding distances (coordinate speed) can appear different from the constant speed c that we would measure locally in a static region. To alleviate the issue, we introduce *conformal time*,

$$\mathrm{d}\eta = \frac{\mathrm{d}t}{a(t)}\,, \tag{2.30}$$

which is useful if we want to study the propagation of light in particular. With this and (2.18) we can write

$$\mathrm{d}s^2 = c^2a^2(\eta)\left[-\mathrm{d}\eta^2 + \left(\mathrm{d}\chi^2 + S_k^2(\chi)d\Omega^2\right)\right]. \tag{2.31}$$

Having factorised the scale factor, this has become the static Minkowski metric multiplied by a time-dependent conformal factor $a(\eta)$. For a null geodesic, $\mathrm{d}s^2 = 0$, and the change in conformal time along the path equals the comoving distance,

$$\Delta\eta = \Delta\chi\,. \tag{2.32}$$

We can see from the above how our choice of coordinate system can make certain cases much more convenient to work with. However, for now we will continue to make use of cosmic time to derive the Friedmann equations.

2.1.2 The Friedmann equations

We can now use the metric (2.25) and the Einstein field equations (1.83) to derive the Friedmann equations. The covariant components $g_{\mu\nu}$ of (2.25) are

$$g_{00} = c^2, \quad g_{11} = -\frac{a^2(t)}{1-Kr^2}, \quad g_{22} = -a^2(t)r^2, \quad g_{33} = -a^2(t)r^2\sin^2\theta\,.$$

The metric is diagonal, so the contravariant components $g^{\mu\nu}$ are just the reciprocals of the covariant components, and the connection is given by the equations in (1.18):

$$\Gamma^{\sigma}{}_{\mu\nu} = \frac{1}{2}g^{\sigma\rho}\left(\partial_\nu g_{\rho\mu} + \partial_\mu g_{\rho\nu} - \partial_\rho g_{\mu\nu}\right).$$

The only non-zero components of this expression are

$$\begin{array}{lll}
\Gamma^0{}_{11} = a\dot{a}/\left[c\left(1-Kr^2\right)\right], & \Gamma^0{}_{22} = a\dot{a}r^2/c\,, & \Gamma^0{}_{33} = \left(a\dot{a}r^2\sin^2\theta\right)/c\,, \\
\Gamma^1{}_{01} = c\dot{a}/a\,, & \Gamma^1{}_{11} = Kr/\left(1-Kr^2\right), & \Gamma^1{}_{33} = -r\left(1-Kr^2\right)\sin^2\theta\,, \\
\Gamma^2{}_{02} = c\dot{a}/a\,, & \Gamma^2{}_{12} = 1/r\,, & \Gamma^2{}_{33} = \sin\theta\cos\theta\,, \\
\Gamma^3{}_{03} = c\dot{a}/a\,, & \Gamma^3{}_{13} = 1/r\,, & \Gamma^3{}_{33} = \cot\theta\,,
\end{array} \tag{2.33}$$

where $\dot{a} = \frac{\mathrm{d}a}{\mathrm{d}t}$. We substitute these into an expression for the Ricci tensor[4]:

$$R_{\mu\nu} = \partial_\nu\Gamma^{\sigma}{}_{\mu\sigma} - \partial_\sigma\Gamma^{\sigma}{}_{\mu\nu} + \Gamma^{\rho}{}_{\mu\sigma}\Gamma^{\sigma}{}_{\rho\nu} - \Gamma^{\rho}{}_{\mu\nu}\Gamma^{\sigma}{}_{\rho\sigma}\,, \tag{2.34}$$

to obtain the non-zero components of the Ricci tensor – in this case the diagonal components:

$$\begin{aligned}
R_{00} &= \frac{3\ddot{a}}{c^2a}\,, \\
R_{11} &= -\frac{a\ddot{a}+2\dot{a}^2+2c^2K}{c^2}\frac{1}{1-Kr^2}\,, \\
R_{22} &= -\frac{a\ddot{a}+2\dot{a}^2+2c^2K}{c^2}r^2\,, \\
R_{33} &= -\frac{a\ddot{a}+2\dot{a}^2+2c^2K}{c^2}r^2\sin^2\theta\,.
\end{aligned}$$

[4]Obtained by combining (1.47) and (1.51).

The scalar curvature is therefore given by

$$R = g^{\mu\nu}R_{\mu\nu} = \frac{6}{c^2}\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{Kc^2}{a^2}\right]. \tag{2.35}$$

The Einstein field equations (1.83) indicate that the value of the scale factor depends on the matter and energy content of the Universe. If we choose to model matter and energy as a perfect fluid, we can use the energy-momentum tensor from (1.56). Since we are working in comoving coordinates, the fluid is at rest and its four-velocity is

$$U^\mu = (c, 0, 0, 0)\,. \tag{2.36}$$

This allows the stress tensor to be written as

$$T^\mu{}_\nu = \mathrm{diag}(-\rho c^2, p, p, p)\,, \tag{2.37}$$

with trace

$$T = T^\mu{}_\mu = -\rho c^2 + 3p\,. \tag{2.38}$$

We now have all the necessary information to find the components of the Einstein field equations. For $\mu, \nu = 0$ we have

$$\begin{aligned} R_{00} &= \frac{8\pi G}{c^4}\left(T_{00} + \frac{1}{2}Tg_{00}\right) + \Lambda g_{00} \\ \therefore -3\frac{\ddot{a}}{a} &= 4\pi G\left(\rho + 3\frac{p}{c^2}\right) + \Lambda c^2\,, \end{aligned} \tag{2.39}$$

and for $\mu, \nu = \sigma$[5] we have

$$\begin{aligned} R_{\sigma\sigma} &= \frac{8\pi G}{c^4}\left(T_{\sigma\sigma} - \frac{1}{2}Tg_{\sigma\sigma}\right) + \Lambda g_{\sigma\sigma} \\ \therefore a\ddot{a} + 2\dot{a}^2 + 2Kc^2 &= 4\pi G\left(\rho - \frac{p}{c^2}\right)a^2 + \Lambda c^2 a^2\,. \end{aligned} \tag{2.40}$$

These are the only independent equations due to the homogeneity and isotropy of the FLRW metric. If we clean up (2.39) and then substitute it into (2.40), we obtain the **Friedmann equations**

$$\boxed{\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{1}{3}\Lambda c^2\,,} \tag{2.41a}$$

$$\boxed{\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{1}{3}\Lambda c^2 - \frac{K}{a^2}c^2\,.} \tag{2.41b}$$

We can derive another useful equation using the conservation of energy-momentum (1.57). The $\nu = 0$ component of (2.37) gives us the continuity equation

$$\begin{aligned} 0 &= \nabla_\mu T^\mu{}_0 \\ &= \partial_\mu T^\mu{}_0 + \Gamma^\mu{}_{\mu\lambda}T^\lambda{}_0 - \Gamma^\lambda{}_{\mu 0}T^\mu{}_\lambda \\ &= -c\dot{\rho} - \frac{3}{c}\frac{\dot{a}}{a}\left(\rho c^2 + p\right). \end{aligned} \tag{2.42}$$

[5]The three spatial field equations are shown to be equivalent when combined with the Ricci tensor.

Substituting the equation of state equation (1.60) gives the **energy conservation equation**:

$$\dot{\rho} + 3\rho(1+w)\frac{\dot{a}}{a} = 0 \implies \dot{\rho} = -3H\rho(1+w)\,, \tag{2.43}$$

with a solution of the form

$$\rho \propto a^{-3(1+w)}\,. \tag{2.44}$$

2.1.3 Cosmological parameters

In a general cosmological model, the total energy density of the Universe is comprised of matter, radiation, and vacuum energy densities. The energy density of the vacuum arises from the cosmological constant term, and can also be modelled as a perfect fluid. The total density is then a sum of these components:

$$\rho(t) = \rho_{\mathrm{m}}(t) + \rho_{\mathrm{r}}(t) + \rho_{\Lambda}(t)\,, \tag{2.45}$$

where the subscripts m, r, and Λ represent the equivalent mass densities of matter, radiation and the vacuum, respectively. All of these components have an equation of state of the form (1.60), which we can write as

$$p_i = w_i \rho_i c^2\,, \tag{2.46}$$

where i is the component label. The requirements on pressure and density from the weak energy condition (1.58) means that the vacuum energy can either be positive or negative and therefore $|w| \leq 1$ [43].

The choice of w leads us to several interesting cases. The simplest is that of the **matter-dominated universe**, where $p = 0$ and hence $w = 0$. This is also known as the dust case, where dust describes collisionless, non-relativistic particles with thermal pressure much less than their rest masses, and zero pressure. This includes everyday baryonic matter and possibly more exotic forms of matter beyond the Standard Model, such as dark matter. Both of these contribute to the total matter density:

$$\rho_{\mathrm{m}}(t) = \rho_{\mathrm{b}}(t) + \rho_{\mathrm{dm}}(t)\,. \tag{2.47}$$

We do not need to differentiate between these types, as the scale factor only depends on the total matter density. Since dust is considered to be pressureless, the energy density results solely from the rest masses of the particles, and $\rho_{\mathrm{m}} \propto a^{-3}$. We expect this since the mass density should be inversely proportional to the volume the mass occupies. We can write this explicitly as

$$\rho_{\mathrm{m}}(a) = \rho_{\mathrm{m},0} a^{-3} \quad \text{or} \quad \rho_{\mathrm{m}}(z) = \rho_{\mathrm{m},0}(1+z)^3\,. \tag{2.48}$$

A **radiation-dominated universe** primarily consists of photons[6] and any particles travelling close to lightspeed (with small or zero rest masses), such as neutrinos:

$$\rho_{\mathrm{r}}(t) = \rho_{\gamma}(t) + \rho_{\nu}(t)\,. \tag{2.49}$$

The energy-momentum tensor vanishes for an electromagnetic field, but since we are also assuming a perfect fluid, we must have $T^{\mu}{}_{\mu} = -\rho c^2 + 3p = 0$, and therefore $w = 1/3$, giving $\rho_{\mathrm{r}} \propto a^{-4}$.

[6] The dominant contribution to the radiation energy density is from the photons in the CMB. As we discussed in the introduction, the radiation is highly uniform with an almost perfect blackbody distribution.

This is because, as well as being inversely proportional to the volume, the particles are redshifted by $1/a$ as the Universe expands. We can show the density's dependence on redshift with

$$\rho_{\mathrm{r}}(a) = \rho_{\mathrm{r},0} a^{-4} \quad \text{or} \quad \rho_{\mathrm{r}}(z) = \rho_{\mathrm{r},0}(1+z)^4 \,. \tag{2.50}$$

In a **vacuum-dominated universe**, the equation of state is simply $p = -c^2\rho$, making $w = -1$ and

$$\rho_{\mathrm{v}} \propto \text{constant} \,, \tag{2.51}$$

because the vacuum energy is unchanging [43]. We can show it remains constant at

$$\rho_\Lambda = \rho_{\Lambda,0} = \frac{\Lambda c^2}{8\pi G} \,. \tag{2.52}$$

If we combine the above results to account for contributions from every component, the function for the **total equivalent mass density** becomes

$$\rho(a) = \rho_{\mathrm{m},0} a^{-3} + \rho_{\mathrm{r},0} a^{-4} + \rho_{\Lambda,0} \,. \tag{2.53}$$

It is clear from (2.53) that the relative contributions from matter, radiation, and the vacuum fluctuate as the Universe evolves. We will expand on these different models in §2.1.4.

Having introduced the scale factor, we can now define the **Hubble parameter**,

$$H = \frac{\dot{a}}{a} \,, \tag{2.54}$$

which originated with the Hubble-Lemaître law[7] after Edwin Hubble observed the expansion of the Universe in 1929 [9]. This law describes the rate at which points in the Universe are receding away from us:

$$v = H(t) D \,, \tag{2.55}$$

where v is the speed of expansion in km s^{-1}, D is the distance to the point in Mpc, and $H(t)$ is the time-dependent Hubble parameter. It has been famously difficult to constrain the current value of the Hubble parameter, Hubble's constant H_0. Its current value is believed to be in the region of $H_0 \sim 70$ km s^{-1} Mpc^{-1} [89, 90]. It is conventional to account for this uncertainty in H_0 by defining a dimensionless parameter h such that $H_0 \equiv 100h$ km s^{-1} Mpc^{-1} and $h \sim 0.7$. We can also define the *Hubble length*, $d_H = H_0^{-1} c$ and the *Hubble time*, $t_H = H_0^{-1}$. In general, and as we will do in later chapters, c is set to one, and H_0^{-1} is referred to as both the Hubble length and the Hubble time.

The Hubble parameter measures the expansion rate at time t for any model obeying the Cosmological Principle but is still dependent on the matter/energy content of the Universe [66]. We can probe further by Taylor expanding the scale factor at times close to t_0:

$$a(t) = a_0 \left[1 + H_0(t - t_0) - \frac{1}{2} q_0 H_0^2 (t - t_0)^2 + \ldots \right] , \tag{2.56}$$

[7] More commonly referred to as Hubble's law.

where

$$q_0 = -\frac{\ddot{a}(t_0)a_0}{\dot{a}^2(t_0)} \tag{2.57}$$

is the deceleration parameter $q_0 = q(t_0)$. Since $\dot{H} = \frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2$, it follows that

$$q(t) = -\left(\frac{\dot{H}}{H^2} + 1\right). \tag{2.58}$$

We can similarly extend this to higher terms, giving us four cosmographic parameters, **deceleration**, **jerk**, **snap**, and **lerk** [8] [91] [92]:

$$\boxed{q \equiv -\frac{1}{aH^2}\frac{\mathrm{d}^2 a}{\mathrm{d}t^2} \quad , \quad j \equiv \frac{1}{aH^3}\frac{\mathrm{d}^3 a}{\mathrm{d}t^3} \quad , \quad s \equiv \frac{1}{aH^4}\frac{\mathrm{d}^4 a}{\mathrm{d}t^4} \quad , \quad l \equiv \frac{1}{aH^5}\frac{\mathrm{d}^5 a}{\mathrm{d}t^5}\,.} \tag{2.59}$$

This result will prove very useful in §4.2, and enables us to write (2.56) as

$$\begin{aligned} a(t) \approx a_0 \Big[1 + H_0(t-t_0) - \frac{1}{2}q_0 H_0^2 (t-t_0)^2 + \frac{1}{6} j_0 H_0^2 (t-t_0)^3 \\ + \frac{1}{24} s_0 H_0^2 (t-t_0)^4 + \frac{1}{120} l_0 H_0^2 (t-t_0)^5 \Big]\,. \end{aligned} \tag{2.60}$$

The choice of defining q as negative is deliberate. It shows that the Hubble parameter decreases with time unless $q < -1$. A positive value of q corresponds to $\ddot{a} < 0$, meaning that the Universe's expansion is decelerating. A negative value implies $\ddot{a} > 0$ and an accelerating expansion [93]. Observations of Type Ia supernovae [14, 94], as well as low redshift observations on the assumption of almost flat spatial sections – which we can justify with CMB data – have determined q to be negative, implying that the expansion of the Universe is accelerating in the current epoch. The rate of change of the scale factor must therefore be positive.

By truncating the expansion at fourth order, we assume that higher-order terms have a negligible impact over the redshift range covered by the data. As noted by Capozziello and Francaviglia [95], these higher-order terms effectively approximate modifications to the field equations. For instance, an R^2 term leads to a fourth-order equation, while an $R\Box R$ term results in a sixth-order equation, and so on. Through a conformal transformation, each second-order derivative corresponds to an additional scalar degree of freedom, meaning that fourth-order gravity can be interpreted as general relativity with an extra scalar field. In principle, this process could generate an infinite hierarchy of equations. However, taking the third derivative of the modified Friedmann equation expresses the result entirely in terms of cosmographic parameters up to the lerk, eliminating the need for higher derivatives. In special cases, such as spatially flat or vacuum cosmologies, the third derivative is unnecessary, and the snap parameter can be omitted. This applies to the spatially flat ΛCDM model, which satisfies the well-known cosmographic condition $j = 1$ [96, 97, 98].

We have been working with the equivalent mass densities thus far, but a more convenient and commonly used quantity is the dimensionless **density parameter**:

$$\boxed{\Omega_i(t) \equiv \frac{8\pi G}{3H^2(t)}\rho_i(t)\,,} \tag{2.61}$$

[8] Lerk (l) is referred to as crackle (c) in some texts.

the derivative of which can be easily calculated using (2.43):

$$\begin{aligned}\dot{\Omega}_i(t) &= \frac{8\pi G}{3H^2}\left(\dot{\rho}_i - \frac{2\dot{H}}{H}\rho_i\right) \\ &= -\Omega_i H\left(3(1+w_i) + \frac{2\dot{H}}{H^2}\right).\end{aligned} \tag{2.62}$$

Note that due to the inclusion of the Hubble parameter $H(t)$, the parameter Ω_Λ is generally a function of time – as opposed to the constant ρ_Λ. The density parameter can also be expressed in terms of the critical density via $\Omega = \rho/\rho_{\text{crit}}$, where

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}\,. \tag{2.63}$$

Equation (2.61) allows us to fix our cosmological model by specifying four, present-day values: H_0, $\Omega_{\text{m},0}$, $\Omega_{\text{r},0}$, and $\Omega_{\Lambda,0}$. As such, constraining these quantities is one of the main goals of observational cosmology. Researchers from the Planck and DESI collaborations [24, 99], among others, have managed to achieve this to a few per cent accuracy. Here, we give them approximately:

$$\boxed{H_0 \approx 70\ \text{km s}^{-1}\,\text{Mpc}^{-1}\,,\quad \Omega_{\text{m},0} \approx 0.3\,,\quad \Omega_{\text{r},0} \approx 5\times10^{-5}\,,\quad \Omega_{\Lambda,0} \approx 0.7\,.} \tag{2.64}$$

We can see that the energy density of the current epoch is dominated by the cosmological constant. Even more intriguing is that only one-sixth of the matter density is occupied by baryonic matter. The Hubble parameter as a function of redshift using these parameters is

$$H^2(z) = H_0^2\left[\Omega_{\text{m},0}(1+z)^3 + \Omega_{\text{r},0}(1+z)^4 + \Omega_{\Lambda,0} + \Omega_{k,0}(1+z)^2\right]. \tag{2.65}$$

The convenience of using the density parameters becomes clear when we substitute them into the Friedmann equations (2.40), obtaining

$$1 = \Omega_{\text{m}} + \Omega_{\text{r}} + \Omega_\Lambda - \frac{c^2K}{H^2a^2}\,. \tag{2.66}$$

The extra term at the end is defined as the curvature density parameter:

$$\Omega_k(t) = -\frac{c^2K}{H^2(t)a^2(t)}\,, \tag{2.67}$$

guaranteeing that, at all times t, we have the satisfying relation

$$\Omega_{\text{m}} + \Omega_{\text{r}} + \Omega_\Lambda + \Omega_k = 1\,. \tag{2.68}$$

The value of $\Omega \equiv \Omega_{\text{m}} + \Omega_{\text{r}} + \Omega_\Lambda$ therefore determines the spatial curvature of the Universe. We can summarise the three cases as

$$\begin{array}{lllllllll}
\Omega < 1 & \leftrightarrow & \rho < \rho_{\text{crit}} & \leftrightarrow & \bar{K} < 0 & \leftrightarrow & K = -1 & \leftrightarrow & \text{open} \\
\Omega = 1 & \leftrightarrow & \rho = \rho_{\text{crit}} & \leftrightarrow & \bar{K} = 0 & \leftrightarrow & K = 0 & \leftrightarrow & \text{flat} \\
\Omega > 1 & \leftrightarrow & \rho > \rho_{\text{crit}} & \leftrightarrow & \bar{K} > 0 & \leftrightarrow & K = +1 & \leftrightarrow & \text{closed}\,.
\end{array}$$

To obtain an analytic expression for the spatial curvature in terms of redshift, we can evaluate (2.67) at $t = t_0$, using the relation $R_0/R = 1+z$ to obtain the general formula

$$\Omega_k(z) = \left[\frac{H_0(1+z)}{H(z)}\right]^2 \Omega_{k,0}\,. \tag{2.69}$$

We can fix the cosmological model by specifying the values of the quantities in (2.64), so it is worthwhile to rewrite the field equations (1.83) in terms of these quantities [44]. The second field equation may be written

$$H^2 = \frac{8\pi G}{3}\left(\sum_i \rho_i\right) - \frac{c^2 K}{a^2}\,. \tag{2.70}$$

Combining this with equations (2.53), (2.61) and (2.67) gives

$$H^2 = H_0^2\left(\Omega_{\mathrm{r},0}a^{-4} + \Omega_{\mathrm{m},0}a^{-3} + \Omega_{k,0}a^{-2} + \Omega_{\Lambda,0}\right), \tag{2.71}$$

which can also be written in terms of the redshift using the relation (2.29). The second Friedmann equation (2.41b) can be multiplied through by $a/\dot{a}^2$ to obtain

$$\frac{a\ddot{a}}{\dot{a}^2} = -\frac{4\pi G}{3H^2}\sum_i \rho_i(1+3w_i)\,.$$

Looking at the left hand side, we can see that this is simply the deceleration parameter. Therefore, by substitung an appropriate w_i value for each component, we obtain an expression for q in terms of the dimensionless parameters:

$$q = \frac{1}{2}(\Omega_{\mathrm{m}} + 2\Omega_{\mathrm{r}} - 2\Omega_\Lambda)\,. \tag{2.72}$$

This result provides another neat expression for the derivative of the density parameter (2.62),

$$\dot{\Omega}_i = \Omega_i H(\Omega_{\mathrm{m}} + 2\Omega_{\mathrm{r}} - 2\Omega_\Lambda - 1 - 3w_i)\,. \tag{2.73}$$

Setting w_i equal to 0, 1/3 or -1 then gives the evolution of the density parameters for matter, radiation, and the vacuum, respectively. If we want to study the behaviour of the spatial curvature, we need to differentiate (2.67) with respect to cosmic time to obtain

$$\begin{aligned}\dot{\Omega}_{\mathrm{k}} &= 2\Omega_k H q \\ &= \Omega_k H(\Omega_{\mathrm{m}} + 2\Omega_{\mathrm{r}} - 2\Omega_\Lambda)\,.\end{aligned}$$

We can see that this expression remains positive if the contribution from Λ is zero. This means that the spatial curvature rapidly evolves away from flat if Ω_k differs even slightly from zero: $\Omega_k \to 1$ for an open universe and $\Omega_k \to -\infty$ for a closed one. However, the presence of a cosmological constant, i.e. a contribution from Ω_Λ alters the evolution completely. In this case the $2\Omega_\Lambda$ term will eventually dominate both the matter and radiation terms at some finite time and cause Ω_k to retreat back to zero. Even if the present-day value $\Omega_{k,0}$ is nowhere near zero, Ω_k must have been extremely close to zero in the distant past (at high redshifts). Since the value today is ~ 1, it must have been very finely tuned to zero in early epochs. While the current spatial flatness is explained somewhat by a positive cosmological model, this particular 'fine-tuning' problem requires new methods that we will consider in §2.2.

2.1.4 Cosmological models

The various choices of density parameters leads to a great number of possible analytical models, of which we cover the most notable. The Friedmann models in particular describe a universe that obeys the cosmological principle of homogeneity and isotropy. They have a zero cosmological constant and a non-zero matter or radiation density. As such they allow for three different types of geometry: flat, open, or closed, depending on the total energy density of the Universe. The

flat, open and closed universes correspond to curvature parameters that are positive, zero, and negative, respectively [93]. For a dust-only case ($\Omega_{\mathrm{r},0}, \Omega_{\Lambda,0} = 0$) with $\Omega_{\mathrm{m}}, K = 0$ we have the *Einstein-de-Sitter* (EdS) model. Energy conservation (2.43) gives $\rho = \rho_0 a^{-3}$ and a Friedmann equation satisfying

$$3H^2 = 3\left(\frac{\dot{a}}{a}\right)^2 = 8\pi G\rho_0 a^{-3}\,,$$
$$\therefore \dot{a} = 8\pi G H_0 a^{-1/2}\,,$$

where we wrote the density as $\rho_0 = 3H_0^2$. This gives the solution

$$a(t) = (12\pi G H_0 t)^{2/3}\,. \tag{2.74}$$

The age of this particular universe is then $t_0 \sim 9 \times 10^9$ yr.

We can also consider a universe without matter or radiation, called the *de Sitter* model. Proposed by Willem de Sitter [100], this model describes an empty universe dominated by dark energy, where the expansion of space accelerates over time. The model is a type of Lemaitre model, defined by $\Omega_{\mathrm{m},0}, \Omega_{\mathrm{r},0} = 0$ and $\Omega_{\Lambda,0} = 1$. Since it assumes that the matter and radiation densities are zero, it is not a true cosmological model. Regardless, it is an important model for understanding the late-time evolution of the Universe and the effects of dark energy. With $K, \rho = 0$, the Friedmann equation is simply

$$3H^2 = \Lambda c^2 \implies \frac{\dot{a}}{a} = \sqrt{\frac{\Lambda}{3}}c\,, \tag{2.75}$$

and the scale factor increases exponentially as

$$a(t) = \exp\left[\sqrt{\Lambda/3}c(t - t_0)\right], \tag{2.76}$$

giving us an inflating universe. Interestingly, this model lacks a big-bang singularity ($a = 0$) at a finite time in the past.

As discussed in §1.3, Einstein initially proposed a static and unchanging universe, with matter distributed uniformly throughout space. Now that we know the Universe is expanding, this model is of little empirical interest, but still of interest to theorists, particularly those investigating emergent universe models [101]. The Friedmann equation for the *Einstein static* universe, with a constant and $H = 0$, is

$$\frac{8\pi G}{c^2}\rho + \Lambda - \frac{3K}{a^2} = 0 \quad , \quad \Lambda = 4\pi G\left(\frac{\rho}{c^2} + 3p\right) > 0\,. \tag{2.77}$$

We can see that the spatial curvature must be positive, giving us a closed universe with Λ and standard matter existing in a state of equality.

Finally, we have the *Milne model*, a simplified cosmological model that assumes no gravitational interactions between objects in the Universe. It's also known as the 'kinematic' or 'empty' universe model. While obviously not an accurate description of our universe, it serves as a useful theoretical experiment. With $\rho = 0$, the Friedman equation is simply $H^2 = -Kc^2/a^2$, so this model has a negative curvature. For $K = -1$ we get $\dot{a} = c$ and a coasting solution $a(t) = ct$. This universe expands linearly with time without the influence of forces like gravity or dark energy slowing it down or accelerating it [66].

The cosmological constant Λ has had a convoluted history since Einstein introduced it to counteract the gravitational attraction of matter. Its property of universal spatial repulsion allowed for the *Einstein static* universe, which has a positive spatial curvature satisfied by the $\rho_\Lambda = \rho_\mathrm{m}/2$ [102, 103]. It was used much later to address the following issues:

- Globular cluster estimates of the age of the Universe conflict with those obtained for $\Lambda = 0$ [104, 105].
- Inflation implies a vanishing curvature $K = 0$ that requires $\Omega_0 + \frac{\Lambda}{3H_0^2} = 1$ [106]. With the most recent observations pointing towards $\Omega_0 \sim 0.3$ [107, 99], Λ would need to be non-zero to agree with observations.
- Number counts of galaxies at high redshifts are inconsistent with a density parameter $\Omega = 1$ that may be avoided by decreasing Ω or with a positive Λ value [108, 109].
- The standard CDM model [110, 111], with $\Omega = 1$, predicts significantly reduced large-scale structure than indicated by observations.

In the second point we have introduced the term *inflation*. This is an intensely-studied subbranch in cosmology with a relative abundance of information in the literature. In the next section it is briefly summarised.

2.2 Inflation

FLRW cosmology predicts an early universe consisting of causally disjointed regions of space, so how did the Universe end up homogeneous and isotropic? In the 1980s, cosmologists developed the idea of inflation [112, 16, 18], a period of accelerated expansion in the history of the Universe during which its initial inhomogeneities were smoothed out [66]. The basic requirement for this is that $\ddot{a} > 0$. We can rework this condition by taking the cosmic time derivative of the Hubble distance (2.27) to get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{H^{-1}}{a}\right) < 0\,. \tag{2.78}$$

This gives a more physically meaningful condition for inflation: the comoving Hubble distance decreases with cosmic time. This means that, from a comoving perspective, the characteristic length scale of the Universe contracts during inflation.

If we suppose that the energy in the early universe was dominated by some form of matter with density ρ and pressure p. The field equation is therefore that of an Einstein static universe,

$$\ddot{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right)a\,. \tag{2.79}$$

Therefore, in order for inflation to occur, we require the matter fluid to satisfy

$$p < -\frac{c^2}{3}\rho\,. \tag{2.80}$$

In other words, the fluid needs to have an equation of state with negative pressure. This is also a solution to the flatness problem, which we will expand on later. We can take the same approach ($\Lambda = 0$) with the first Friedmann equation (2.41a), obtaining

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - c^2 K\,. \tag{2.81}$$

This means that shortly after the Big Bang ($\sim 10^{-35}$ s), the cosmological constant must have been totally dominant, allowing the Universe to expand according to

$$\frac{\dot{a}}{a} = \left(\frac{8\pi\rho_\Lambda}{3}\right)^{1/2} . \tag{2.82}$$

Were this to last long enough, a tiny region could conceivably expand to the size of the current observable universe. If the initial region had been homogenised somehow prior to inflation, it would explain the current homogeneity. Inflation also allows us to resolve two glaring issues: the flatness problem and the horizon problem.

2.2.1 The flatness and horizon problems

The *flatness problem* stems from studying the Friedmann equation in a universe devoid of vacuum energy. According to (2.70), its Friedmann equation is therefore

$$1 = \Omega - \frac{c^2 K}{H^2 a^2}, \quad \text{where} \quad \Omega = \Omega_\mathrm{m} + \Omega_\mathrm{r} = \frac{8\pi G\rho}{3H^2} , \tag{2.83}$$

where we have neglected the contribution from Λ. If we consider the derivative of Ω with respect to $\ln a$, and substitute the Friedmann equation, we obtain

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\ln a} = \Omega(1+3w)(\Omega - 1) . \tag{2.84}$$

While Ω is a solution here, it turns out to be unstable. We can show this by perturbing the solution by an infinitesimal amount ε, i.e. $\Omega(t) = 1 \pm \varepsilon(t)$. If we neglect higher orders of ε, (2.84) becomes

$$\frac{\mathrm{d}\varepsilon}{\mathrm{d}\ln a} = (1+3w)\varepsilon , \tag{2.85}$$

with solution $\varepsilon(a) = Ca^{1+3w}$, where C is the integration constant. We can see that there is an instability for $1+3w > 0$ as the perturbation increases with a, i.e. as the Universe expands. If we keep $1+3w < 0$, the perturbation vanishes and $\Omega = 1$ is a stable solution. Therefore if the Universe was dominated by matter or radiation – with equations of state obeying $1+3w > 0$ – throughout its history, the value of $\Omega_0 \approx 1$ today means that $\Omega(a)$ in the early universe must have been extremely close to unity. This ties into another issue known as fine-tuning: the fact that Ω is almost perfectly unity at early times is suspiciously coincidental. This is expanded upon in §2.3.2.

The flatness problem arises from the seemingly nonsensical equality between the densities of today and the early universe, especially given that a has increased by a factor of $\sim 10^{30}$ since the Planck epoch [43]. We showed that $\Omega = 1$ is a repulsive fixed point in a matter/radiation-dominated universe, with any deviations causing a runaway expansion. As such, it seems odd that Ω is still ~ 1 today. Fortunately, an inflationary universe predicts an Ω_0 value very close to unity. If we suppose there was some period of accelerated expansion ($\ddot{a} > 0$) in the very early universe, driven by a component that is neither matter nor radiation, we obtain

$$\Omega(t_\mathrm{P}) \simeq 1 + (\Omega_0 - 1)10^{60} , \tag{2.86}$$

where $t_\mathrm{P} \simeq 10^{-43}$ s is the Planck time.

If the early universe was indeed inhomogeneous, how do completely disconnected regions display the same physical characteristics, such as uniform CMB temperature? The regions have such

vast distances separating them that standard cosmology, with finite light speed, dictates they could never have been in causal contact. This paradox is known as the *horizon problem*. We can examine the path of light using the conformal system from (2.31) in an isotropic spacetime. In this case, we can define the coordinate system such that light only travels in the radial direction and $\theta = \phi =$ constant. The line element is then

$$\mathrm{d}s^2 = c^2 a^2(\tau)\left[-\mathrm{d}\tau^2 + \mathrm{d}\chi^2\right], \tag{2.87}$$

and the path of the null geodesic is

$$\Delta\chi(\tau) = \pm\Delta\tau\,, \tag{2.88}$$

where a positive $\Delta\tau$ indicates outgoing photons. The advantage of this coordinate system is that light rays always correspond to 45° lines – as they do in special relativity – as opposed to curved lines, which is the case when we use cosmic time t [113]. Before continuing, we need to clarify the difference between an event horizon and a particle horizon. Let us consider an observer O at $\chi = 0$ and a second comoving observer E at $\chi = \chi_i$. If E emits a photon at time t_i, the only signals O receives by a time t must be from radial coordinates $\chi < \chi_i$. This distance is known as the *particle horizon*, which we can express using (2.88) as

$$\chi_{\mathrm{ph}}(\tau) = \tau - \tau_i = c\int_{t_i}^{t} \frac{\mathrm{d}t}{a(t)}\,. \tag{2.89}$$

We can visualise the size of the particle horizon as the intersection of an observer's past light cone with the spacelike surface $\tau = \tau_1$. All causal effects must originate in this region.

The *event horizon* is well known, particularly in the context of black holes. All event horizons represent the maximum distance at which an observer at time t_f will no longer receive signals emitted any time after time t,

$$\chi_{\mathrm{eh}}(\tau) = \tau_f - \tau = c\int_{t}^{t_f} \frac{\mathrm{d}t}{a(t)}\,. \tag{2.90}$$

This is the comoving event horizon [114]. Finally we have the Hubble distance, which we briefly defined earlier in (2.27). We can see that this is not synonymous with the particle horizon, despite the two often being used interchangeably[9]. The horizon problem stems from the isotropy present in the CMB to such a high degree of precision, despite it containing vastly separated points that are completely outside of each other's horizons. This seems impossible since the proper distance between two comoving objects increases with scale factor while their particle horizons increase at a higher rate [44]. We must now consider the amount of inflation necessary to solve the flatness and horizon problems. To address this, we can divide the history of our universe into a number of phase transitions with varying growth factors and temperatures. If we assume the early universe followed the radiation-dominated Friedmann model since inflation began at t_*,

$$R(t) \propto t^{1/2} \propto \frac{1}{T(t)} \quad \Longrightarrow \quad \frac{R_*}{R_0} \sim \left(\frac{t_*}{t_0}\right)^{1/2} \sim \frac{T_0}{T_*}\,, \tag{2.91}$$

where we have related $T(t)$ to particle energy $T \sim E/k_{\mathrm{B}}$ and $T_0 \sim 3$ K is the current temperature of the CMB. Equation 2.69 allows us to equate the ratio of the inflation-era spatial curvature to the present one:

$$\frac{\Omega_{k,*}}{\Omega_{k,0}} = \left(\frac{H_0}{H_*}\right)^2\left(\frac{R_0}{R_*}\right)^2 \sim \frac{t_*}{t_0}\,, \tag{2.92}$$

[9] This is because on length scales $\ll d_{\mathrm{H}}$, Newtonian theory is still viable and $d_{\mathrm{p}}(t) \sim cH^{-1}(t) \sim d_{\mathrm{H}}$

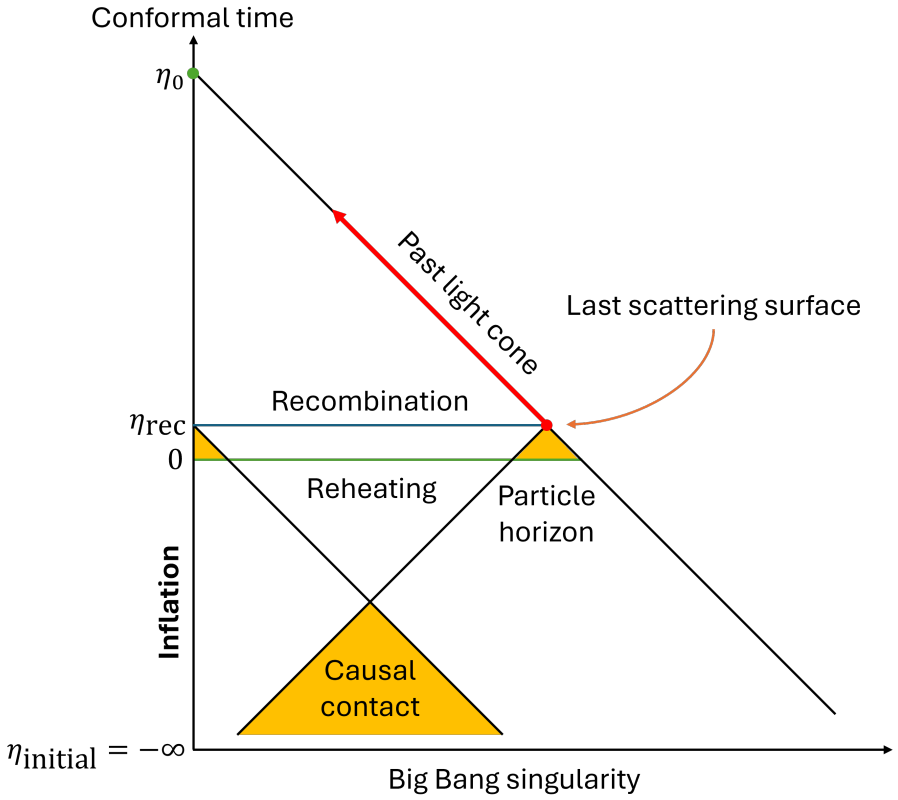


Figure 2.2: In an inflationary FLRW model, the light cones on a causal diagram lie at angles of $\pm 45°$. The worldlines representing comoving matter are vertical while particle horizons are horizontal. Illustrated here is the particle horizon for the CMB. In a non-inflationary universe, conformal time would only extend back to a limit denoted by τ_0, meaning that different regions of the CMB we observe today would not have shared causal contact in the past. However, inflation extends conformal time back to the Big Bang singularity, allowing these regions to have been causally connected at some point along our past light cone. Thus, inflation explains why the observed CMB regions share similar properties, as they would have been in causal contact early on. Image adapted from Debono and Smoot (2016) [115].

where we used the fact that $H_0/H_* \sim t_*/t_0$. If we assume inflation to have occurred some time between the Planck era and GUT phase (see §2.3.1), this ratio is approximately 10^{-60}–10^{-54}. Since the value of the curvature parameter has to be unity today, the scale factor would need to grow by a factor of 10^{27}–10^{30} during this time. We therefore require at least 60–70 e-foldings of the scale-factor to solve the flatness problem.

To examine the horizon problem, we look at the radiation-dominated Friedmann model in its earliest stages. The particle horizon at the inflationary epoch would be $d_{\mathrm{p},*} = 2ct_* \sim 10^{-33}$–$10^{-27}$ m. Equation (2.91) tells us that the size of this causally connected region today is $\sim 10^{-3}$–1 m. However, our universe is currently approximately the size of the Hubble distance $d_{\mathrm{H},0} = cH_0^{-1} \sim 10^{26}$. This means that far-flung galaxies cannot possibly be causally connected, unless the scale factor grew by a factor of 10^{26}–10^{29}. The number of e-foldings necessary to solve the horizon problem is also at least 60–70. Before inflation, the observable universe was limited to the size of the causal horizon at that time, approximately $ct_{\mathrm{P}} \sim 10^{-35}$ m; the size after inflation was therefore $ct_{\mathrm{P}}e^{60} \sim 10^{-9}$ m [44]. This marked the transition from quantum scales to more physical ones.

We can now conclude that a sustained acceleration over a minimum of 60 e-folds is sufficient to solve the horizon and flatness problems [44]. We have also shown that inflation naturally explains the observed flatness of the Universe today. The key question now is what mechanism drove this inflationary epoch and how it eventually ended. Most models propose that, in the early universe, inflation was driven by the vacuum energy density of a scalar field, $V(\phi)$. This allows for an almost exponential expansion when $V \approx$ constant. The various theoretical models for the origin of inflation are explored in detail in Chapter 7 of Coles [66]. In the following section, we will examine the mechanisms responsible for bringing inflation to an end.

2.2.2 Scalar fields

To obtain an inflationary phase in the very early universe, we require some negative-pressure material. The simplest form of this is a scalar field ϕ describing spin-0 particles, the potential of which yields the necessary vacuum energy. Since its potential energy $V(\phi)$ may redshift extremely slowly as the Universe expands, this field possesses an effective equation of state with negative pressure [21]. This hypothetical scalar field believed to be responsible for the initial inflation is called the *inflaton* [16, 18]. The Lagrangian for a scalar field has the usual form $\mathcal{L} = T - V = \frac{1}{2}mg_{ab}\dot{q}^a\dot{q}^b - V$, but with the generalised coordinates replaced by the field $\phi(x^\mu)$ and the time derivatives replaced by spacetime derivatives with respect to spacetime position:

$$\mathcal{L} = \frac{1}{2}g^{\mu\nu}(\nabla_\mu\phi)(\nabla_\nu\phi) - V(\phi)\,, \tag{2.93}$$

since $\mathcal{L}$ must be a scalar function of spacetime position. The corresponding action is

$$S = \int_{\mathcal{R}} \left[\frac{1}{2}g^{\mu\nu}\left(\nabla_\mu\phi\right)\left(\nabla_\nu\phi\right) - V(\phi)\right]\sqrt{-g}\mathrm{d}^4x\,. \tag{2.94}$$

To vary the action with respect to ϕ, we can use this form of the Euler-Lagrange equation:

$$\frac{\delta\mathcal{L}}{\delta\Phi^a} = \frac{\partial\mathcal{L}}{\partial\Phi}^a - \nabla_\mu\left[\frac{\partial\mathcal{L}}{\partial(\nabla_\mu\Phi^a)}\right] = 0\,, \tag{2.95}$$

to obtain

$$\frac{\partial L}{\partial\phi} = -\frac{\mathrm{d}V}{\mathrm{d}\phi} \quad \text{and} \quad \frac{\partial L}{\partial\left(\nabla_\mu\phi\right)} = \frac{\partial}{\partial\left(\nabla_\mu\phi\right)}\left[\frac{1}{2}g^{\rho\sigma}\left(\nabla_\rho\phi\right)\left(\nabla_\sigma\phi\right)\right]\,, \tag{2.96}$$

where the dummy indices in the second equation have been relabelled to make the differentiation clear. Evaluating the second equation gives

$$\frac{\partial L}{\partial\left(\nabla_\mu\phi\right)} = \frac{1}{2}g^{\rho\sigma}\left[\delta^\mu_\rho\left(\nabla_\sigma\phi\right) + \left(\nabla_\rho\phi\right)\delta^\mu_\sigma\right] = \frac{1}{2}\left(g^{\mu\sigma}\nabla_\sigma\phi + g^{\rho\mu}\nabla_\rho\phi\right) = g^{\mu\nu}\nabla_\nu\phi\,, \tag{2.97}$$

allowing us to write equation (2.95) as

$$-\frac{\mathrm{d}V}{\mathrm{d}\phi} - \nabla_\mu\left(g^{\mu\nu}\nabla_\nu\phi\right) = 0\,. \tag{2.98}$$

Recall that the covariant derivative of the metric tensor is zero, so we can simplify this to

$$\Box^2\phi + \frac{\mathrm{d}V}{\mathrm{d}\phi} = 0\,, \tag{2.99}$$

where $\Box^2 \equiv \nabla^\mu\nabla_\mu = g^{\mu\nu}\nabla_\mu\nabla_\nu$ is the covariant d'Alembertian operator. If we choose the potential to be $V = \frac{1}{2}m^2\phi^2$, where m is some constant that influences the dynamics by characterising how steep the potential is as a function of the scalar field – sometimes called the effective mass. The 'mass' also impacts the duration of the inflationary phase; if m is small relative to the value of ϕ during inflation, it leads to a flatter potential and longer inflation. Substituting this scalar field yields the *Klein-Gordon* equation:

$$\nabla^2\phi + m^2\phi = 0. \tag{2.100}$$

This is the case for a free relativistic scalar field, but for now it is preferable to keep the potential function $V(\phi)$ general. We can show that the scalar field acts as a perfect fluid with energy density and pressure[10]

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi) + \frac{1}{2}(\vec{\nabla}\phi)^2, \tag{2.101a}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi) - \frac{1}{6}(\vec{\nabla}\phi)^2\,. \tag{2.101b}$$

If the scalar field only interacts gravitationally with other matter or radiation, it obeys the equation of motion

$$\dot{\rho} + 3(\rho + p)\frac{\dot{a}}{a} = 0\,, \tag{2.102}$$

into which we can substitute the pressure and density from (2.101) to obtain the equation of motion for a scalar field:

$$\boxed{\ddot{\phi} + 3H\dot{\phi} + \frac{\mathrm{d}V}{\mathrm{d}\phi} = 0\,.} \tag{2.103}$$

This equation describes the interplay between the acceleration of the field, the damping effect of cosmic expansion, and the driving force from the potential. If we assume that the scalar field dominates the energy density of the Universe for a period of time such that we can neglect the curvature term in the cosmological field equation (2.81), we can write

$$H^2 = \frac{1}{3}\left[\frac{1}{2}\dot{\phi}^2 + V(\phi)\right]. \tag{2.104}$$

Equations (2.103) and (2.104) provide a set of coupled dynamical equations in ϕ and H that govern the evolution of the scalar field and the scale factor during inflaton domination. Since inflation occurs when $\ddot{a} > 0$, we must also have

$$\dot{\phi}^2 < V(\phi)\,. \tag{2.105}$$

[10] See chapter 7 of Liddle and Lyth (2000) [21] for a comprehensive derivation.

To obtain an analytical solution to the inflation equations (2.103) and (2.104), we make use of the *slow-roll approximation*, where we assume that $\dot{\phi}^2 \ll V(\phi)$. Differentiating this implies that $\ddot{\phi} \ll \mathrm{d}V/\mathrm{d}\phi$, allowing us to neglect the $\ddot{\phi}$ term in equation (2.103). This yields

$$3H\dot{\phi} = -\frac{\mathrm{d}V}{\mathrm{d}\phi} \quad \text{and} \quad H^2 = \frac{1}{3}V(\phi)\,. \tag{2.106}$$

If we differentiate the second equation with respect to t and combine it with the first, we get

$$\dot{H} = -\frac{1}{2}\dot{\phi}^2\,. \tag{2.107}$$

Demanding that $\dot{\phi}^2 \ll V(\phi)$, using the above equations we can obtain two conditions,

$$\epsilon \equiv \frac{1}{2}\left(\frac{1}{V}\frac{\mathrm{d}V}{\mathrm{d}\phi}\right)^2 \ll 1 \tag{2.108a}$$

$$\eta \equiv \frac{1}{V}\frac{\mathrm{d}^2V}{\mathrm{d}\phi^2} \ll 1\,, \tag{2.108b}$$

where ϵ and η are known as the *slow-roll parameters*. We can also consider the special case in which the potential $V(\phi)$ is so flat as to be approximately constant. From equation (2.106), we can see that in this case the Hubble parameter remains constant and the scale factor grows exponentially:

$$a(t) \propto \exp\left(\sqrt{\frac{1}{3}V(\phi)}t\right). \tag{2.109}$$

At some early time we assume the potential is large and almost entirely flat, such that the scalar field ϕ 'rolls' very slowly down the potential $V(\phi)$ due to the potential gradient in (2.103), causing the Hubble constant to decrease slowly and and the Universe to experience generally exponential inflation [116]. The Hubble friction term $3H\dot{\phi}$ prevents the inflaton from picking up too much speed, but as the Hubble constant decreases, the friction term decreases and $\dot{\phi}$ increases until the condition (2.105) no longer holds. Equivalently, the conditions (2.108) in the slow-roll approximation are no longer satisfied.

If the potential possesses a local minimum – as it does in most inflationary models – the field, upon reaching it, will oscillate about the minimum as it is gradually damped by the friction term. As the oscillations decay, the energy of the scalar field is converted into energetic elementary particles. For the inflaton (or any scalar field) to decay into elementary particles, there must be an interaction that facilitates this process, described by coupling terms in the Lagrangian. The decay occurs via interactions where the inflaton field transfers its energy to other fields, leaving the Universe with normal matter and radiation in a state of thermal equilibrium. This provides the initial conditions for the standard cosmological model. If the inflaton were completely decoupled from the Standard Model, it would remain as an oscillating condensate, unable to transfer its energy to radiation or matter. For this reason, the process of scalar field decay into other particles is known as *reheating* [44]. With the Universe now filled with a hot plasma of particles, the standard evolution described by the ΛCDM model can take over, governing the expansion and structure formation that led to the Universe we observe today.

2.3 Our universe

The current standard model ΛCDM uses the FLRW model with the inclusion of dark energy (Λ) and cold dark matter (CDM) [110]. It successfully explains a wide range of observations, including the cosmic microwave background radiation, large-scale structure formation in the Universe,

and the accelerated expansion observed through supernovae data. Regarding curvature, CMB and large-scale structure observations [107, 11] suggest that $K = 0$ is the most likely solution: the Universe is geometrically flat on large scales and fluctuates in size over time due to $a(t)$. They also point to a universe filled with all the aforementioned components: matter, radiation and dark energy. The matter component consists of 5% baryonic matter and 27% cold dark matter. The dark energy equation of state suggests a cosmological constant of $w \approx -1$. Looking at the proportions in (2.64), we can see the curvature contribution is no more than 1%; at earlier times this would have been even less, as the curvature part only increases as a^{-2}.

The different scalings we looked at §2.1.3 imply that the Universe was dominated by single components throughout its history: first radiation (a^{-4}), then matter (a^{-3}), then vacuum energy (a^0). The evolution of these components throughout the history of the early universe can be seen in figure 2.3. A flat, single-component universe would have the Friedmann equation reduce to

$$\frac{\dot{a}}{a} = H_0 \sqrt{\Omega_i} a^{-\frac{3}{2}(1+w_i)} \,, \tag{2.110}$$

where, upon integrating we get

$$a(t) \propto \begin{cases} t^{\frac{2}{3}} & \text{matter-dominated} \\ t^{\frac{1}{2}} & \text{radiation-dominated} \\ e^{H_0 t} & \text{dark energy-dominated} \end{cases} . \tag{2.111}$$

Matter and radiation had equal density at $a_{\text{eq}} \equiv \Omega_{\text{r}}/\Omega_{\text{m}} \approx 3 \times 10^{-4}$, shortly before the CMB dissipated through space[11].

[11] Note that this is not the case for a spatially open universe.

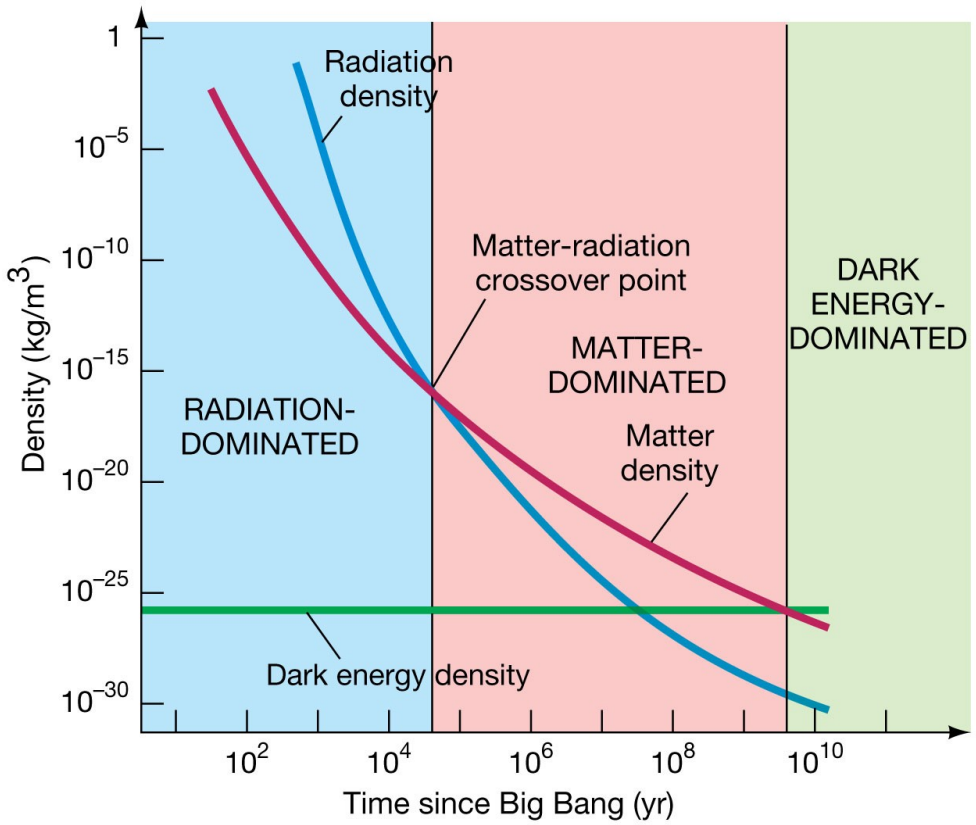


Figure 2.3: Timeline of radiation (blue), matter (red), and dark energy (green) dominance in the Universe, showing the density in km m^{-3} against the age of the Universe in years. Radiation dominated until about 47,000 years after the Big Bang. After matter-radiation equivalence $z_{\text{eq}} \sim 3600$, matter dominated for another ~ 10 Gyr. As the Universe continued to expand, the density of both matter and radiation decreased until dark matter became dominant. Image from Blau (2017) [48] and Pearson Education Inc. (2011).

2.3.1 Structure formation

The formation of galaxies and more complex structures is likely to be the result of coalescing overdensities from the early universe that arose from gravitational instabilities. Primordial fluctuations in matter and radiation evolved into galaxies and clusters as a result of their self-gravity. In the early universe, the growth of these fluctuations (δ) was very small; the equations governing these perturbations can therefore be expanded to first order in δ, allowing us to neglect higher-order, non-linear terms. This is the basis for linear perturbation theory, which will be covered in the next chapter. In this section we will go over the processes in the very early universe, before the emergence of large-scale structure.

Our universe began as an extremely hot and highly compressed region of approximately 10^{30} K [117]. At these scales, a theory of quantum gravity is required to suitably describe the processes occurring inside this volume. As such, we cannot conclude whether the Universe at $t = 0$ was a true singularity, such as those of black holes. Hawking [15], inspired by the Penrose singularity theorem in black holes, discovered another theorem proving that all timelike geodesics can be traced back to a singularity at $t = 0$. Both of these are referred to as the Hawking-Penrose singularity theorems. The primordial singularity began to expand, for reasons we do not yet know. This event is widely referred to as the Big Bang, though the implication that this was some kind of explosion is misleading. It is more fitting to think of it as the inception of space itself [45].

The first stage in the Universe's history was the Planck era. This period corresponds to a temperature $T_{\mathrm{P}} \simeq 10^{19}$ GeV[12], and is the earliest period of time, beginning with the Big Bang [66]. At the end of the Planck era, gravity separated from the other fundamental forces. This period was followed by the grand unification epoch at $10^{-43} < t < t_{\mathrm{GUT}} \simeq 10^{-36}$ s, during which electromagnetism, the strong interaction, and the weak interaction were unified as the electronuclear force – the Grand Unified Theory (GUT) [93]. This era was followed by the electroweak epoch $10^{-36} < t < t_{\mathrm{EW}} \simeq 10^{-11}$ s, at which point the Universe dropped below 10^{15} GeV and the strong nuclear force became distinct from the electroweak force, likely instigating an almost instant period of rapid inflation. By $\sim 10^{-32}$ seconds, the scale factor had increased at least 10^{26}-fold. Immediately afterwards, the Universe cooled down and settled into an era of slow expansion. Whatever was responsible for the rapid expansion previously appears to have diminished as a result of reheating [45], which we covered in §2.2.2.

By the end of this era, the electroweak force had separated into the weak nuclear and electromagnetic forces. At $t_{\mathrm{QH}} \simeq 10^{-5}$ s we have the quark-hadron phase transition, during which quark confinement took place. By this time the Universe contained the four fundamental interactions we know today: gravity, electromagnetism, and the weak and strong nuclear forces. As the Universe cooled and expanded, the energy per particle dropped to a level where the strong force could no longer be overcome. This cooling forced quarks into stable, confined states, leading to the formation of protons, neutrons, and other hadrons, and marked the transition from a quark-gluon plasma to the hadron-dominated universe we have today.

All of this took place within 1 second of the Big Bang occurring; at approximately this time, neutrinos decoupled from matter to form the cosmic neutrino background (CνB) and end the lepton epoch (see bottom of figure 2.4). This background exists at a temperature of ~ 1.9 K, and while there is indirect evidence for its existence [118], it remains undetectable due to neutrinos' extremely small cross section, i.e. their low rate of reaction. The high temperature and plasma density present at the early stages of the Universe imply that atomic nuclei

[12] At extremely high temperatures we often treat temperature and energy as directly interchangeable: 1 GeV roughly corresponds to 1.16×10^{13} K, according to $E = k_B T$.

could have formed during that time. Within the first few minutes after the Big Bang, some of the lightest atomic nuclei were created, a process referred to as Big Bang nucleosynthesis (BBN).

Once Big Bang nucleosynthesis was complete at $t \sim 20$ minutes, having introduced the 'light elements' ^{4}He, D, ^{3}He and ^{7}Li, the Universe became a hot ($\sim 10^{10}$ K) whirlpool of protons, dark matter, light elements, photons and neutrinos. The last two of these – the relativistic components – dominated the energy density. By this time neutrinos and dark matter had already decoupled from the plasma, but the rest of the particles shared the same temperature as a result of the interactions between photons and charged particles. The universe transitioned from a radiation-dominated to a matter-dominated regime, allowing the eventual expansion to be driven by pressureless matter, i.e. dust.

Before inflation began, the scale of gravitational perturbations λ grew with the scale factor a while $\dot{a}/a$ decreased with the slowdown of the expansion. When inflation started, λ increased exponentially while $\dot{a}/a$ plateaued. Ellis and Rothman [113] pedagogically let λ be the distance to a visible quark. This distance soon exceeded the Hubble radius – or the speed-of-light sphere – c/H, so $\lambda H/c \gg 1$, causing the quark to leave the horizon. After inflation, H decreased until eventually $\lambda H/c \ll 1$ and the quark reentered the horizon, now as a constituent of a galaxy. Gravitational perturbations, depending on their λ, therefore reentered the horizon once the horizon had expanded to the size of the fluctuations [119].

The constituent fluids of the Universe depend on the model we are using. For our purposes, a fluid is simply a set of particles dense enough that we can treat them as continuous. Particles under consideration include CDM, baryons, photons, and massless and massive neutrinos, the latter of which is a contender for hot dark matter (HDM). CDM and baryons can be treated as collisionless and collisional[13] fluids, respectively, while photons and neutrinos require a phase-space description governed by the Boltzmann transport equation. The CDM model treats cold dark matter as a pressureless perfect fluid, a.k.a. cold dust, which interacts with other particles via gravity only. For massless neutrinos, the energy density and pressure are given by $\rho_\nu = 3p_\nu = T^i_{\ i}$.

Photons behave differently before and after recombination. Before recombination – and after neutrino decoupling – photons were tightly coupled with baryons, interacting primarily via Thomson scattering

$$\gamma + e^- \to \gamma + e^- \,, \tag{2.112}$$

in which the photon energy is taken to be much less than the electron rest mass, i.e. $h\nu \ll m_e$. It is treated as elastic since the recoil of the electron in its initial rest frame is neglected. The cross section for these collisions is tiny:

$$\sigma_{\mathrm{T}} = \frac{1}{6\pi\epsilon_0^2}\left(\frac{e^2}{m_e c^2}\right)^2 \sim 10^{-25}\mathrm{cm}^2 \,. \tag{2.113}$$

The mean free path for photons is $\lambda = (n_e\sigma_{\mathrm{T}})^{-1}$, making the rate at which a photon undergoes scattering

$$\Gamma_{\mathrm{T,e}} = \frac{c}{\lambda} = c n_e \sigma_{\mathrm{T}} \,. \tag{2.114}$$

We can see that Thomson scattering is most apparent at high free electron densities n_e. The optical depth is found by integrating the scattering rate over time:

$$\tau_{\mathrm{T}} = \int \Gamma_{\mathrm{T,e}}(t)\, \mathrm{d}t \,. \tag{2.115}$$

[13] Collisional gases support particle collisions that occur at a high enough rate to establish a local thermal equilibrium.

The Thomson opacity was high enough that baryons and photons were tightly coupled, fielding the transfer of energy and momentum between them and ensuring they remain at the same temperature. Once recombination began, the Universe cooled further, causing protons and He^{2+} nuclei to combine with free electrons to form neutral atoms. The rapidly declining density of free electrons caused the photon scattering rate $\Gamma_{T,e}$ to drop by at least a factor of 10^3 below the expansion rate H. This caused an increase in the photons' mean free path λ, allowing them to travel freely. Consequently, the Universe transitioned from being opaque to being transparent to radiation [120]. Crucially, this transition represents the limit for any photon-collecting telescope. Recombination is the earliest time we can probe with electromagnetic radiation. The last photons to scatter off electrons at this time, at $z_{dec} = 1090$, make up the CMB radiation that we receive today. This is known as the *time of last scattering*.

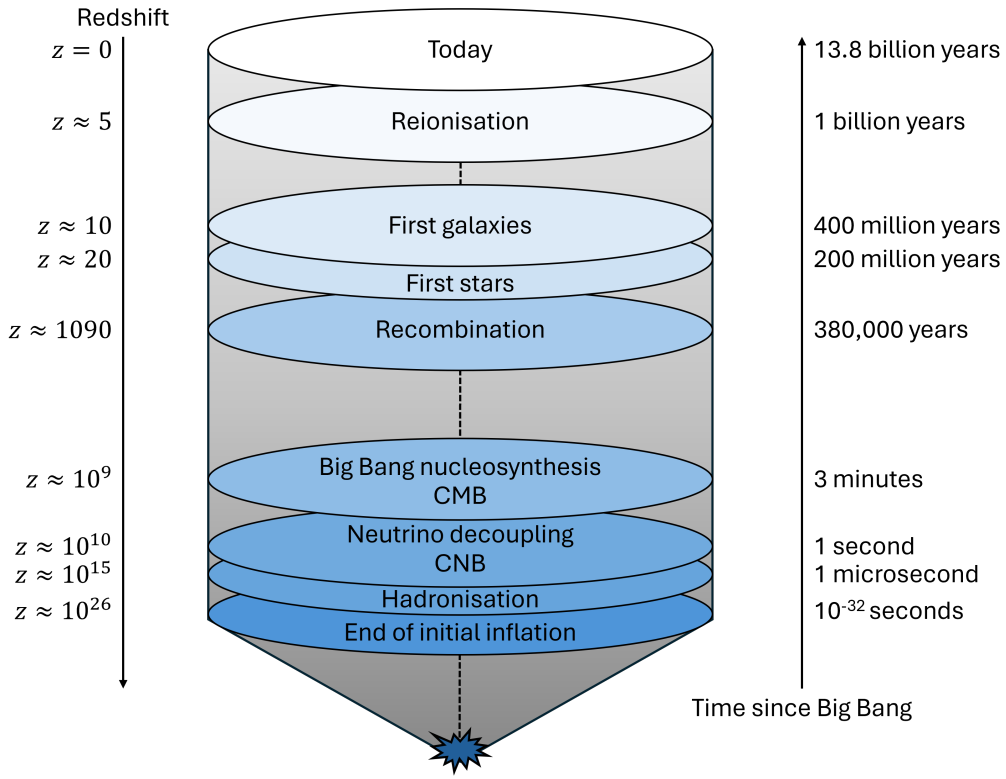


Figure 2.4: The history of the Universe as a light cone looking back from $z = 0$ to the end of the initial inflation following the Big Bang at $z \approx 10^{26}$, with relevant milestones labelled. The arrow on the left shows the redshift, and the arrow on the right shows the time that has passed since the Big Bang for each event. Image adapted from Debono and Smoot (2016) [115].

Observations, when interpreted within the FLRW framework, strongly suggest that the Universe is experiencing another phase of accelerated expansion [14, 94]. This discovery was particularly surprising to astrophysicists, who had reasonably assumed that the expansion of the Universe would continue to slow down as gravitational attraction between objects would dominate, potentially even leading to eventual gravitational collapse. The renewal of accelerated expansion suggests that something is creating negative 'pressure' on the Universe. As with the first, many hypotheses exist to explain the second period of inflation. The most popular of these is the theory of dark energy, a unique type of energy that exerts a positive pressure on the Universe counteracting the inward pull of gravity. Other theories – the subject of this thesis – attempt to modify the theory of general relativity to explain the unexpected acceleration. In the next section we examine dark energy and dark matter more closely.

2.3.2 Dark matter and dark energy

Particles of cold dark matter belong to the classes of cold thermal or non-thermal relics. Thermal relics remain in thermal equilibrium with the other components until they decouple, such as massless neutrinos. These relics can be either hot or cold. The former become relativistic upon decoupling while the latter do not. Non-thermal relics on the other hand do not exist in equilibrium, and comprise the more exotic types such as monopoles, axions and cosmic strings. The upshot of this is that fluctuations produced by CDM are relatively low, and therefore unaffected by photon diffusion [66]. This means structure can survive on small scales. Since the particles are relatively slow-moving, they can clump together to form structure. ΛCDM is referred to as a 'bottom-up' model because it describes a universe where small structures form first and gradually merge to create larger and more complex structures over time, driven primarily by this behaviour of cold dark matter. The other important aspect of dark matter is it that it must interact weakly with the other components; this would explain why it has evaded detection for so long. Virtually every particle in the current Standard Model has been ruled out as a candidate for dark matter. As such, the nature of dark matter remains the white whale for observational cosmologists.

Dark energy dominates the expansion dynamics of the cosmos today, but – like dark matter – its precise nature remains a mystery. The simplest model of dark energy is the cosmological constant, with an equation of state parameter $w \approx -1$. While observations support a flat geometry consistent with ΛCDM, this explanation faces severe theoretical challenges. Attempts to estimate the vacuum energy density from quantum mechanics produce values about 120 orders of magnitude larger than what cosmology indicates, highlighting a huge gap in our understanding [120]. Dark energy reveals itself solely through its influence on cosmic expansion, as it is smoothly distributed through space, unlike dark matter, which clumps under gravity.

To probe dark energy, observational techniques like gravitational wave detection with LISA offer a promising approach. LISA could measure the luminosity distance of merging black holes at high redshift with great precision, providing an independent method to explore cosmic expansion without relying on the traditional cosmic distance ladder. Baryon acoustic oscillations (BAO), which use the acoustic scale as a standard ruler, are another tool to investigate dark energy. By studying how the BAO 'bump' varies with redshift, cosmologists can determine whether dark energy behaves like a cosmological constant or has a time-varying equation of state [93]. Other approaches include the Hubble diagram of Type Ia supernovae and cosmic shear measurements, both of which are being pursued in upcoming space-based missions. These observations are critical for determining whether dark energy is static or evolving. Ground-based projects, such as the Large Synoptic Survey Telescope (LSST), aim to gather data that will further refine these methods, increasing our sensitivity to changes in the equation of state parameter, w [120].

Despite these ongoing efforts, the mystery of dark energy remains unresolved. While plausible ideas about dark matter exist and may soon be testable, dark energy's connection to quantum physics and gravity is far more elusive. Solving this problem may require a new theory that unifies quantum mechanics and gravity, similar to the way electromagnetism and the weak force were unified in the Standard Model. For now, dark energy stands as one of the greatest challenges to fundamental physics, requiring both innovative theoretical models and high-precision observational cosmology.

2.3.3 Problems with ΛCDM

For decades, the cosmological constant has served as a compelling explanation for the Universe's current accelerating expansion, though it is not without significant challenges. Dicke first highlighted the issue of 'fine-tuning' in a cosmological context in 1961 [27]: if Λ were even slightly larger, large-scale structure formation would be impossible, as the Universe would expand too rapidly for matter to clump under gravity. Conversely, a smaller Λ might have caused the Universe to expand more slowly, potentially leading to recollapse before large-scale structures could form. Additionally, if Λ dominated over matter and radiation, the Universe would expand too quickly for gravitational collapse to form galaxies and clusters. This balance between Λ and matter/radiation introduces its own questions. The current dark energy density's similar magnitude to that of matter, despite their different evolution histories, has been termed the 'coincidence problem' [121][14]. Furthermore, quantum physics has revealed a substantial discrepancy between the observed vacuum energy density and the vastly larger zero-point energy predicted by quantum field theory [123]. Lastly, recent DESI results appear to support a theory of evolving dark energy, as parameterised by the w_0w_aCDM model [124]. The ongoing and future observations from DESI offer an exciting opportunity to refine our understanding of dark energy, potentially revealing deviations from a cosmological constant and providing new insights into the fundamental physics governing cosmic acceleration.

Another consideration is the various tensions ($> 2\sigma$) between astrophysical data and the ΛCDM model specified by the Planck parameters [107], which may hint at the need for extensions beyond a simple cosmological constant. The most infamous of these is the Hubble tension. Local measurements of the Hubble constant (H_0), based on the distance ladder approach, yield values significantly higher than those inferred from the cosmic microwave background (CMB) under the ΛCDM model. Specifically, local direct measurements by the SH0ES Team, using Cepheid calibrators, find $H_0 \approx 73.04 \pm 1.04$ km s^{-1} Mpc^{-1} [125], which is in a 5σ tension with the CMB-inferred Planck/ΛCDM value of $H_0 \approx 67.4 \pm 0.5$ km s^{-1} Mpc^{-1} [24]. There are a plethora of measurements for H_0, though they generally centre around one of these quoted values, in a bimodal distribution; the most recent are given in figure 2.5. Included in the figure is an additional set of values that lies around $H_0 \simeq 69.8$ km s^{-1} Mpc^{-1}, which was obtained by Freedman [126] using the tip of the red giant branch (TRGB) method, and will certainly benefit from additional data from Gaia and JWST [127]. Theoretical solutions to the tension can generally be divided into three categories: pre-recombination adjustments, modifying early-universe physics to recalibrate the standard ruler used in CMB analysis; late-time modifications, which involve altering the Hubble rate $H(z)$ at late times; and luminosity recalibration, by proposing new physics that affects the absolute luminosity of Type Ia supernovae at late times [128].

Other tensions include the growth tension, which refers to the 2–3σ discrepancy between direct measurements of the growth rate of cosmological perturbations (using weak lensing, redshift space distortions, and cluster counts) and the higher growth rate predicted by Planck/ΛCDM parameters [129]; and CMB anistropy anomalies, which encapsulate a series of 2–3σ deviations from what is expected from ΛCDM in the CMB [130]. Lesser known dilemmas include cosmic dipoles, the BAO Lyα tension, the lithium problem, age-of-the-universe issues, etc., all of which are covered in comprehensive reviews by Perivolaropoulos and Skara [128] and Bull et al. [131].

Finally, the ΛCDM model faces several 'small-scale' problems arising from the comparison of increasingly detailed simulations and improved observational constraints [131].

[14] A suggested solution for these is the anthropic principle, which posits that the 'coincidences' are actually the result of human selection bias in a multiverse [122].

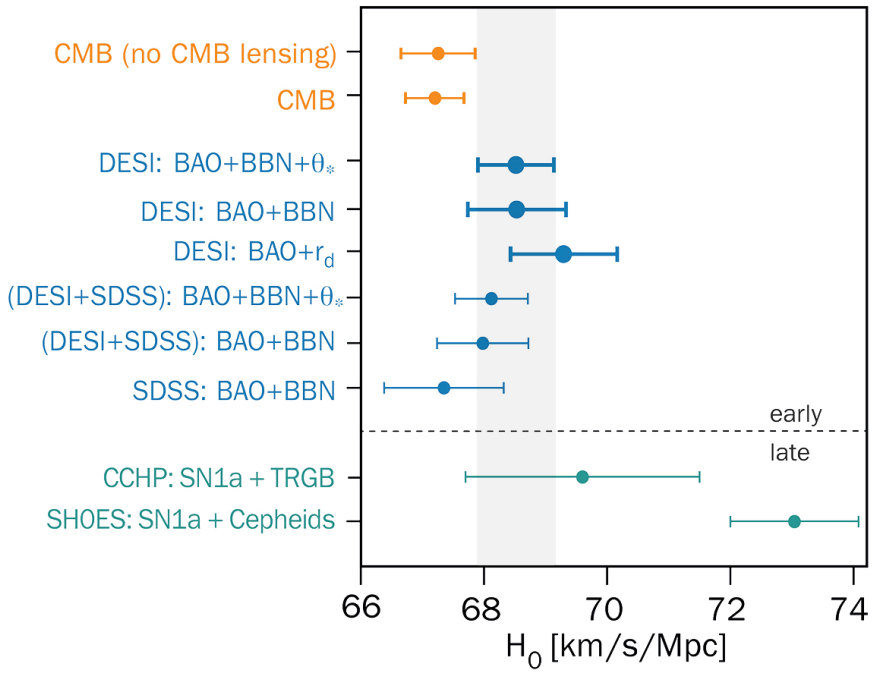


Figure 2.5: Plot of H_0 measurements showing 68% credible interval constraints, assuming a flat ΛCDM model. It shows the CMB anisotropy measurements from Planck and the Atacama Cosmology Telescope (orange); the DESI BAO measurements in combination with other data (blue, bold); corresponding results from SDSS combined with DESI (blue); and measurements using cepheid and TRGB methods (green). Image from the DESI Collaboration (2024) [99].

Missing Satellites: Predicted dark matter halo substructures outnumber observed satellite galaxies [132, 133].

Too Big to Fail (TBTF): Observed densities in dwarf galaxies are lower than the higher densities of massive subhalos predicted by ΛCDM [134] This issue was refuted in [135], which claimed statistical gaps are to be expected.

Emptiness of Voids: Observed galaxy velocities in voids (from ALFALFA[15]) differ from those in comparable ΛCDM -simulated regions [136, 137].

Core/Cusp Problem: ΛCDM predicts dense, central 'cusps' in halos, but dwarf and low-surface-brightness galaxies often exhibit flatter, 'cored' profiles [138, 139]. More recent papers [140] found issue with the methods laid out in Walker et al., and that until more precise data is collected, the issue remains inconclusive [141].

Void Phenomenon: Galaxies in voids display properties similar to those in dense environments, despite differing expected histories [142, 143]. Conflicting research claims this phenomenon is predicted by ΛCDM [144].

Satellite Planes: The thin, rotating alignments of Milky Way and Andromeda satellite galaxies are rare in ΛCDM simulations [145, 146]. However, recent research has found ΛCDM to be consistent with observations [147, 148], suggesting this issue may be resolved.

Baryonic Tully–Fisher Relation (BTFR): The observed tight relationship between galaxy circular velocity and baryonic mass does not match ΛCDM expectations [149].

[15]The Arecibo Legacy Fast ALFA (Arecibo L-band Feed Array) Survey.

Mass-Discrepancy vs. Acceleration Relation (MDA): A strong correlation between mass discrepancy and acceleration in galaxies is unpredicted by ΛCDM [150].

Proposed solutions include improved baryonic physics modelling, alternative dark matter types (e.g., warm, interacting, or self-interacting dark matter), and modified gravity theories, like MOND or MG models, either supplementing or replacing dark matter. Modified gravity theories, particularly $f(R)$ models, will be the subject of the chapter 4.

2.4 Summary

In this chapter, we explored the foundational concepts of cosmology, beginning with the cosmological principle, which posits the Universe's large-scale homogeneity and isotropy – a premise supported most directly by the CMB. Building on these principles, we developed the FLRW geometry to describe a universal framework, using symmetric spaces and curvature equations to model the Universe's shape and scale. The derivation of the FLRW metric, along with the introduction of the scale function and redefined coordinates, enabled us to characterise distances and time scales central to cosmological studies, such as the Hubble and comoving distances and the dimensionless scale factor. The application of the metric – in addition to the Einstein field equations from chapter 1 – facilitated the derivation of the Friedmann, continuity, and energy conservation equations. These equations led us to define essential cosmological parameters and examine universe models dominated by matter, radiation, and vacuum energy. With the Hubble parameter and its higher-order extensions (deceleration, jerk, snap, and lerk), we gained tools to measure cosmic expansion, while the density parameters and Ω values helped us visualise how variations in density and curvature shape cosmic geometry. Through a review of models like the Einstein-de Sitter, de Sitter, Einstein static, and Milne universes, we highlighted different theoretical frameworks for cosmic evolution and introduced the cosmological constant, Λ, and its profound role in expanding the Universe. The concept of inflation emerged as a solution to the flatness and horizon problems, and this process allowed us to distinguish the event horizon and particle horizon from the Hubble distance. Finally, we assessed the evolution of our own universe, from the Big Bang through to large-scale structure formation, and considered the roles of dark matter and dark energy in the ongoing expansion. This chapter has only briefly touched on the complexities of structure formation, primarily assuming an idealised, uniform universe. In the next chapter, we will delve into the role of initial inhomogeneities – seeds that, through the application of perturbation theory, became the galaxies, clusters, and vast structures we observe today.

Chapter 3

Growth of Large-scale Structure

Theories accounting for the current accelerated expansion of the Universe lie at a crossroads: we can either introduce a new, smooth component with sufficiently negative pressure – dubbed dark energy – or modify the laws of gravity such that this acceleration is allowed. Depending on which path we take, the evolution of matter perturbations will be affected differently. Despite the additional work this presents, we can use it as a consistency check for our results; perturbation theory proves to be a powerful tool for differentiating models that do or do not adhere to general relativity (GR) [38]. Henceforth, to streamline calculations, we let $c = 1$ unless specified.

Any reasonable cosmological model should be able to account for the structure we observe in the Universe. We have discussed the homogeneity that manifests at very large scales, but we know that zooming in reveals a number of 'small' details ranging from hyperdense neutron stars only a few kilometres across, to great walls extending for hundreds of megaparsecs. These structures give us information about the state of the early universe, and their great variety is likely due to gravitational instability. Gravity is the only – relatively well-understood – action we know of that acts at universe-scale distances and can aggregate matter into discernible structure. Gravitational instability has the ability to amplify the growth of small density fluctuations in the early universe, making it the prominent candidate for understanding the origin of structure formation. In our case, the observable universe is well described by the FLRW universe [151]. This means that all structure would have evolved from initial small density perturbations – arising from gravitational instability or cosmic explosions – in the FLRW background. Those stemming from initial conditions just after the big bang, i.e. within 10^{-34} seconds are known as *primordial* perturbations.

We refer to this perturbation-inclusive model as an *almost*-FLRW universe. We know that the perfectly homogeneous and isotropic FLRW universe is applicable on large scales, but remains unrealistic on small scales due to the presence of galaxies, clusters, and cosmic structures. An almost-FLRW universe is a more physical model that incorporates small deviations from homogeneity and isotropy, allowing for local inhomogeneities and anisotropies such as density fluctuations, cosmic structures, or small anisotropies in the cosmic microwave background. Perturbations can also be classified into scalar, vector, and tensor types [152]. The first is the most significant; scalar perturbations correspond to fluctuations in the energy density, which create gravitational potential wells that seed the formation of large-scale structure. These same potential wells also imprint anisotropies on the CMB via the *Sachs-Wolfe effect* [153], where photons from high-density regions at last scattering lose energy as they climb out of gravitational wells, leading to redshifts in their observed temperature. This redshift causes a fractional variation of the temperature:

$$\frac{\Delta T}{T} = \frac{\delta\phi}{c^2}\,, \tag{3.1}$$

where $\delta\phi$ is the fluctuation in gravitational potential. The photons also experience a time dilation effect dependent on the scale factor,

$$\frac{\Delta T}{T} = -\frac{\delta a}{a}\,. \tag{3.2}$$

The relative contributions of these two terms depend on the behaviour of the perturbations [66]. As a result, CMB radiation has become one of the most important probes of Big Bang cosmology. Vector perturbations arise from vorticity or rotational motions in the fluid, and typically decay over time in an expanding universe. Tensor perturbations are fluctuations in spacetime, i.e. gravitational waves [154, 155]. These waves travel through the Universe relatively unaffected by matter, preserving information about the early universe, and comprise a fascinating area of cosmological research.

A simple linear expression for the behaviour of matter perturbations is of the form

$$\delta \equiv \frac{\delta\rho}{\bar{\rho}} = \frac{\rho - \bar{\rho}}{\bar{\rho}}\,, \tag{3.3}$$

also known as the *density contrast*, which measures the fractional deviation of the local density ρ from the background density $\bar{\rho}$. Positive values ($\delta > 0$) indicate regions where the density is higher than average (overdensities), which could collapse under gravity to form structures like galaxies. Negative values ($\delta < 0$) indicate underdense regions (voids). Small values ($|\delta| \ll 1$) imply a weak gravitational field, which can be studied using linear theory, or the Newtonian approach. However, this breaks down on scales exceeding the Hubble radius (2.27) because the metric perturbations dominate the dynamics at these scales. Because Newtonian theory neglects perturbations of the metric, it is not viable for studying perturbations on large scales.

The CMB provides us with a road map of the Universe back to the time of last scattering (see §2.3.1). Observation of the CMB anisotropy should ideally enable us to infer the properties of primordial perturbations. Such inferences began over 50 years ago when Sachs and Wolfe [153] calculated the perturbed redshift back to last scattering along null geodesics for a perturbed universe using (3.1) and (3.2). The calculation of CMB anisotropies is hampered by the gauge-freedom that arises when specifying a map Φ between the real universe S and the unperturbed background model $\bar{S}$ [109], i.e., FLRW. The map can be specified by imposing coordinate conditions on S and $\bar{S}$, and the resulting residual freedom means that i) map-dependent quantities such as the density perturbation are not necessarily observable, as the map cannot be reconstructed from observations in the real universe; and ii) points in $\bar{S}$ can be mapped to physically incompatible points in S, allowing unphysical gauge mode solutions to exist.

3.1 The gauge problem

To illustrate the significance of metric fluctuations on large scales, we can consider a *separate universe* argument [156, 157]. This states that on scales much larger than the Hubble radius, each region of the Universe evolves as if it were a separate, unperturbed FRLW universe. In other words, when dealing with these *super-horizon* perturbations[1], each region of space evolves according to its own local Hubble expansion without being influenced by causal effects from other regions. This is because regions separated by distances larger than the Hubble radius cannot exchange information. The separate universe argument implies that large-scale perturbations – such as those associated with density fluctuations in the early universe – can be treated as local

[1]Scales greater than or less than the Hubble radius are termed 'super-horizon' or 'sub-horizon' scales, respectively.

changes in the parameters of a homogeneous universe model, rather than as local deviations from a homogeneous background. This allows us to analyse the perturbations using the dynamics of homogeneous, isotropic universes, rather than dealing with the more complex, inhomogeneous solutions.

There are no particular coordinates we need to use for perturbed spacetime, but we do want them to reduce to the standard coordinate system in the limit where the perturbations disappear. We refer to coordinates that satisfy this requirement as a *gauge*. When defining perturbations, we consider both an unperturbed, background universe $\bar{S}$ and a more accurate, inhomogeneous universe S. The perturbation of quantities such as pressure or density can then be determined by taking the difference between these quantities at any given spacetime point q. The perturbation *field* results from considering all points in spacetime [109]. By treating the background and perturbed spacetime as two different spacetimes, we can define a coordinate system on the perturbed spacetime with the metric

$$\delta g_{ab} = g_{ab} - \bar{g}_{ab}\,, \quad |\delta g_{ab}| \ll 1\,, \tag{3.4}$$

where $\bar{g}_{ab}$ is the metric of the background spacetime and the energy-density perturbation, for example, is $\delta\mu = \mu - \bar{\mu}$, where the bar denotes the quantity in the idealised spacetime $\bar{S}$. Taking the perturbation – as well as its first and second partial derivatives[2] – to be very small allows us to ignore any terms beyond first order.

However, equation 3.4 misrepresents the situation. Subtraction between two vectors on separate manifolds cannot be covariantly defined. We need some kind of map that allows us to determine if points in $\bar{S}$ and S are valid for vector subtraction [154]. We can see that the only constraint we have in relating the two models in (3.4) is δg_{ab}, which must necessarily be very small. This does not seem like enough information to recover a unique spacetime $\bar{S}$ from S. A more coherent method is to define a map Φ from $\bar{S}$ to S that allows us to map the density in the background spacetime to a background density in the perturbed spacetime, as seen in figure 3.1. The choice of coordinates determines the map from $\bar{S}$ into S, and the freedom of this coordinate choice in S is known as the gauge freedom [109]. If we want to change the correspondence between the background and perturbed spacetimes, we need to use a *gauge transformation*. More detailed explanations of the gauge problem can be found in Carroll [43], Ellis and Bruni [109], and Stewart [154].

[2]It is not always necessary to keep the second derivative small, but it does make the process easier.

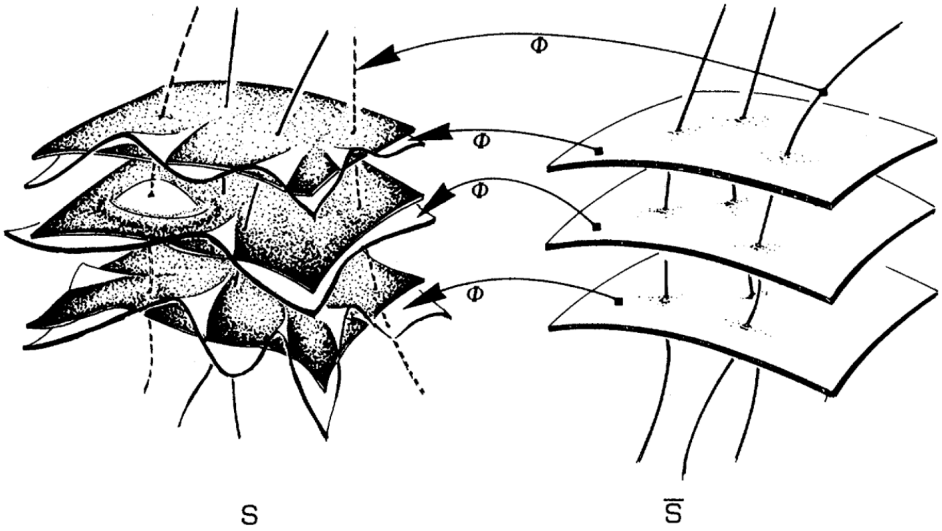


Figure 3.1: The mapping Φ between the unperturbed, background universe $\bar{S}$ and the perturbed universe S. Image from Ellis and Bruni (1989) [109].

Equation 3.4 allows us to write the curvature and energy tensors as

$$G^a_b = \bar{G}^a_b + \delta G^a_b\,, \quad |\delta G^a_b| \ll 1 \tag{3.5}$$
$$T^a_b = \bar{T}^a_b + \delta T^a_b\,, \quad |\delta T^a_b| \ll 1\,. \tag{3.6}$$

Substituting the Einstein equations of the two spacetimes, $G^a_b = 8\pi G T^a_b$ and $\bar{G}^a_b = 8\pi G \bar{T}^a_b$, gives us the field equation for the perturbations

$$\delta G^a_b = 8\pi G \delta T^a_b\,. \tag{3.7}$$

To compare spacetimes, we need any point in the background spacetime to correspond to a point in the perturbed spacetime, and vice versa. Since many coordinate systems exist for which (3.4) holds, our choice of gauge will prove important.

In linear, or first-order, perturbation theory, we can drop all products of the small quantities, δg_{ab}, $\delta g_{ab,c}$ and $\delta g_{ab,cd}$. The field equation then becomes a linear differential equation. This could be simplified even further by using Minkowski space for the background spacetime, for which $\bar{g}_{ab} = \eta_{ab}$ and $\bar{G}^a_b = \bar{T}^a_b = 0$ [158]. $\bar{G}^a_b$ and $\bar{G}^a_b$ do not satisfy the perturbation equation exactly, since GR is nonlinear and perturbations affect the evolution of the mean; this is known as *backreaction*. In first-order linear perturbation theory, however, we can assume that they are solutions. In cosmological perturbation theory, the background spacetime is typically taken to be the FLRW universe, which is homogeneous and isotropic but evolves with time due to cosmic expansion. Fortunately, the Euclidean geometry of a flat FLRW universe greatly simplifies calculations. This background spacetime provides the foundation for studying perturbations.

A quantity is gauge invariant if it is independent of δg_{ab} (3.4). Unfortunately, most perturbations are gauge dependent. The first systematic approach to handling this – the *metric approach* – was introduced by Lifshitz in 1946 [152], who defined small perturbations around an FLRW metric. However, his method led to unphysical gauge modes, limiting its applicability. In 1980, Bardeen [159] developed a more robust gauge-invariant formulation. This method involves a scalar-vector-tensor (SVT) decomposition of the perturbations, which allows one to treat the Einstein equations for scalars, vectors, and tensors separately at linear order. However, these perturbations are not unique and require a choice of gauge. One solution is to define a set of

metric perturbations that do not transform under coordinate transformation, i.e., the *Bardeen variables*[3]. In this method, the variables are kept as the dynamical degrees of freedom [159], allowing gauge modes to be eliminated automatically from the perturbation equations. However, the linearity of the equations means that the variables can only be defined for small variations from FLRW symmetry. The variables themselves are only gauge-invariant for gauge transformations that respect the SVT decomposition. The last issue is that the Bardeen variables, while gauge-invariant, fail to characterise the perturbations in a way that can be easily interpreted physically [60, 160]. One such example is the uniform density gauge, where $\delta\rho$ is set to zero. In the post-Newtonian approach, ρ is already gauge-invariant, so attempting to impose this condition via a gauge transformation doesn't make sense as the density is already independent of the gauge choice [161].

Another method is to fix the gauge and keep track of the perturbations. This can become cumbersome when dealing with small perturbations or complex geometries, and requires careful treatment of gauge freedoms. Moreover, because the metric approach considers small variations of the metric tensor, extracting a physical interpretation from the results can be difficult. The metric tensor is not a physically significant quantity, so it cannot be directly measured; only its second derivatives can. This, in addition to the aforementioned gauge issues, are the primary limitations of metric-based perturbation theory. The metric approaches, both gauge-dependent and -invariant, are expanded upon in [114] and [60]. To preserve general covariance under perturbations, we must pay attention to how gauge transformations affect the perturbations at each order, and attempt to eliminate gauge degrees of freedom. For a tensor to be gauge-invariant at a certain order, it must vanish (or be constant) at all lower orders. This ensures that any physical implications derived from it do not depend on the choice of gauge.

An alternative to the metric approach here is the $1+3$ covariant approach. Its advantages include being locally defined; being covariant; and using variables defined such that they are automatically scalar, vector, or tensor modes [162]. As the name implies, this approach involves the covariant $1+3$ splitting of spacetime, which does not involve a choice of coordinates. This makes it distinct from $3+1$ approaches, where the choice of three-dimensional spatial sections is not unique. The covariant method is an apt choice for analysing gauge-invariant perturbations; collectively this is known as the covariant and gauge-invariant approach (CGI). A key distinction between this and the standard approach is that CGI begins with the fully non-linear equations – before any linearisation takes place – instead of with the background. This is advantageous when considering non-linear equations.

At first-order perturbations, any of the $1+3$ objects with an index – in addition to scalars – must vanish in the background to maintain symmetry; they must therefore be gauge invariant according to the *Stewart-Walker lemma*. This theorem asserts that a tensor is gauge-invariant as long as it vanishes in the background. The effect of a gauge transformation generated by an infinitesimal vector field ξ on a tensor $\bar{T}$ in the background spacetime is equivalent to the Lie derivative (1.30) of T along ξ:

$$\bar{T}' = \bar{T} + \pounds_\xi \bar{T} \implies \delta T' = \delta T + \pounds_\xi \bar{T}\,, \quad \pounds_\xi \bar{T} = 0 \implies \delta T' = \delta T\,. \tag{3.8}$$

We can then apply the Stewart and Walker Lemma: a perturbation to a geometrical background quantity T is gauge-invariant if and only if $\bar{T}$ (i) is a constant scalar, (ii) vanishes, or (iii) is a linear combination of products of Kronecker deltas with constant coefficients [163]. The core idea developed by Ellis and Bruni [109] is to define covariant quantities T such that the corresponding $\bar{T}$ vanishes in an FLRW universe. These variables are inherently gauge-invariant perturbations

[3]This method is laid out comprehensively in Peter and Uzan [42].

in almost-FLRW universes, with their physical meaning clearly tied to their covariant definition. However, this does not mean that all perturbations disappear in a different gauge. For example, density perturbations are gauge-dependent, meaning their specific numerical value and evolution depend on the choice of gauge, but the physical inhomogeneities they represent persist in all gauges.

3.2 The CGI approach

In cosmology there are physically defined observers and hence reference frames from which we examine physical quantities. All observable quantities inherently depend on the choice of such an observer. These quantities are independent of any coordinate system, thus preserving general covariance. This has led to the $1+3$ covariant approach to cosmological perturbation theory, which carefully distinguishes between gauge invariance and frame invariance [162]. The semi-tetrad conversion in the $1+3$ approach converts the GR field equations into evolution and constraint equations that rely on covariant objects (tensors) defined relative to an observer's velocity field u^a, its associated spatial projection tensor $h^{ab} = g^{ab} + u^a u^b$, and the volume element $\varepsilon_{abc} = u^d \eta_{abcd}$. These tensors can be irreducibly decomposed into scalars, vectors, and projected symmetric and trace-free (PSTF) tensors which collectively govern the dynamics of the system. This leads to a description of the background spacetime in terms of a family of invariant scalars, such as energy density and expansion rate, assuming u^a represents the fluid's velocity. Gauge-invariant quantities arise naturally when considering perturbations of objects that vanish in the unperturbed background. The vorticity, shear, and electric and magnetic parts of the Weyl tensor all vanish in a homogeneous FLRW background, ensuring their perturbations remain gauge-invariant. Other gauge-invariant perturbations can be obtained from the spatial gradients of various scalar quantities. A more formal derivation can be found in [154]. We can now define new gauge-invariant quantities for a single-component fluid.

The following is based on work by Hawking [164], Olsen [165], and Ellis and Bruni [109] who make use of gauge-invariant variables that are covariantly defined in the real universe. It will also closely follow the procedures laid out in Ellis and van Elst [166][4]. The four-velocity is the spacetime vector tangent to particle worldlines; these represent the average velocity of matter at every point in spacetime:

$$u^a = \frac{\mathrm{d}x^a}{\mathrm{d}\tau}, \quad u_a u^a = -1\,, \tag{3.9}$$

where the proper time τ is measured along the fundamental worldlines [166]. The four-acceleration is defined as

$$a^a \equiv \dot{u}^a = u^b \nabla_b u^a\,. \tag{3.10}$$

The projection tensors, which project arbitrary spacetime vectors into the space of vectors tangent to a spacelike hypersurface, are uniquely defined as

$$U^a{}_b = u^a u_b \quad \Rightarrow \quad U^a{}_c U^c{}_b = U^a{}_b\,,\ U^a{}_a = 1\,,\ U_{ab}u^b = u_a \tag{3.11a}$$

$$h_{ab} = g_{ab} + u_a u_b \quad \Rightarrow \quad h^a{}_c h^c{}_b = h^a{}_b\,,\ h^a{}_a = 3\,,\ h_{ab}u^b = 0\,. \tag{3.11b}$$

The covariant time derivative along the fundamental worldlines is simply the covariant derivative along u^a:

$$\dot{T}^{ab}{}_{cd} = u^e \nabla_e T^{ab}{}_{cd}\,, \tag{3.12}$$

[4] Based on Ellis' original Cargèse lectures [167].

and the fully orthogonally projected covariant derivative $\widetilde{\nabla}_a$ behaves as

$$\widetilde{\nabla}_e T^{ab}{}_{cd} = h^a{}_f h^b{}_g h^p{}_c h^q{}_d h^r{}_e \nabla_r T^{fg}{}_{pq} \tag{3.13}$$

for any tensor $T^{ab}{}_{cd}$. We can see that the $\widetilde{\nabla}_a$ operator is useful in avoiding excess indices. The irreducible parts of the spatial derivative of projected, symmetric, trace-free tensors are the *divergence*, *curl*, and *distortion* [168]:

$$\text{div } T_{b\ldots c} = \widetilde{\nabla}^a T_{ab\ldots c} \tag{3.14}$$

$$\text{curl } T_{ab\ldots c} = \varepsilon_{de\langle a} \widetilde{\nabla}^d T_{b\ldots c\rangle}{}^e \tag{3.15}$$

$$\text{dis } T_{ca\ldots b} = \widetilde{\nabla}_{\langle c} T_{a\ldots b\rangle}\,, \tag{3.16}$$

where $\varepsilon = u^d \eta_{abcd}$ is the volume element. Splitting the first covariant derivative of u_a into its irreducible parts gives

$$\nabla_a u_b = -u_a \dot{u}_b + \widetilde{\nabla}_a u_b = -u_a \dot{u}_b + \frac{1}{3}\Theta h_{ab} + \sigma_{ab} + \omega_{ab}\,, \tag{3.17}$$

where the trace $\Theta = \nabla_a u^a$ is the scalar rate of volume expansion of the fluid, $\sigma_{ab} = \nabla_{\langle a} u_{b\rangle}$ is the trace-free symmetric rate of shear tensor describing the distortion rate of matter flow, and $\omega_{ab} = \nabla_{[a} u_{b]}$ is the anti-symmetric vorticity tensor describing the rotation of the matter relative to a non-rotating frame. The volume expansion can also be expressed as

$$\Theta \equiv \nabla_a u^a = 3\frac{\dot{a}}{a} = 3H(t)\,, \tag{3.18}$$

where a is the scale factor. This definition will be useful later. The derivative of the volume expansion is then simply $\dot{\Theta} = u^b \nabla_b (\nabla_a u^a)$. Recalling the identity

$$\left(\nabla_c \nabla_d + \nabla_d \nabla_c\right) u^a = R^a{}_{bcd} u^b\,, \tag{3.19}$$

and the magnitudes of the shear and vorticity tensors,

$$\sigma^2 = \frac{1}{2}(\sigma_{ab}\sigma^{ab})\,, \quad \omega^2 = \frac{1}{2}(\omega_{ab}\omega^{ab})\,, \tag{3.20}$$

one can derive the Raychaudhuri equation, beginning by contracting the indices and multiplying by u^c:

$$\begin{aligned} R_{dc} u^d u^c &= u^c \nabla_a \nabla_c - u^c \nabla_c \Theta \\ &= \nabla_a (\nabla_c u^a u^c) - \nabla_c u^a \nabla_a u^c - u^c \nabla_c \Theta \\ &= \nabla_a \dot{u}^a - \left(u_a \dot{u}_b u^a \dot{u}^b + \frac{1}{3}\Theta^2 + 2\sigma^2 - 2\omega^2 \right) - \dot{\Theta}\,. \end{aligned}$$

The left-hand side can be found by substituting the Einstein Field Equations for a perfect fluid (1.83):

$$\begin{aligned} R_{ab} u^a u^b &= K T_{ab} u^a u^b - \Lambda g_{ab} u^a u^b + \frac{1}{2} R g_{ab} u^a u^b \\ &= K\rho u_a u_b u^a u^b + K p h_{ab} u^a u^b - \Lambda \\ &= \frac{1}{2} K(\rho + 3p) - \Lambda\,. \end{aligned}$$

The **Raychaudhuri equation** can then be written as

$$\dot{\Theta} - \nabla_a \dot{u}^a = -\frac{1}{3}\Theta^2 + \dot{u}_b \dot{u}^b - 2(\sigma^2 - \omega^2) - \frac{1}{2}(\rho + 3p) + \Lambda \,. \tag{3.21}$$

This is the *linearised* Raychaudhuri equation. This equation shows the evolution of Θ along the fluid flow lines and is the fundamental equation describing gravitational attraction. The $\rho + 3p$ term represents the effective source of gravitational effects in the fluid. If we impose homogeneity and isotropy on (3.21), the shear and vorticity vanish. We can also treat the product of the acceleration with itself as negligible, as we do not expect any great deviation from geodesic motion, and absorb it into the covariant derivative. This leaves us with

$$\dot{\Theta} - \nabla_a \dot{u}^a = -\frac{1}{3}\Theta^2 - \frac{1}{2}(\rho + 3p) + \Lambda \,. \tag{3.22}$$

We can express this in terms of the scale factor using (3.18) to obtain what is sometimes known as the *acceleration equation*:

$$3\frac{\ddot{a}}{a} - \nabla_a \dot{u}^a = -\frac{1}{2}(\rho + 3p) + \Lambda \,. \tag{3.23}$$

This form shows that the rate of expansion depends on the energy density and pressure of the matter content in the Universe. If $\rho + 3p > 0$, the Universe decelerates; if $\rho + 3p < 0$, the Universe is dominated by the cosmological constant and accelerates. We can recognise these as the weak- and strong energy conditions from §1.2.2. A singularity occurs when the expansion scalar Θ goes to negative infinity in a finite time, leading to a state where the geodesics converge to a point, and quantities like density and curvature become infinite. The Raychaudhuri equation suggests this will happen under the following conditions:

- The matter in the Universe satisfies the strong energy condition (1.59), implying an attractive gravity which causes convergence of geodesics.
- There is initially negative expansion ($\Theta < 0$), which will cause Θ to continue to decrease over time.

It therefore reveals that singularities are a generic feature in general relativity, arising when geodesics converge under the influence of attractive gravity and the strong energy condition. These singularities signal points where spacetime curvature becomes infinite, leading to a breakdown in the laws of physics. If we start from an expanding state, the Raychaudhuri equation predicts that as we go backward in time, $\Theta \to -\infty$. This corresponds to a state of infinite density and curvature which we know as the Big Bang singularity.

3.2.1 Linearised propagation equations

In an FLRW universe, we assume a perfect fluid spacetime in which

$$\sigma_{ab} = \omega_{ab} = \dot{u}^a = 0 \,, \tag{3.24}$$

implying that the density ρ, pressure p, and fluid expansion Θ are all functions of the cosmic time, defined by the fluid flow vector as $u_a = -\partial_a t$. For a perfect fluid, there is a unique covariant four-velocity, to which the projection tensor h_{ab} projects orthogonally. An almost-FLRW universe admits the following gauge-invariant variables: the shear, vorticity and acceleration,

$$\sigma_{ab} \equiv \nabla_{\langle a} u_{b\rangle} = h_a^c h_b^d \nabla_c u_d - \frac{1}{3}\theta h_{ab} \,, \tag{3.25}$$

$$\omega_{ab} \equiv h_a^c h_b^d \nabla_{[c} u_{d]} \,, \tag{3.26}$$

$$\dot{u}^a \equiv \nabla_b u^a u^b \,, \tag{3.27}$$

where the angle brackets represent the projected symmetric trace-free part of a tensor, i.e.

$$\nabla_{\langle c}u_{d\rangle} - \frac{1}{3}\nabla_c u^c h_{ab}\,; \tag{3.28}$$

the gravito-electric and gravito-magnetic components of the Weyl tensor C_{abcd},

$$E_{ab} \equiv C_{abcd}u^c u^d\,, \tag{3.29}$$

$$H_{ab} \equiv \frac{1}{2}\eta_{acd}C^{cd}{}_{be}u^e\,, \tag{3.30}$$

which are associated with gravitational radiation; and the matter tensor components,

$$q_a \equiv -h^a_c T_{cd}u^d\,, \tag{3.31}$$

$$\pi_{ab} \equiv h^c_a h^d_b T_{cd} - \frac{1}{3}\big(h^{cd}T_{cd}\big)h_{ab}\,. \tag{3.32}$$

All of the above variables vanish identically in an FLRW background, therefore any perturbation of them is automatically gauge-invariant by the Stewart-Walker lemma. However, the quantities representing the variation of the zeroth-order variables, i.e. the energy density ρ, pressure p, and fluid expansion Θ, are typically non-zero in expanding models, meaning they cannot be gauge invariant. To find associated gauge-invariant quantities, Ellis and Bruni [109] introduced the gauge-invariant orthogonal spatial gradients:

$$X_a \equiv \widetilde{\nabla}_a\rho\,, \tag{3.33a}$$

$$Y_a \equiv \widetilde{\nabla}_a p\,, \tag{3.33b}$$

$$Z_a \equiv \widetilde{\nabla}_a\Theta\,. \tag{3.33c}$$

These allow us to analyse the evolution of density perturbations on an FLRW background in a general case. Any first-order tensor X_{ab} orthogonal to u^a can be decomposed by repeated application of the $\widetilde{\nabla}_a \equiv h^b{}_a\nabla_b$ operator (3.13) to extract the scalar part of the perturbation variables. Many more gauge-invariant quantities can be found in the more complex invariantly defined quantities that vanish in an FLRW universe. However, only two more will be necessary, namely, the divergence of the acceleration and its spatial gradient:

$$A \equiv \nabla_\mu\dot{u}^\mu\,, \quad A_\mu \equiv \widetilde{\nabla}_\mu A\,. \tag{3.34}$$

To obtain the linearised equations of propagation, we begin by finding the energy and momentum conservation equations for a perfect fluid, where T_{ab} takes the form

$$T_{ab} = \rho u_a u_b + p h_{ab}\,. \tag{3.35}$$

Applying the conservation equation, $\nabla_b T_{ab} = 0$, and (3.11b) allows us to write

$$\begin{aligned}
0 &= \nabla_b(\rho u^a u^b + p h^{ab}) \\
&= \nabla_b(\rho u^a u^b) + \nabla_b\left[p(g^{ab} + u^a u^b)\right] \\
&= u^a u^b\nabla_b\rho + \rho u^b\nabla_b u^a + \rho u^a\nabla_b u^b + h^{ab}\nabla_b p + \rho\nabla_b g^{ab} + p u^b\nabla_b u^a + p u^a\nabla_b u^b \\
&= u^a\left[\Theta(\rho + p) + \dot{\rho}\right] + \dot{u}^a(\rho + p) + Y_a\,.
\end{aligned}$$

From this we can see that the energy and momentum conservation equations for perfect fluids take the forms

$$\Theta(\rho + p) + \dot{\rho} = 0\,, \tag{3.36a}$$

$$\dot{u}^a(\rho + p) + Y_a = 0\,. \tag{3.36b}$$

Taking the spatial gradient of (3.36a) then gives us the propagation equation for X_a (3.33b),

$$\begin{aligned}\widetilde{\nabla}_a[\Theta(\rho+p)+\dot{\rho}] &= (\rho+p)h^b{}_a\nabla_b\Theta + \Theta h^b{}_a\nabla(\rho+p) + h^b{}_a\nabla_b(u^c\nabla_c\rho)\\ &= (\rho+p)Z_a + \Theta(X_a+Y_a) + h^b{}_a\nabla_c\rho\nabla_b u^c + h^b{}_a u^c\nabla_b\nabla_c\rho\\ &= (\rho+p)Z_a + \Theta(X_a+Y_a) + h^b{}_a\nabla_c\rho\left(-u_b\dot{u}^c + \frac{1}{3}\Theta h^c{}_b + \sigma^c{}_b + \omega^c{}_b\right)\\ &= (\rho+p)Z_a + \Theta(X_a+Y_a) + \frac{1}{3}\Theta X_a + h^b{}_a u^c\nabla_b\nabla_c\rho\,.\end{aligned}$$

The last term on the right-hand side can be expressed as

$$h^b{}_a u^c\nabla_b\nabla_c\rho = h^b{}_a\nabla_b(u^c\nabla_c\rho) - h^b{}_a\nabla_b u^c\nabla_c\rho\,.$$

This can be simplified by expanding $h^b{}_a = h^d{}_a h^b{}_d$ to obtain

$$\begin{aligned}h^b{}_a u^c\nabla_b\nabla_c\rho &= h^d{}_a\nabla_c(u^c h^b{}_d\nabla_b\rho) - h^d{}_a(\nabla_c\rho h^b{}_d\nabla_b u^c)\\ &= h^d{}_a(X_a)^{\cdot} - \dot{\rho}\dot{u}^c\,.\end{aligned}$$

Here, the last term can be rewritten by utilising the acceleration equation,

$$\begin{aligned}\dot{u}^a &= \frac{Y^a}{\rho+p} = \frac{Y^a}{\dot{\rho}/\Theta}\\ \therefore\ \dot{\rho}\dot{u}^a &= Y^a\Theta\,,\end{aligned}$$

which cancels out the Y_a terms. Since the left-hand-side is zero from the conservation equations, we can write

$$\frac{4}{3}\Theta X_a + h^d{}_a(X_a)^{\cdot} + (\rho+p)Z_a = 0\,. \tag{3.37}$$

For the next step, we introduce three new gauge-invariant quantities: the fractional density gradient $\mathcal{X}_a$, comoving fractional density gradient $\mathcal{D}_a$, and comoving expansion gradient $\mathcal{Z}_a$. These are defined as:

$$\mathcal{X}_a \equiv \frac{X_a}{\rho}\,, \tag{3.38a}$$

$$\mathcal{D}_a \equiv a\mathcal{X}_a = \frac{a}{\rho}X_a\,, \tag{3.38b}$$

$$\mathcal{Z}_a \equiv aZ_a\,. \tag{3.38c}$$

The latter two quantities represent density variations at a fixed *comoving* scale, as opposed to changes to a *fixed* distance. This allows us to investigate relative density growth between neighbouring fluid comoving volumes. The vector $\mathcal{D}_a$ can be decomposed into magnitude $\mathcal{D}$ and direction e_a:

$$\mathcal{D}_a = \mathcal{D}e_a\,,\ e_a e^a = 1 \implies \mathcal{D} = (\mathcal{D}_a\mathcal{D}^a)^{1/2}\,. \tag{3.39}$$

This vector is associated with density variations, and can be decomposed into two parts: a solenoidal (or divergence-free) part and a gradient part. This means that $\mathcal{D}_a$ can be split into components based on whether or not they have divergence. The solenoidal part corresponds to the curl or vorticity aspects of the vector field, and does not contribute to changes in density due to compressive effects. The gradient part can be derived from a potential function, and it directly affects density variations due to compression or expansion. This component contributes to the divergence and thus to the density changes in the fluid.

Returning to the conservation equation, we can rewrite (3.37) as

$$\begin{aligned} 0 &= h^d{}_a(X_d)^{\cdot} + \frac{4}{3}X_a + (\rho+p)Z_a \\ &= \frac{a}{\rho}h^d{}_a(X_d)^{\cdot} + \frac{4}{3}\frac{a\Theta}{\rho}X_a + \frac{a(\rho+p)}{\rho}Z_a\,. \end{aligned}$$

To calculate the first term on the right-hand side, we make use of the relation

$$\begin{aligned} \left(\frac{a}{\rho}X_d\right)^{\cdot} &= \frac{\dot{a}}{\rho}X_d - \frac{a}{\rho^2}\dot{\rho}X_d + \frac{a}{\rho}(X_d)^{\cdot} \\ \therefore \frac{a}{\rho}(X_d)^{\cdot} &= (\mathcal{D}_d)^{\cdot} - \frac{a\Theta}{3\rho}X_d + \frac{\dot{\rho}}{\rho}\mathcal{D}_d \\ &= (\mathcal{D}_d)^{\cdot} - \frac{\Theta}{3}\mathcal{D}_d - (1+w)\Theta\mathcal{D}_d\,. \end{aligned}$$

Finally, we can derive the propagation equation for the comoving fractional density and expansion gradients,

$$h^d{}_a(\mathcal{D}_d)^{\cdot} = \Theta w\mathcal{D}_a - (1+w)\mathcal{Z}_a\,, \tag{3.40}$$

$$h^d{}_a(\mathcal{Z}_d)^{\cdot} = -\frac{2}{3}\Theta\mathcal{Z}_a - \frac{1}{2}K\rho\mathcal{D}_a + a\left(\frac{1}{2}\dot{u}_a\widetilde{R} + A_a\right). \tag{3.41}$$

We have defined $\widetilde{R}$ as the 3-Ricci scalar

$$\widetilde{R} = 2\left(-\frac{1}{3}\Theta^2 + A + K\rho + \Lambda\right) = \frac{6K}{a^2}\,, \tag{3.42}$$

where $K = 0, \pm 1$ is the curvature constant for the background FLRW universe. Because this model has spatial geometry with constant curvature, the commutation of spatial derivatives brings in the background curvature term Ka^{-2} through the spatial Ricci identity. This is a result of constant-curvature property of the spatial 3-Riemann tensor (see §2.1.1):

$${}^{(3)}R_{abcd} = 6Ka^{-2}(\gamma_{ac}\gamma_{bd} - \gamma_{ad}\gamma_{bc})\,, \tag{3.43}$$

where $\gamma_{ab} = a^{-2}g_{ab}$ is the metric tensor for a three-space of uniform spatial curvature K [159].

When the vorticity is zero, $\widetilde{R}$ becomes the Ricci scalar ${}^{(3)}R$ of the three-dimensional spaces orthogonal to the fluid flow [169]. When the vorticity of the fluid flow is zero, the flow is irrotational, allowing for of a family of hypersurfaces that are everywhere orthogonal to the flow lines. These hypersurfaces can then be treated as three-dimensional 'space-like' surfaces, essentially allowing a slicing of spacetime into three-dimensional spatial sections that are orthogonal to the four-velocity of the fluid. However, for a non-zero vorticity, the fluid flow has rotation, and it becomes impossible to define a family of hypersurfaces that are everywhere orthogonal to the flow. The flow lines 'twist' around each other, and there is no single spatial surface that can be orthogonal to the flow everywhere in space. Fortunately, we will be considering only irrotational flow.

We now have entirely general variables, which means that solving their evolution equations will be a time-consuming task. Since current observations suggest a universe very close to that of FLRW, we can save time by limiting our analysis to cases in which the real universe differs only slightly from the background. We therefore treat the vanishing (gauge-invariant) quantities as first order and the quantities that are non-vanishing in the background as zeroth order. This

leaves us with only the terms that are first order in the gauge-invariant perturbation variables. When we linearise the equations, we drop products because they are second order and keep quantities like $\Theta\mathcal{D}_a$, since Θ is a background quantity. Linearisation of equations (3.40) and (3.41) gives

$$\dot{\mathcal{D}}_a = \Theta w \mathcal{D}_a - (1+w)\mathcal{Z}_a\,, \tag{3.44}$$

$$\dot{\mathcal{Z}}_a = -\frac{2}{3}\Theta\mathcal{Z}_a - \frac{1}{2}K\rho\mathcal{D}_a + a\left(\frac{3K}{a^2} + A_a\right), \tag{3.45}$$

where the covariant derivatives can be those of the background model [109]. We now use the definition of A (3.34), the acceleration equation (3.2.1), and the gauge-invariant quantities in (3.38) to rewrite A_a as

$$A_a = \widetilde{\nabla}_a\left[\nabla^b\left(-\frac{\widetilde{\nabla}_b p}{\rho+p}\right)\right] \tag{3.46}$$

$$= -\frac{c_s^2}{a(1+w)}\widetilde{\nabla}_a\widetilde{\nabla}^b\mathcal{D}_b\,. \tag{3.47}$$

To obtain the second equality, we assumed an adiabatic evolution such that $c_s^2 \equiv \frac{\partial p}{\partial\rho}$, giving adiabatic perturbations.

To obtain the second-order equations, we follow Bardeen [159] by defining

$$\left(\frac{p}{\rho}\right)^{\cdot} \equiv \dot{w} = -(1+w)(c_s^2 - w)\Theta\,. \tag{3.48}$$

The linear second-order equation then follows from (3.44):

$$\ddot{\mathcal{D}}_a + \left[\frac{2}{3} - w + \left(c_s^2 - w\right)\right]\Theta\dot{\mathcal{D}}_a$$
$$- \left[\left(\frac{1}{2}(1-w)(3w+1) - 3\left(c_s^2 - w\right)\right)\rho + \left(c_s^2 - w\right)\frac{12K}{a^2}\right]\mathcal{D}_a + c_s^2\widetilde{\nabla}_a\left(\widetilde{\nabla}^b\mathcal{D}_b\right) = 0\,. \tag{3.49}$$

This allows for a time-varying equation of state w, but if we treat it as a constant $c_s^2 \approx w$ we obtain

$$\ddot{\mathcal{D}}_a + \left(\frac{2}{3} - w\right)\Theta\dot{\mathcal{D}}_a - \frac{1}{2}(1-w)(1+3w)\rho\mathcal{D}_a + w\left(\frac{2K}{a^2}\mathcal{D}_a - \widetilde{\nabla}^2\mathcal{D}_a\right) = 0\,, \tag{3.50}$$

where we used (B.2) to rewrite the last term. Instead of directly performing a harmonic decomposition of $\mathcal{D}_a$, we first take its spatial divergence to define a new scalar quantity,

$$\Delta = \widetilde{\nabla}^a\mathcal{D}_a\,. \tag{3.51}$$

This represents the local (covariant) density perturbation and remains gauge-invariant. Since it is a scalar, it is more convenient for applying a harmonic decomposition. Equation (3.50) therefore becomes

$$\ddot{\Delta} + \left(\frac{2}{3} - w\right)\Theta\dot{\Delta} - \frac{1}{2}(1-w)(1+3w)\rho\Delta - w\widetilde{\nabla}^2\Delta = 0\,. \tag{3.52}$$

For the dust-only model ($w = 0$) introduced in §2.1.3, (3.52) simplifies to

$$\ddot{\Delta} + \frac{2}{3}\Theta\dot{\Delta} - \frac{1}{2}\rho\Delta = 0\,. \tag{3.53}$$

3.2.2 Harmonic decomposition

It is standard practice [164, 159] to decompose the variables harmonically, effectively isolating the time and space variations. This transforms the differential equations governing the time variance of the perturbations into separate time equations for each component of spatial variation. Instead of a complex equation that includes both time and spatial derivatives, this method yields separate equations for time variation for each mode associated with a different spatial frequency. This separation simplifies the equations, as each mode evolves independently over time. Each component of spatial variation is associated with a Fourier wavenumber k which characterises the scale of each perturbation mode in a way that remains consistent as the Universe expands, i.e. a matter-comoving wavelength. We achieve this by writing $\mathcal{D}_a$ in terms of the comoving harmonic vectors $Q_a^{(n)}$. Given a second-order harmonic equation

$$\ddot{X} + A\dot{X} + BX + C = 0\,, \tag{3.54}$$

we can separate the variable X into $X(t,x) = X(t)Q(x)$. This can then be written as

$$X(t,\mathbf{x}) = \sum X^{(k)}(t)Q^{(k)}(\mathbf{x})\,, \tag{3.55}$$

where the $1+3$ covariant scalar harmonics $Q^{(n)}$ are defined as

$$\dot{Q}^{(k)} = 0\,, \quad \widetilde{\nabla}^2 Q^{(k)} = -\frac{k^2}{a^2}Q^{(n)}\,, \tag{3.56}$$

and k is the eigenvalue associated with the Fourier decomposition, which in the case of $K = 0$ is the *comoving wavenumber* – also known as the comoving momentum – related to the comoving wavelength λ via $k = 2\pi/\lambda$. Due to expansion, the comoving wavenumber is not the physical momentum; this is given by k/a. The normalisation of the scale factor to unity today means that the comoving momentum of a Fourier mode is equal to the physical momentum it has today [170]. We also see from the first equality that the spatial eigenfunctions $Q^{(n)}$ are time independent. This definition is just Bardeen's scalar Helmholtz equation [159],

$$\nabla^2 Q^{(0)} + k^2 Q^{(0)} = 0\,, \tag{3.57}$$

expressed covariantly as obtained by Hawking [164]. The eigenfunctions of the Laplacian operator ∇^2 form a complete set of scalar, vector, and tensor modes [171]:

$$\begin{aligned}\nabla^2 Q^{(0)} &= -k^2 Q^{(0)}\,,\\ \nabla^2 Q_i^{(\pm1)} &= -k^2 Q_i^{(\pm1)}\,,\\ \nabla^2 Q_{ij}^{(\pm2)} &= -k^2 Q_{ij}^{(\pm2)}\,.\end{aligned}$$

We express Δ in terms of the harmonics, getting

$$\Delta = \sum_k \Delta^{(k)} Q^{(k)}\,, \tag{3.58}$$

where $\Delta^{(k)}$ is the harmonic component of Δ corresponding to the matter-comoving wavenumber k. While $\Delta^{(k)}$ is often related to the density contrast, it is more accurately the covariant analogue of $\delta\rho^{(k)}/\bar{\rho}$. Substituting this decomposition into (3.52), the harmonics decouple and we obtain the k-th harmonic equation

$$0 = \ddot{\Delta}^{(k)} + \left(\frac{2}{3} - w\right)\Theta\dot{\Delta}^{(k)} - \left[\frac{1}{2}(1-w)(1+3w)\mu - w\frac{k^2}{a^2}\right]\Delta^{(k)}\,, \tag{3.59}$$

valid for $k \geq 0$. This demonstrates how the perturbative growth depends on the matter-comoving wavelength.

3.2.3 Jeans instability

The Jeans criterion determines the conditions under which small perturbations in a medium will grow due to gravitational instability, leading to collapse, or oscillate and disperse due to pressure. The former occurs if the effect of gravity exceeds the stabilising pressure force, which are related to the mass density and Laplace terms in (3.50), respectively. Collapse therefore tends to occur when the matter term – responsible for the gravitational pull – is greater than the Laplace term – which resists collapse, i.e. if

$$\frac{1}{2}(1-w)(1+3w)\rho\Delta > w\left(\frac{2K}{a^2}\Delta - \widetilde{\nabla}^2\Delta\right) \tag{3.60}$$

when $c_s^2 = w$. With the harmonic decomposition used for (3.59), this translates to the inequality

$$\frac{1}{2}(1-w)(1+3w)\rho > w\frac{k^2}{a^2}\,. \tag{3.61}$$

We can therefore define the Jeans wavenumber in this case as

$$k_J \equiv \left[(1-w)\left(\frac{1}{w}+3\right)\frac{\rho(t)}{2}\right]^{1/2} a(t) > k\,, \tag{3.62}$$

which gives a wavelength – called the **Jeans length**:

$$\lambda_J \equiv \frac{2\pi a(t)}{k_J} = c_s c\sqrt{\frac{\pi}{G\rho(t)}\frac{1}{(1-w)(1+3w)}}\,, \tag{3.63}$$

where, here only, we have reinstituted the speed of light c and gravitational constant, so that $w = (c_s/c)^2$. Despite inhomogeneities having a constant matter-comoving size and thus a constant comoving wavelength, the Jeans length still varies with time. During the radiation era, $a \propto t^{1/2}$ and $\rho_m \propto t^{-3/2}$, giving a comoving Jeans length of $\lambda_{J,\mathrm{rad}} = \sqrt{\pi t^2/3}$ which increases to a maximum as the Universe approaches matter-radiation equality. The Jeans mass will therefore grow as

$$M_J = \frac{4\pi}{3}\rho_m\lambda_J^3 \propto t^{3/2}\,, \tag{3.64}$$

until recombination is reached, at which point it will stay constant. After this the Universe is matter-dominated and the sound speed of the coupled fluid becomes dependent on the matter density. Following recombination, the Jeans length and mass will rapidly decline due to decoupling of matter and radiation, leading to $c_s \to 0$ and thus $\lambda_J \to 0$. Wavelength exceeding $\lambda_J^{\max}$ will grow until it is overtaken by the Jeans wavelength, at which point it will undergo acoustic oscillations until decoupling begins and the Jeans length drops to as matter-dominated growth starts. When the Universe becomes curvature-dominated at late times – dependent on Ω_0 – the growth of small perturbations dies down.

We generally assume that perturbations remain unaffected by strong interactions in the early universe present after the initial inflation. We therefore only need to consider the times (i) t_J, when wavelengths drop below the Jeans length and acoustic oscillations begin to decay as $t^{-1/6}$; (ii) t_{eq}, the beginning of matter domination; and (iii) t_{dec}, when decoupling takes place and the acoustic oscillations have died away. In a ΛCDM universe, the CDM fluctuations start to grow at t_{eq}, earlier than baryon fluctuations [166]. It is these CDM perturbations that attract the baryons to potential wells and seed the growth of inhomogeneities.

3.3 Matter power spectrum

Another way to solve differential equations is to express δ in terms of its Fourier components. If we consider the Universe to be a box of comoving volume V, the fluctuation field can be written as [93]

$$\delta(k) = \frac{V}{(2\pi)^3} \int \delta_k e^{-ik\cdot k} \mathrm{d}^3 k \,, \tag{3.65}$$

where each Fourier mode δ_k is calculated by performing the integral

$$\delta_k = \frac{1}{V} \int \delta(k) e^{-ik\cdot k} \mathrm{d}^3 r \,. \tag{3.66}$$

The solution is of the form

$$\delta_k = |\delta_k| e^{i\phi_k} \,. \tag{3.67}$$

The Fourier mode of Δ is related to the Fourier mode of the density contrast via $\Delta_k = -k^2\delta_k$. For very small fluctuations, $|\delta_k| \ll 1$, the phase ϕ_k remains small and each Fourier component obeys

$$\ddot{\delta}_k + 2H\dot{\delta}_k - \frac{3}{2}\Omega_m H^2 \delta_k = 0 \,. \tag{3.68}$$

The full general-relativistic equations governing the evolution of these density fluctuations under gravity can be derived by applying the same approach used for a perfect fluid instead of a scalar field. The resulting equations are identical to those derived from Newtonian theory, except for a term that only affects super-horizon scales [44]. Therefore, on sub-horizon scales, we can approximate these potential fluctuations as obeying the perturbed Poisson equation in Newtonian gravity:

$$\delta\Phi_k = -\frac{4\pi G\bar{\rho}a^2}{k^2}\delta_k \,. \tag{3.69}$$

When the phases ϕ_k of the Fourier components are uncorrelated, it is known as a Gaussian field. Interestingly, if a Gaussian field is also homogeneous and isotropic, its statistical properties are fully described by the power spectrum $P(k)$. Since a number of inflationary scenarios posit that the density fluctuations constitute a homogeneous, isotropic field, Gaussian density fields are of great interest to cosmologist. Returning to equation 3.69, substituting $\bar{\rho} = 3H^2/(8\pi G)$ yields

$$\delta_k = -\frac{2}{3}\left(\frac{k}{aH}\right)^2 \delta\Phi_k \,, \tag{3.70}$$

from which we can deduce that $\langle|\delta(k)|\rangle \propto k^4 \langle|\delta\Phi_k|^2\rangle$. We therefore define the **matter power spectrum** as the mean square amplitude of the Fourier components:

$$\boxed{P(k) \equiv |\delta^2(k)|} \,, \tag{3.71}$$

where the average is taken over all possible wavenumber orientations. The power spectrum is generally parameterised at $P_\delta(k) \propto k^n$, where n is the primordial spectral index. Inflation favours a power-law index $n = 1$, known as the Harrison-Zel'dovich spectrum. The value of n can be deduced from temperature fluctuations in the CMB on large scales, and those measured by Planck [107] suggest a value of $n = 0.97 \pm 0.01$, only slightly off the Harrison-Zel'dovich spectrum [93].

Just after its inception, the Universe expanded extremely rapidly – by at least 60 e-folds in a fraction of a second, as we showed in §2.2.1. This caused many perturbation scales to be stretched beyond the Hubble horizon and become unaffected by causal processes. As expansion

slowed down, the Hubble radius grew and the stretched perturbations began to 're-enter' the horizon, such that their scales became smaller than the Hubble radius again. Looking again at equation 3.70, we note that if we do not define the spectrum at a specific cosmic time but rather evaluate it when a given scale re-enters the horizon $k = aH$, we can write

$$\delta_k \propto \Phi_k \propto \frac{1}{(aH)^2}\,. \tag{3.72}$$

A radiation-dominated model exhibits $a \propto t^{1/2}$ and $H = 1/(2t)$, while a matter-dominated model has $a \propto t^{2/3}$ and $H = 2/(3t)$ [44]. We therefore find that

$$\delta_k(t) \propto \begin{cases} t & \text{radiation-dominated,} \\ t^{2/3} & \text{matter-dominated.} \end{cases}$$

Every fluctuation, regardless of wavelength, will eventually cross the horizon. This happens because the Hubble radius increases linearly over time, while the proper wavelength $a\lambda$ grows at a slower rate. Consequently, each fluctuation will inevitably shift from a super-horizon scale ($k \ll aH$) to a sub-horizon scale ($k \gg aH$) [172]. At super-horizon scales, the perturbation amplitudes grow with time before they reach the Hubble horizon, as which point growth stops.

The power spectrum provides a statistical measure of density fluctuations in the matter distribution. It is therefore not an observable quantity but can be easily reconstructed from galaxy surveys. To obtain it, we need to consider the effects of dark matter and the background cosmology on the evolution of perturbations. The two primary modes of fluctuation are adiabatic and isocurvature. The entropy of the former does not vary spatially due to the coupled fluctuations in the matter and radiation components. The latter does involve entropy fluctuations, with zero net fluctuation in the energy density [66]. Most modern models predict adiabatic fluctuations.

Per the classical Jeans instability criterion, pressure inhibits structure growth on scales smaller than Jeans length

$$\lambda_J = c_s\sqrt{\frac{\pi}{G\rho}}\,, \tag{3.73}$$

where G has been reinstated for dimensional clarity. This is simply (3.63) with $w = 1$, since we are working in the dust case for now. This is the distance travelled by an acoustic wave during the free-fall collapse time of a perturbation. Oscillations on scales larger than this become unstable to gravitational collapse. When a fluctuation becomes sub-horizon, dissipative processes alter the shape of the spectrum depending on its scale. If collisionless particles of HDM or relativistic particles and radiation are present, they can traverse through the background rapidly, damping perturbations (see figure 3.2). Imperfect couplings of photons and baryons may also cause perturbations in the baryonic component to dissipate. For uniform initial Gaussian fluctuations, the cumulative result of these processes is a deformation of the original power spectrum's shape. This change is described by the transfer function $T(k)$, a simple function of the wavenumber, which relates the power spectrum $P(k)$ to its initial value $P_0(k)$ via $P(k) = P_0(k) \times T^2(k)$.

Relativistic particles at matter-radiation equality, like light neutrinos, are known as hot dark matter (HDM); their large free-streaming length significantly damps the power spectrum across a wide range of k values. Weakly Interacting Massive Particles (WIMPs), which are nonrelativistic at matter-radiation equality and called cold dark matter (CDM), have minimal impact on the power spectrum. Baryons, tightly coupled to radiation by electron scattering before recombination, experience increased photon mean-free paths once recombination occurs. As photons escape dense regions, they pull baryons along, smoothing out small-scale density fluctuations in

a process called *Silk damping*. This results in damped oscillations of the baryon-photon fluid on sub-horizon scales [66]. We can interpret these results graphically in figure 3.2 with the transfer functions of particular models. Free-streaming effects of HDM particles eliminate small-scale structure such that $T(k) \to 0$ exponentially for large k. CDM experiences similar to a lesser degree, where kinematic suppression occurs on scales less than the horizon size at matter-radiation equality. These two scenarios have great implications for late-stage structure formation: HDM initially produces superclusters which fragment into galaxies in a 'top-down' scheme; conversely, CDM begins with small-scale structures that coalesce into larger ones in a 'bottom-up' scheme.

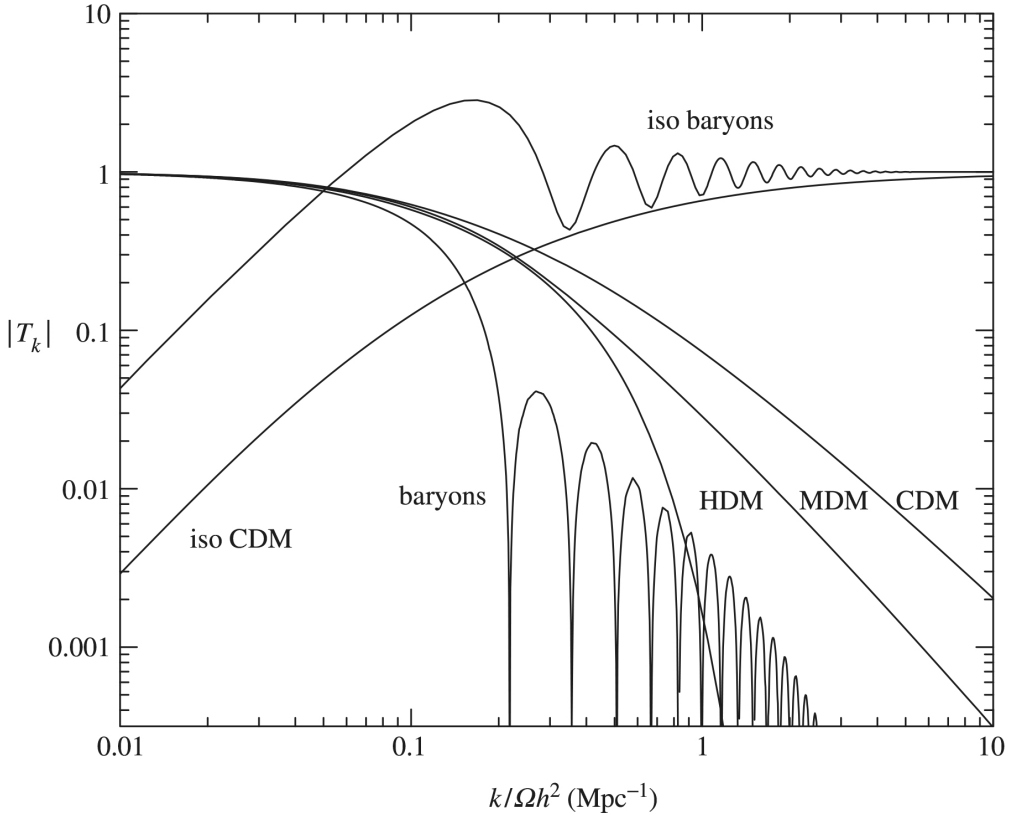


Figure 3.2: Examples of adiabatic transfer functions – for which $T_k \to 1$ at small k – for baryons, hot dark matter (HDM), cold dark matter (CDM) and mixed dark matter (MDM), as well as isocurvature modes (iso baryons and iso CDM). Image from Coles and Lucchin (2002) [66], adapted from Peacock (2003) [173].

3.4 Summary

In this chapter, we acknowledged that the Universe is by no means an ideal, smooth space, and we need to take this into account when studying structure formation. The universe we know actually evolved from initial inhomogeneities – overdensities that created potential wells for the seeds of large-scale structure. To begin, we looked at the types of perturbations and their formation mechanisms, inevitably leading to a discussion of the gauge problem. To avoid the gauge issues in metric perturbation theory, we turned to the $1 + 3$ covariant perturbation approach. This allowed us to recover key results such as the Raychaudhuri, continuity, and conservation equations. The next step involved utilising gauge-invariant variables to analyse perturbations, where we ultimately derived the differential equations for the comoving fractional density and expansion gradients. By adopting a harmonic decomposition of variables, we isolated

the time and spatial variations, enabling us to derive Jeans' instability criteria and examine the significance of the Jeans length in the context of cosmic evolution. We finished with a simplified analysis using Newtonian perturbation theory, which allowed us to derive the density contrast equation. We then showed how to derive the power spectrum from the Fourier expansion, and what the spectrum tells us about perturbations in the early universe. Throughout the chapter, we remained within the framework of general relativity. In the next chapter, we will apply all of this to modified gravity – namely $f(R)$.

Chapter 4

Modified Gravity

In this chapter, we formulate a model-independent, cosmography-based approach to $f(R)$ cosmologies that mimic ΛCDM [36], and analyse the growth factor. The last three chapters have covered all the tools and techniques we need to tackle the main problem of the thesis. By expressing terms involving $f(R)$ and its derivatives in terms of cosmographic parameters (q, j, s, l) within the $f(R)$ $1+3$ covariant perturbation scheme, we will be able to formulate a model-independent analysis of $f(R)$ without pre-specifying the functional form of f.

We first review the model-independent analysis of the background, highlighting the effectiveness of cosmography in describing the evolution of the cosmological expansion history. The exact linear perturbation equations are then reformulated, and the quasi-static (QS) assumptions are separately applied to derive the semi- and fully quasi-static approximation equations. The cosmographic condition $j = 1$ is adopted to specify a background cosmic evolution that adheres to ΛCDM. Following this, we analyse the dynamics of matter perturbations using the $1+3$ CGI formalism. We finally turn to the evolution of the growth rate function and the growth index parameter across different scales under exact, semi-QS, and full-QS approximations, comparing the results to model-specific $f(R)$ predictions. We also explore the scale-dependent dispersion of the growth index parameter, a hallmark of modified gravity theories.

4.1 $f(R)$ gravity

The most prevalent of the modified gravity theories, $f(R)$, is the focus of this chapter. In this model, the Ricci scalar in the Einstein-Hilbert action is replaced with a general function of itself. Since this will show itself in large-scale clustering, accurate observations are required in both the background evolution and large-scale structure to reconstruct the theory. Traditional methods first specify the form of the modified gravity model, e.g. Hu-Sawicki [174], Starobinsky [175], etc., and fit the parameters using observational data. This method is not particularly efficient, as one needs to correctly estimate a model before even constructing a dynamical system. The added observational errors also make it difficult to pin down a specific form of the model.

In $f(R)$ gravity, we want to generalise the Lagrangian of the four-dimensional Einstein-Hilbert action (1.64). The Ricci scalar $R = g^{\mu\nu}R_{\mu\nu}$ is replaced by a function of the Ricci scalar, giving

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2}f(R) + \mathcal{L}_m(g_{\mu\nu}, \Psi_m)\right], \tag{4.1}$$

where g is the determinant of the metric $g_{\mu\nu}$ and $\mathcal{L}_m$ is the matter Lagrangian, which is dependent on the metric and the matter fields Ψ_m. These matter fields obey standard conservation equations and the metric $g_{\mu\nu}$ therefore corresponds to the physical frame. There are two methods to derive the field equations from the action [64]:

Metric: In the metric formalism, the connections $\Gamma^{\alpha}_{\beta\gamma}$ are the standard connections defined in terms of the metric $g_{\mu\nu}$. The field equations are obtained by varying the action (4.1) with respect to $g_{\mu\nu}$.

Palatini: In the Palatini formalism, $\Gamma^{\alpha}_{\beta\gamma}$ and $g_{\mu\nu}$ are treated as independent variables. The action (4.1) is varied with respect to both the metric $g_{\mu\nu}$ and the connection $\Gamma^{\alpha}_{\beta\gamma}$.

While the metric approach typically results in higher-order derivative field equations, the Palatini formalism yields gravitational field equations that are consistently second-order partial differential equations. We will be making use of the former.

By the principle of least action[1], the variation of the action (4.1) with respect to the inverse must yield zero, and we can use this to derive the equations of motion:

$$\begin{aligned}
0 &= \delta S \\
&= \delta \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2}f(R) + \mathcal{L}_m(g_{\mu\nu}, \Psi_m)\right] \\
&= \int \mathrm{d}^4x \left[\frac{1}{2}\frac{\delta(\sqrt{-g}f(R))}{\delta g^{\mu\nu}} + \frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu} \\
&= \int \mathrm{d}^4x \left[\frac{1}{2}\left(\sqrt{-g}\frac{\delta f(R)}{\delta g^{\mu\nu}} + f(R)\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}}\right) + \frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu} \\
&= \int \mathrm{d}^4x \left[\frac{1}{2}\left(\frac{\delta f(R)}{\delta g^{\mu\nu}} + \frac{f(R)}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}}\right) + \frac{1}{\sqrt{-g}}\frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\right]\sqrt{-g}\delta g^{\mu\nu}\,.
\end{aligned}$$

Since this equation should hold for all $\delta g^{\mu\nu}$, the two terms in the square brackets should sum to zero, thus

$$\frac{\delta f(R)}{\delta g^{\mu\nu}} + \frac{f(R)}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}} = -\frac{2}{\sqrt{-g}}\frac{\delta\left(\sqrt{-g}\mathcal{L}_m\right)}{\delta g^{\mu\nu}}\,.$$

In the case of scalar fields, the energy-momentum tensor can be written as (1.66) [48], leaving us with

$$\frac{\delta f(R)}{\delta g^{\mu\nu}} + \frac{f(R)}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}} = T_{\mu\nu}\,. \tag{4.2}$$

The variation of $f(R)$ yields $f'(R)\delta R$, so we need to calculate the variations of both the metric and the Ricci scalar. Returning to the variation of the action, we can substitute (1.66), (1.68) and (1.75) to obtain

$$\begin{aligned}
\delta S &= \int \mathrm{d}^4x\left[\frac{1}{2}\left(\delta f(R)\sqrt{-g} + f(R)\delta\sqrt{-g}\right) - \frac{1}{2}T_{\mu\nu}\sqrt{-g}\delta g^{\mu\nu}\right] \\
&= \int \mathrm{d}^4x\left[\frac{1}{2}\left(F(R)\delta R\sqrt{-g} - \frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}f(R)\right) - \frac{1}{2}T_{\mu\nu}\sqrt{-g}\delta g^{\mu\nu}\right] \\
&= \int \mathrm{d}^4x\left[\frac{1}{2}\sqrt{-g}\left(F(R)(R_{\mu\nu}\delta g^{\mu\nu} + g_{\mu\nu}\Box\delta g^{\mu\nu} - \nabla_\mu\nabla_\nu\delta g^{\mu\nu}) - \frac{1}{2}g_{\mu\nu}\delta g^{\mu\nu}f(R) - \frac{1}{2}T_{\mu\nu}\delta g^{\mu\nu}\right)\right],
\end{aligned}$$

where $F(R) \equiv f'(R)$. Integration by parts on the second and third terms – while discarding total divergences – yields

$$\delta S = \int \mathrm{d}^4x\sqrt{-g}\delta g^{\mu\nu}\left[\frac{1}{2}\left(F(R)R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}f(R) + [g_{\mu\nu}\Box - \nabla_\mu\nabla_\nu]F(R)\right) - \frac{1}{2}T_{\mu\nu}\right]. \tag{4.3}$$

[1]The solutions for the equations of motion are stationary points of the action.

Demanding that the action remains invariant under variations of the metric, i.e. $\delta S/\delta g^{\mu\nu} = 0$, we obtain the ***f*(*R*) field equations**:

$$F(R)R_{\mu\nu} - \frac{1}{2}f(R)g_{\mu\nu} + [g_{\mu\nu}\Box - \nabla_\mu\nabla_\nu]\,F(R) = T_{\mu\nu}\,. \tag{4.4}$$

The trace of this equation is

$$3\Box F(R) + RF(R) - 2f(R) = T\,, \tag{4.5}$$

where $T = g^{\mu\nu}T_{\mu\nu}$ and $\Box F = (1/\sqrt{-g})\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu F)$. This gives us the propagating scalar degree of freedom $F(R)$, referred to hereafter as the *scalaron*. In the case of GR the scalaron is simply $F(R) = 1$ since $f(R) = R - 3\Lambda \to R$ [47], and the $\Box F(R)$ term vanishes[2]. In the presence of a perfect fluid ($p = w\rho$) with energy density ρ and pressure p, the $f(R)$ gravity field equations become

$$3F\left(H^2 + \frac{K}{a^2}\right) = \rho + \rho_R\,, \tag{4.6a}$$

$$-F\left(2\dot{H} + 3H^2 + \frac{K}{a^2}\right) = p + p_R\,, \tag{4.6b}$$

where the scalaron energy density and pressure are defined as

$$\rho_R \equiv \frac{1}{2}\left(RF - f\right) - 3H\dot{F}\,, \tag{4.7a}$$

$$p_R \equiv \ddot{F} + 2H\dot{F} - \frac{1}{2}\left(RF - f\right)\,, \tag{4.7b}$$

where dots represent differentiation with respect to time t. We can combine equations 4.6b and 4.7b to obtain the **modified Friedmann** and **Raychaudhuri equations** for a flat ($K = 0$) FLRW universe:

$$3H^2 = \frac{1}{f'}\left[\rho_m + \frac{1}{2}\left(Rf' - f\right) - 3Hf''\dot{R}\right], \tag{4.8}$$

$$3\dot{H} + 3H^2 = -\frac{1}{2f'}\left[\rho_m + 3p_m + f - f'R + 3Hf''\dot{R} + 3f'''\dot{R}^2 + 3f''\ddot{R}\right], \tag{4.9}$$

where ρ_m represents the energy density of standard matter, H and R represent the Hubble parameter and Ricci scalar defined in (2.54) and (1.52), respectively, and a prime represents the derivative with respect to R.

4.1.1 Viability of $f(R)$

In terms of cosmological dynamics, an $f(R)$ theory should display smooth transitions between cosmological eras. It must provide an inflationary period which solves the flatness, horizon, and monopole problems. This needs to be followed by a radiation-dominated epoch which leads to a matter-dominated era. The evolution should finally evolve towards a de Sitter-type expansion. The stability of $f(R)$ gravity depends on the following conditions:

- $f'(R) > 0$ for all R to avoid a ghost degree of freedom that appears when $f'(R) < 0$. This also guarantees that gravity is attractive with a positive effective gravitational constant.

[2]Contrary to the metric formalism, the Palatini approach does not introduce new degrees of freedom and the scalaron does not propagate freely [176].

- $f''(R) > 0$ during the early epoch of matter domination, as $f''(R) < 0$ gives rise to the Dolgov-Kawasaki instability[3] [177]. This may result in a weak sudden singularity if $f''(R) = 0$ for a finite scalar curvature.

The existence of a stable Newtonian limit is linked to the presence of a stable ground state in the theory, be it Minkowski, de Sitter, or anti-de Sitter. To obtain a stable weak-field limit – applicable in regimes for R where Newtonian gravity holds – $f(R)$ must satisfy

$$|f(R) - R| \ll R \implies |f'(R) - 1| \ll 1\,, \tag{4.10}$$

$$Rf''(R) \ll 1\,, \tag{4.11}$$

$$f(R) \to R - 2\Lambda\,, \tag{4.12}$$

for $R \gg R_0$, guaranteeing that deviations from GR remain minimal [31]. Additional constraints arise from the cosmological and solar system tests, requiring

$$\lim_{R\to\infty} \tilde{f}(R) = \text{const.} \quad \text{and} \quad \lim_{R\to 0} \tilde{f}(R) = 0\,. \tag{4.13}$$

To summarise, $f(R)$ must accurately reproduce the gravitational field in and around compact objects; accelerate the Universe at late times; generate an expansion history consistent with observations; and align with the growth of large-scale structures. One of the first models to exhibit inflation in the early universe was proposed by Starobinsky in 1980 [112], $f(R) = R + \alpha R^2$ ($\alpha > 0$). The quadratic term leads to an accelerated expansion and an asymptotically exact de Sitter solution [47]. However, this term is also negligible relative to R at the present epoch, so the model cannot account for the current accelerated expansion. To mitigate this, models of the form $f(R) = R + \alpha R^n$ ($\alpha, n > 0$) were proposed. While these models could drive cosmic acceleration, they often failed to satisfy local gravity constraints due to the instability associated with negative $f''(R)$, which can lead to a tachyonic scalar degree of freedom. Additionally, many such models struggled to reproduce a standard matter-dominated epoch [64]. One challenge in $f(R)$ gravity is the degeneracy among models; many can be tuned to match specific observational constraints but may fail others, such as local gravity tests or structure formation requirements, making it difficult to distinguish the most viable ones.

Moreover, some models may require fine-tuning of parameters to match observations, raising questions about their validity. For instance, the value of certain parameters may need to be very specific to avoid conflicts with data. The Newtonian limit and adherence to energy conditions impose additional constraints, while deviations from GR – in the Hubble and deceleration parameters – typically emerge only at low redshifts ($z < 1$), limiting their distinct signatures. Moreover, the scalar field introduced by $f(R)$ theories can lead to complexities such as higher-order field equations in the metric formulation, or additional algebraic relations in the Palatini formulation. Rigorous observational tests, including those using the CMB, weak gravitational lensing, the Integrated Sachs-Wolfe (ISW) effect, and gravitational wave propagation, are essential for constraining $f(R)$ models. $f(R)$ gravity still requires extensive observational analysis to address these limitations and identify the most viable theories. Finally, though beyond the scope of this thesis, we also need to consider the fact that many modified gravity theories, including $f(R)$, have not been thoroughly quantised, leading to uncertainty surrounding their consistency at quantum scales. The lack of a well-defined quantum theory could limit their applicability in high-energy regimes.

[3] This implies a modification of the form $1/R$, which causes a temporal instability in the growth of curvature perturbations.

4.2 Dynamical systems analysis

One can solve the field equations (4.6b) for a chosen $f(R)$ theory, but it can be beneficial to determine the form of $f(R)$ that yields a desired solution. This is achieved via the various *reconstruction* methods of $f(R)$ gravity, described by Nojiri and Odintsov in [178]. In this approach, it is assumed that the Universe's expansion history is precisely known, and the field equations are inverted to identify the class of $f(R)$ theories that can produce a given FLRW model. Such solutions are particularly significant because, in FLRW backgrounds, they often represent asymptotic or intermediate states within the broader phase-space of potential cosmological evolutions. The reconstruction methods call for the Ricci scalar R to be recast as a function of a cosmological variable in such a way that the function can be inverted. Different reconstruction approaches express R in terms of different variables. Even if $R(t)$ is non-invertible, $R = R(a)$, $R(H)$, or $R(\tau)$ might be. The particular method must be chosen well, since a number fail to deliver the desired cosmology where others succeed [179].

If a reconstruction differential equation is successfully derived, there is no guarantee that a compact form of the general solution can be found, and if it is, it may involve functions too complex to be solved analytically. For instance, as the ΛCDM model aligns best with observations, efforts have been made to reconstruct $f(R)$ theories that replicate the ΛCDM evolution history precisely without introducing a Λ-term. These attempts often result in compact solutions expressed through Hypergeometric functions [180]. However, when delving deeper into the perturbative regime to determine if perturbation-dependent observables can differentiate between the $R - 2\Lambda$ and $f(R)$ models, hypergeometric functions pose significant challenges due to their complexity [36]. In the following sections we will attempt to understand how the linear perturbation scheme of ΛCDM differs from $F(R)$ theories while bypassing the reconstruction method completely, as done in [36, 37].

The complexity of the fourth-order field equations in these theories often hinders efforts to fully understand their physics, making it challenging to derive exact or numerical solutions that can be directly compared with observations. However, one particularly effective approach leverages the theory of dynamical systems [181, 182] has demonstrated great success in providing a straightforward framework for deriving exact solutions and qualitatively describing the global dynamics of these models. Bahamonde et al. provide a thorough review of the dynamical systems approach and its application to various GR and modified gravity theories [183]. The most common method for formulating a dynamical system in terms of expansion-normalised variables was first used in [184], based on a paper by Goliath and Ellis [185]. The expansion-normalised dynamical dimensionless variables present in $f(R)$ gravity can be expressed in terms of the new variables

$$x = \frac{\dot{F}}{HF}\,, \quad y = \frac{R}{6H^2}\,, \quad z = \frac{f}{6FH^2}\,, \quad \Omega = \frac{\rho}{3FH^2}\,, \tag{4.14}$$

where, for a general system, the Ricci scalar is

$$R = 6\left(\dot{H} + 2H^2 + \frac{K}{a^2}\right)\,. \tag{4.15}$$

The Friedmann constraint equation then becomes

$$-x + y - z - K + \Omega = 1\,. \tag{4.16}$$

To form the dynamical system, these variables (4.14) are differentiated with respect to the number of e-foldings, $\eta = \ln a$, using (4.6b), the fact that $\mathrm{d}\eta = H\mathrm{d}t$, and the time derivatives

$$\dot{R} = 6\left(\ddot{H} + 4H\dot{H} - \frac{2KH}{a^2}\right), \tag{4.17}$$

$$\dot{\rho} = -3H\left(1+w\right)\rho\,. \tag{4.18}$$

K is then eliminated from the system using (4.16) to give

$$\frac{\mathrm{d}x}{\mathrm{d}\eta} = -2x^2 + 2y - (z+2)x - 4z + \Omega(x + 1 - 3w)\,, \tag{4.19a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}\eta} = y\left[x(\Gamma - 2) - 2(z-1) + 2\Omega\right], \tag{4.19b}$$

$$\frac{\mathrm{d}z}{\mathrm{d}\eta} = xy\Gamma + z(-3x - 2z + 2\Omega + 2)\,, \tag{4.19c}$$

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\eta} = \Omega(-3x - 2z + 2\Omega - 3w - 1)\,, \tag{4.19d}$$

where $\Gamma = \Gamma(R)$ is defined as

$$\Gamma(R) \equiv \frac{\mathrm{d}\ln R}{\mathrm{d}\ln F} = \frac{F}{RF'} = \frac{\dot{R}F}{R\dot{F}}\,. \tag{4.20}$$

To close the system, we need to express $\Gamma(R)$ in terms of the dynamical variables x, y, z, and Ω. This requires determining R as a function of y/z and subsequently $\Gamma(y/z)$, which depends on whether the relation $y/z = RF/y$ is invertible. In other words, R must be uniquely solvable for every combination of y and z. This constraint has historically limited the types of $f(R)$ models that can be studied using this approach. To address this issue, alternative methods have been proposed, including introducing additional parameters or variables. In [186], the dynamical system was reformulated using the variables

$$A = \left(\frac{H}{m}\right)^2, \quad Q = \frac{3}{2H}\frac{\mathrm{d}H}{\mathrm{d}\eta}, \quad J = \frac{1}{4H}\frac{\mathrm{d}^2H}{\mathrm{d}\eta^2}, \tag{4.21}$$

where m is a constant with units of mass, specific to the $f(R)$ theory under consideration. The quantities Q and J are related to the cosmographic deceleration (q) and jerk (j) parameters, respectively. Reformulating the system in terms of these variables results in a four-dimensional phase space. This approach introduces two auxiliary quantities, X and Y, replacing the single Γ used in earlier formulations. By expressing X and Y in terms of H^2 and R, the system becomes closed and autonomous, eliminating the need for an invertible function $R(y/z)$. While this method can accommodate any $f(R)$ form, the resulting expressions for $X(A, y)$ and $Y(A, y)$ may become highly complex. A drawback of this and similar formulations is their reliance on a top-down approach, requiring the functional form of $f(R)$ to be specified in advance before analysis. This model-by-model investigation is a time-consuming process. Ideally, one would prefer to reconstruct the gravitational theory directly from observations or at least constrain it to ensure consistency with empirical data.

For this reason, we will make use of a bottom-up approach, where the form of $f(R)$ is left entirely general but must satisfy specific cosmological conditions. To make the dynamical system (4.19) autonomous, the definition of Γ (4.20) can be used, along with the Friedmann constraint (4.16),

to obtain

$$\begin{aligned} xy\Gamma &= \frac{\dot{R}}{6H^3} \\ &= \frac{\ddot{H}}{H^3} + 4\,(y - 2 - K) - 2K \\ &= \frac{\ddot{H}}{H^3} - 2(y+1) + 6(x + z - \Omega)\,. \end{aligned} \tag{4.22}$$

However, the first term is expressed in a way which is unsuitable for dynamical systems analysis. This can be remedied by introducing four cosmographic parameters, deceleration, jerk, snap, and lerk defined in (2.59), which can be expressed as

$$q = -1 - \frac{\dot{H}}{H^2}\,, \tag{4.23a}$$

$$j = \frac{\ddot{H}}{H^3} - 3q - 2\,, \tag{4.23b}$$

$$s = \frac{\dddot{H}}{H^4} + 4j + 3q(q+4) + 6\,, \tag{4.23c}$$

$$l = \frac{H^{(4)}}{H^5} + 5s - 10(j+3q)(q+2) - 24\,. \tag{4.23d}$$

Their derivatives are

$$\frac{\mathrm{d}q}{\mathrm{d}\eta} = 2q^2 + q - j\,, \tag{4.24a}$$

$$\frac{\mathrm{d}j}{\mathrm{d}\eta} = s + j(2+3q)\,, \tag{4.24b}$$

$$\frac{\mathrm{d}s}{\mathrm{d}\eta} = l + s(3+4q)\,, \tag{4.24c}$$

$$\frac{\mathrm{d}l}{\mathrm{d}\eta} = m + l(4+5q)\,, \tag{4.24d}$$

which can be related to each other via

$$j = 2q^2 + q - \frac{\mathrm{d}q}{\mathrm{d}\eta}\,, \tag{4.25a}$$

$$s = \frac{\mathrm{d}j}{\mathrm{d}\eta} - j(2+3q)\,, \tag{4.25b}$$

$$l = \frac{\mathrm{d}s}{\mathrm{d}\eta} - s(3+4q)\,. \tag{4.25c}$$

As we discussed when deriving equations (2.59), it's cosmographic parameters all the way down, but any solution of some $f(R)$ theory can be expressed as a finite set of algebraic equations relating the cosmographic parameters. This makes it unnecessary to extend our analysis beyond lerk. A specific cosmic evolution, whether derived from GR or an $f(R)$ theory, can always be described by an algebraic relationship involving a finite set of cosmographic parameters. To illustrate this, consider a straightforward example from GR. The cosmic evolution associated with the standard ΛCDM model of cosmology is a solution of GR with a cosmological constant and satisfies the corresponding field equation

$$H^2 + \frac{K}{a^2} = \frac{\rho}{3a^2} + \frac{\Lambda}{3}\,. \tag{4.26}$$

Following [91], we use the method of eliminating parameters and redefining them as constants of integration. We consider (4.26) and its first two time derivatives:

$$\dot{H} - \frac{K}{a^2} = -\frac{\rho_0}{2a^3}\,, \tag{4.27}$$

$$\frac{\ddot{H}}{H} + \frac{2K}{a^2} = \frac{3\rho_0}{2a^3}\,. \tag{4.28}$$

Since we are working with dust, energy conservation means that $\rho_0 a^3$ is constant. We can therefore treat (4.26) and (4.28) as a system of algebraic equations for the constants $(K/a^2, \rho_0/a^3, \Lambda)$, which can be expressed in terms of $(H, \dot{H}, \ddot{H})$. If we take the third derivative of (4.26) and substitute these expressions for the constants, the resulting equation can be expressed using only the cosmographic parameters:

$$s + 2(q + j) + qj = 0\,. \tag{4.29}$$

4.2.1 A model-independent dynamical system

Rewriting the first term on the right-hand side of (4.22) in terms of these parameters gives $\ddot{H}/H^3 = 2 + 3q + j$. $xy\Gamma$ can therefore be written as

$$xy\Gamma = -2(y + 1) + 6(x + z - \Omega) + 2 + 3q + j\,. \tag{4.30}$$

The definitions of y (4.14) and the Ricci scalar (4.15) provide another constraint

$$\begin{aligned} \frac{\dot{H}}{H^2} &= \frac{R}{6H^2} - 2 - K \\ \therefore y &= \frac{\dot{H}}{H^2} + 2 + K \\ &= 1 - q + K\,. \end{aligned} \tag{4.31}$$

This allows the Friedmann constraint to be rewritten as

$$z = -x + \Omega - q\,. \tag{4.32}$$

Using the constraint equations (4.31) and (4.32) allows y and z to be eliminated from the dynamical system (4.19).

$$\frac{\mathrm{d}x}{\mathrm{d}\eta} = -x(x - q) + 2(x + K + q) - 3\Omega(1 + w) + 2\,, \tag{4.33a}$$

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\eta} = -\Omega(x - 2q + 3w + 1)\,, \tag{4.33b}$$

$$\frac{\mathrm{d}K}{\mathrm{d}\eta} = 2Kq\,, \tag{4.33c}$$

$$\frac{\mathrm{d}q}{\mathrm{d}\eta} = 2q^2 + q - j\,, \tag{4.33d}$$

$$\frac{\mathrm{d}j}{\mathrm{d}\eta} = j(3q + 2) + s\,. \tag{4.33e}$$

It will also prove beneficial to use the constraint equations to eliminate y and z from (4.30):

$$xy\Gamma = -2K - q + j - 2\,. \tag{4.34}$$

The dynamical system (4.33) takes on a simpler form than previously (4.19), and does not require one to specify the functional form of $f(R)$ to make the system autonomous. The last two equations are completely kinematical in nature, having no dependence on the dimensionless variables. They depend only on how the Universe evolves rather than the inherent dynamical properties responsible for its evolution. In the context of the present universe, $q < 0$ due to acceleration. This means that $|K| \leq 1$, i.e. $1 + K \geq 0$, leading to the conclusion that $y > 0$ in the present universe. $y = 0$ is therefore an invariant submanifold of the dynamical system, dividing the phase space into two disjoint regions, $y > 0$ ($R > 0$) and $y < 0$ ($R < 0$), the former of which contains the relevant physical dynamics leading to an accelerated universe. Additionally, the presence of an invariant submanifold implies that there can exist no global attractor or repeller anywhere other than $y = 0$.

As mentioned previously, the physical viability of any $f(R)$ theory requires that $F > 0$ throughout the physically relevant region of the phase space, and that $F' \geq 0$ at least within the locality of the fixed point corresponding to the matter-dominated epoch [36]. One can obtain a term of the form F'/F by dividing x by $xy\Gamma$, giving

$$\frac{x}{-2K - q + j - 2} = \frac{6H^2F'}{F}\,. \tag{4.35}$$

Assuming $F > 0$ and constraining $F' \geq 0$ yields the necessary constraint

$$\frac{x}{-2K - q + j - 2} \geq 0\,. \tag{4.36}$$

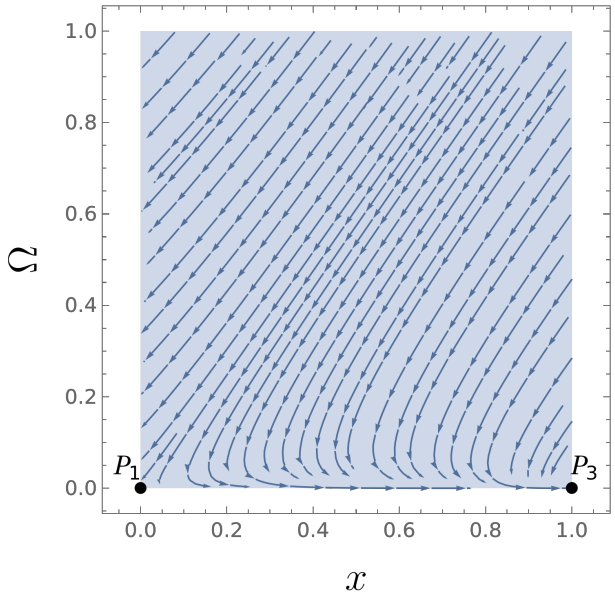
Table 4.1 shows the fixed points of the system. P_1 and P_3 are a saddle point and attractor, respectively, which give solutions for exponential expansion driven by the scalar field. P_2 is a saddle point where expansion is driven by matter, and the scale factor evolves as a power of time — this is useful for studying the matter-dominated era. P_4 and P_5 are a repeller and a saddle, respectively, where the dynamics are primarily influenced by the scalaron field and the scale factor evolves according to a power-law relation. This type of solution could describe various phases of cosmic evolution, including both early inflation and late-time acceleration, depending on the specific form of the $f(R)$ function and the exponent. P_6 and P_7 are fixed points that do not reside on the invariant submanifold $K = 0$, giving an unphysical Milne solution.

Fixed point	Coordinates $(x^*, \Omega^*, K^*, q^*)$	Stability	Cosmological solution
P_1	$(0, 0, 0, -1)$	Saddle	Scalaron-dominated De-Sitter
P_2	$(0, 1, 0, \frac{1}{2})$	Saddle	Matter-dominated power law
P_3	$(1, 0, 0, -1)$	Attractor	Scalaron-dominated De-Sitter
P_4	$(\frac{5-\sqrt{73}}{4}, 0, 0, \frac{1}{2})$	Repeller	Scalaron-dominated power law
P_5	$(\frac{5+\sqrt{73}}{4}, 0, 0, \frac{1}{2})$	Saddle	Scalaron-dominated power law
P_6	$(0, 0, -1, 0)$	Saddle	Milne solution
P_7	$(2, 0, -1, 0)$	Saddle	Milne solution

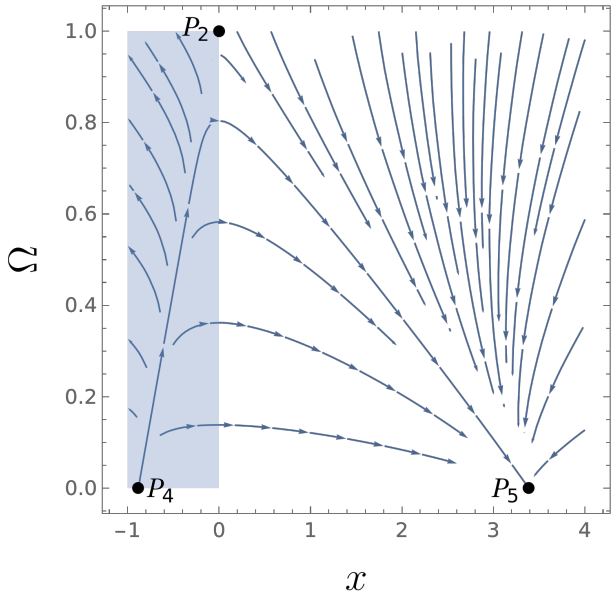
Table 4.1: Fixed points of the dynamical system (4.33e) with their respective coordinates, stabilities and cosmological solutions. P_3 and P_5 are fixed points that do not satisfy constraint (4.36) for $j = 1$, due to division by zero and a negative value, respectively. P_6 and P_7 are fixed points that yield an unphysical solution.

Figure 4.1 presents phase portrait projections on the slices $q = -1$ (4.1a), $q = 1/2$ (4.1b), $\Omega = 0$ (4.1c), and $\Omega = 1$ (4.1d). The shaded region represents the area where the condition $F' > 0$ is satisfied, given that $F > 0$. The faster a phase trajectory exits this shaded region, the more likely the corresponding cosmic evolution is to encounter the Dolgov-Kawasaki instability, making it less physically viable. Trajectories originating from P_3 and moving toward the acceleration submanifold at $q = -1$ indicate possible evolutionary paths from a decelerating, matter-dominated phase to a late-time accelerating phase. Thus, the trajectories on the left of (4.1b) are the most viable. Meanwhile, all trajectories passing near the saddle point P_2 eventually reach P_5, which lies outside the viable region, making these paths undesirable [4].

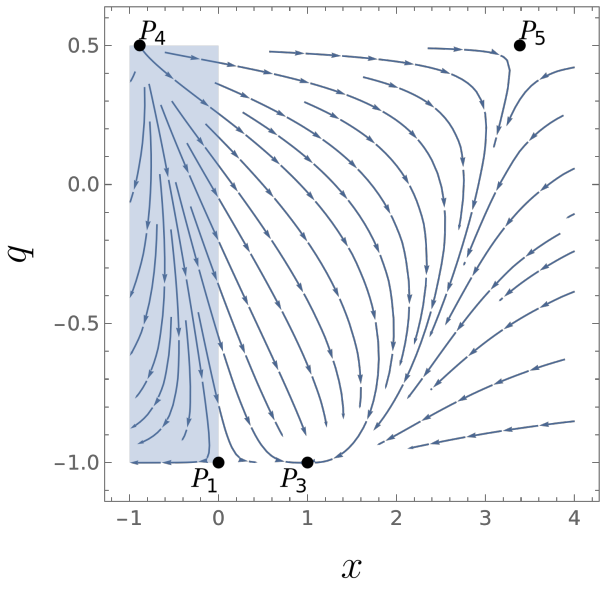
[4] All plots in this chapter were generated using Mathematica v14 (Wolfram Research).



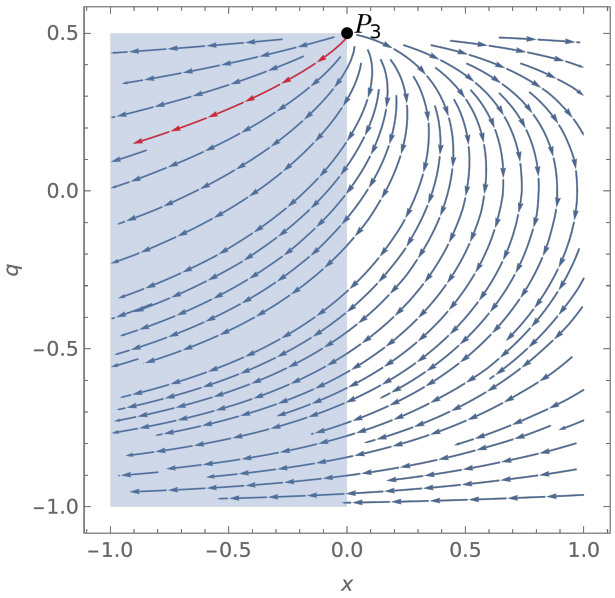
(a) $q = -1$



(b) $q = \frac{1}{2}$



(c) $\Omega = 0$



(d) $\Omega = 1$

Figure 4.1: Phase diagrams showing projections on the slices (a) $q = -1$, (b) $q = 1/2$, (c) $\Omega = 0$, and (d) $\Omega = 1$. The shaded region represents the region in which the condition $F', F > 0$ is satisfied. The red line in (d) indicates a viable trajectory, which has evolved from the initial conditions in (4.134).

4.3 Structure growth in $f(R)$

This section examines the treatment of perturbations in $f(R)$ gravity, starting from the propagation equations that govern the evolution of scalar, vector, and tensor modes. This allows us to introduce the covariant gauge-invariant formalism and construct a model-independent dynamical system.

4.3.1 Propagation equations

For a general tensor T_{ab} we can write

$$T_{ab} = \rho u_a u_b + p h_{ab} + 2q_{(a}u_{b)} + \pi_{ab}\,, \tag{4.37}$$

where ρ and p are the energy density and isotropic pressure, respectively, $q_a = q_{\langle a\rangle}$ is the energy flux and $\pi_{ab} = \pi_{\langle ab\rangle}$ is the anisotropic pressure. Applying this decomposition to the effective energy-momentum tensor gives us

$$\begin{aligned}
\rho^{\rm tot} &= T^{\rm tot}_{ab}u^a u^b = \tilde{\rho}^m + \rho^R\,,\\
p^{\rm tot} &= \frac{1}{3}T^{\rm tot}_{ab}a^{ab} = \tilde{p}^m + p^R\,,\\
q^{\rm tot}_a &= -T^{\rm tot}_{bc}h^b_a u^c = \tilde{q}^m_a + q^R_a\,,\\
\pi^{\rm tot}_{ab} &= T^{\rm tot}_{cd}h^c_{<a}h^d_{b>} = \tilde{\pi}^m_{ab} + \pi^R_{ab}\,,
\end{aligned}$$

relative to u^m_a, where

$$\begin{aligned}
\tilde{\rho}^m &= \frac{\rho^m}{f'}\,, \quad \tilde{p}^m = \frac{p^m}{f'}\,,\\
\tilde{q}^m_a &= \frac{q^m_a}{f'}\,, \quad \tilde{\pi}^m_{ab} = \frac{\pi^m_{ab}}{f'}\,.
\end{aligned}$$

If we assume standard matter to be a perfect fluid, q^m_a and π^m_{ab} – and consequently $\tilde{q}^m_a$ and $\tilde{\pi}^m_{ab}$ – vanish. The effective thermodynamical quantities for the curvature fluid are

$$\rho^R = \frac{1}{f'}\left[\frac{1}{2}\left(Rf'-f\right) - \Theta f''\dot{R} + f''\widetilde{\nabla}^2 R + f''\dot{u}_b\widetilde{\nabla}^b R\right]\,, \tag{4.38}$$

$$\begin{aligned}
p^R = \frac{1}{f'}\Bigg[&\frac{1}{2}\left(f - Rf'\right) + f''\ddot{R} + f'''\dot{R}^2 + \frac{2}{3}\Theta f''\dot{R}\\
&-\frac{2}{3}f''\widetilde{\nabla}^2 R - \frac{2}{3}f'''\widetilde{\nabla}^a R\widetilde{\nabla}_a R - \frac{1}{3}f''\dot{u}_b\widetilde{\nabla}^b R\Bigg]\,,
\end{aligned} \tag{4.39}$$

$$q^R_a = -\frac{1}{f'}\left[f'''\dot{R}\widetilde{\nabla}_a R + f''\widetilde{\nabla}_a\dot{R} - \frac{1}{3}f''\widetilde{\nabla}_a R\right]\,, \tag{4.40}$$

$$\pi^R_{ab} = \frac{1}{f'}\left[f''\widetilde{\nabla}_{\langle a}\widetilde{\nabla}_{b\rangle}R + f'''\widetilde{\nabla}_{\langle a}R\widetilde{\nabla}_{b\rangle}R - \sigma_{ab}\dot{R}\right]\,. \tag{4.41}$$

The twice-contracted Bianchi identities for $f(R)$

$$\nabla^b\widetilde{T}^M_{ab} = \frac{\nabla^b T^m_{ab}}{f'} - \frac{f''}{f'^2}T^m_{ab}\nabla^b R\,, \tag{4.42}$$

$$\nabla^b T^R_{ab} = \frac{f''}{f'^2}\widetilde{T}^M_{ab}\nabla^b R\,, \tag{4.43}$$

yield differential equations for the quantities ρ^m, ρ^R, q_a^R:

$$\dot{\rho}^m = -\Theta\left(\rho^m + p^m\right) , \tag{4.44}$$

$$\dot{\rho}^R + \widetilde{\nabla}^a q_a^R = -\Theta\left(\rho^R + p^R\right) - 2\left(\dot{u}^a q_a^R\right) - \left(\sigma^{ab}\pi_{ba}^R\right) + \rho^m \frac{f''\dot{R}}{f^2} , \tag{4.45}$$

$$\dot{q}_{\langle a\rangle}^R + \widetilde{\nabla}_a p^R + \widetilde{\nabla}^b \pi_{ab}^R = -\frac{4}{3}\Theta q_a^R - \sigma_a^b q_b^R - \left(\rho^R + p^R\right)\dot{u}_a - \dot{u}^b \pi_{ab}^R - \eta_a^{bc}\omega_b q_c^R + \rho^m \frac{f''\widetilde{\nabla}_a R}{f^R} . \tag{4.46}$$

A relation between the acceleration $\dot{u}_a$ and the density and pressure terms then follows from the momentum conservation equation (3.36):

$$\widetilde{\nabla}^a p^m = -\left(\rho^m + p^m\right)\dot{u}^a . \tag{4.47}$$

These equations (4.44)–(4.47) comprise the linear conservation equations for matter and curvature. Together with the nonlinear $1+3$ equations for fourth-order gravity (found in Appendix A), these govern the dynamics of matter and gravitaional fields in fourth order gravity.

We choose our background spacetime to be that of FLRW, in which – as discussed in §3.2 – first-order and gauge-invariant variables vanish. The Friedmann background is devoid of all inhomogeneous and anisotropic quantities q_a^R and π_{ab}^R, which establishes the order of terms in the $1+3$ equations and the linearisation process. Specifically, quantities that vanish in the background are treated as first order in the linearisation scheme. When we linearise the background equations, we drop all terms of second order and higher. The Stuart and Walker lemma guarantees that because these quantities are zero in the background, they are inherently gauge-invariant [154]. The cosmological equations for the background are

$$\Theta^2 = 3\tilde{\rho}^m + 3\rho^R - \frac{\widetilde{R}}{2} , \tag{4.48}$$

$$\dot{\Theta} + \frac{1}{3}\Theta^2 + \frac{1}{2}\left(\tilde{\rho}^m + 3\tilde{p}^m\right) + \frac{1}{2}\left(\rho^R + 3p^R\right) = 0 , \tag{4.49}$$

$$\dot{\rho}^m + \Theta\left(\rho^m + p^m\right) = 0 , \tag{4.50}$$

where ρ^R and p^R are the zeroth order energy density and pressure of the curvature fluid and $\widetilde{R}$ is the 3-Ricci scalar defined in (3.42). Linearisation of the exact propagation (4.51)–(4.55) and

constraint equations (4.56)–(4.60) about the FLRW background yields the system [187]

$$\dot{\Theta} + \frac{1}{3}\Theta^2 - \widetilde{\nabla}^a A_a + \frac{1}{2}\left(\tilde{\rho}^m + 3\tilde{p}^m\right) = -\frac{1}{2}\left(\rho^R + 3p^R\right) , \tag{4.51}$$

$$\dot{\omega}_a + 2H\omega_a + \frac{1}{2}\operatorname{curl} A_a = 0 , \tag{4.52}$$

$$\dot{\sigma}_{ab} + 2H\sigma_{ab} + E_{ab} - \widetilde{\nabla}_{\langle a} A_{b\rangle} = -q^R_a , \tag{4.53}$$

$$\begin{aligned}\dot{E}_{ab} + 3HE_{ab} - \operatorname{curl} H_{ab} + \frac{1}{2}\left(\tilde{\rho}^m + \tilde{p}^m\right)\sigma_{ab} \\ = -\frac{1}{2}\left(\rho^R + p^R\right)\sigma_{ab} - \frac{1}{2}\dot{\pi}^R_{\langle ab\rangle} - \frac{1}{2}\widetilde{\nabla}_{\langle a} q^R_{b\rangle} - \frac{1}{6}\Theta\pi^R_{ab} .\end{aligned} \tag{4.54}$$

$$\dot{H}_{ab} + 3HH_{ab} + \operatorname{curl} E_{ab} = \frac{1}{2}\operatorname{curl}\pi^R_{ab} , \tag{4.55}$$

$$\widetilde{\nabla}^b\sigma_{ab} - \operatorname{curl}\omega_a - \frac{2}{3}\widetilde{\nabla}_a\Theta = -q^R_a , \tag{4.56}$$

$$\operatorname{curl}\sigma_{ab} + \widetilde{\nabla}_{\langle a}\omega_{b\rangle} - H_{ab} = 0 , \tag{4.57}$$

$$\widetilde{\nabla}^b E_{ab} - \frac{1}{3}\widetilde{\nabla}_a\tilde{\rho}^m = -\frac{1}{2}\widetilde{\nabla}^b\pi^R_{ab} + \frac{1}{3}\widetilde{\nabla}_a\rho^R - \frac{1}{3}\Theta q^R_a , \tag{4.58}$$

$$\widetilde{\nabla}^b H_{ab} - (\tilde{\rho}^m + \tilde{p}^m)\omega_a = -\frac{1}{2}\operatorname{curl} q^R_a + (\rho^R + p^R)\omega_a , \tag{4.59}$$

$$\widetilde{\nabla}^a\omega_a = 0 , \tag{4.60}$$

and the linearised conservation equations, obtained from (4.44)–(4.47), are

$$\dot{\rho}^m = -\Theta(\rho^m + p^m) , \tag{4.61}$$

$$\widetilde{\nabla}^a p^m = -(\rho^m + p^m)\dot{u}^a , \tag{4.62}$$

$$\dot{\rho}^R + \widetilde{\nabla}^a q^R_a = -\Theta(\rho^R + p^R) + \rho^m\frac{f''\dot{R}}{f'^2} , \tag{4.63}$$

$$\dot{q}^R_{\langle a\rangle} + \widetilde{\nabla}_a p^R + \widetilde{\nabla}^b\pi^R_{ab} = -\frac{4}{3}\Theta q^R_a - (\rho^R + p^R)\dot{u}_a + \rho^m\frac{f''\widetilde{\nabla}_a R}{f^2} . \tag{4.64}$$

4.3.2 CGI density perturbations

We can now analyse the growth of large-scale structure using the CGI approach introduced by Ellis and Bruni [109] and summarised in §3.2. To derive the perturbation equations in the general case, we define the density and expansion gradients in the same way as [109, 187, 188]:

$$D^m_a = \frac{a}{\rho^m}\widetilde{\nabla}_a\rho^m , \quad Z_a = a\widetilde{\nabla}_a\Theta , \quad C_a = a\widetilde{\nabla}_a\widetilde{R} , \tag{4.65}$$

and the dimensional gradients that describe the inhomogeneities in the Ricci scalar R and its momentum $\dot{R}$ as

$$\mathcal{R}_a = a\widetilde{\nabla}_a R , \quad \mathfrak{R}_a = a\widetilde{\nabla}_a\dot{R} . \tag{4.66}$$

Using equations (4.38), (4.39) and the identities in Appendix B, we can obtain the evolution equations for the variables in (4.65) (barring C_a, the derivative of which is redundant in this case). The simplest is $\dot{D}^m_a$, which becomes

$$\dot{D}^m_a = w\Theta D^m_a - (1+w)Z_a , \tag{4.67}$$

while $\dot{Z}_a$ is more complicated:

$$\begin{aligned}\dot{Z}_a = &-\frac{2}{3}\Theta Z_a + \frac{a\rho^m f''}{2f'^2}(1+3w)\widetilde{\nabla}_a R - \frac{a}{2f'}\widetilde{\nabla}_a\left[\rho^m(1+3w)\right] \\ &- \frac{a}{2}\widetilde{\nabla}_a\left(\rho^R + 3p^R\right) + a\widetilde{\nabla}_a\widetilde{\nabla}^a\dot{u}_a + a\dot{\Theta}\dot{u}_a\,. \end{aligned} \tag{4.68}$$

To continue, we will need to simplify the $\widetilde{\nabla}_a\rho^R$ and $\widetilde{\nabla}_a p^R$ terms. This will require the use of the following derivatives:

$$\widetilde{\nabla}_a f' = f''\widetilde{\nabla}_a R, \quad \widetilde{\nabla}_a f'' = f'''\widetilde{\nabla}_a R, \quad \widetilde{\nabla}_a f''' = f^{(4)}\widetilde{\nabla}_a R, \quad \widetilde{\nabla}_a\frac{1}{f'} = -\frac{f''}{f'^2}\widetilde{\nabla}_a R\,. \tag{4.69}$$

The pressure-density term in (4.68) can then be expanded to

$$\begin{aligned}\frac{1}{2}\left(a\widetilde{\nabla}_a\rho^R + 3a\widetilde{\nabla}_a p^R\right) = &-\frac{f''}{2f'}\left(\rho^R + 3p^R\right)\mathcal{R}_a - \frac{f''}{2f'}R\mathcal{R}_a + \frac{f''}{2f'}\dot{R}Z_a + \frac{f''}{2f'}\Theta\mathfrak{R} + \frac{f'''}{2f'}\dot{R}\Theta\mathcal{R}_a \\ &- \frac{f'''}{2f'}\widetilde{\nabla}^2 R\mathcal{R}_a - \frac{f''}{2f'}a\widetilde{\nabla}_a\left(\widetilde{\nabla}^2 R\right) + \frac{3f''}{2f'}a\widetilde{\nabla}_a\ddot{R} + \frac{3f'''}{2f'}\ddot{R}\mathcal{R}_a + \frac{3f'''}{f'}\dot{R}\mathfrak{R} \\ &+ \frac{3f^{(4)}}{2f'}\dot{R}^2\mathcal{R}_a - \frac{f'''}{f'}a\widetilde{\nabla}_a\left(\widetilde{\nabla}^b R\widetilde{\nabla}_b R\right) - \frac{f^{(4)}}{f'}\left(\widetilde{\nabla}^b R\widetilde{\nabla}_b R\right)\mathcal{R}_a\,, \end{aligned}$$

within which are are terms in need of further simplification, namely $\widetilde{\nabla}_a(\dot{u}_b\widetilde{\nabla}^b R)$ and $\widetilde{\nabla}_a\ddot{R}$, to ascertain whether they contain second-order terms. The former is fairly straightforward:

$$\begin{aligned}\dot{u}_b &= -\frac{\widetilde{\nabla}_b p^m}{\rho^m + p^m} = -\frac{w\widetilde{\nabla}_b\rho^m}{\rho^m(1+w)} \\ \therefore \widetilde{\nabla}_a\left(\dot{u}_b\widetilde{\nabla}^b R\right) &= -\frac{w}{1+w}\widetilde{\nabla}_a\left(\frac{\widetilde{\nabla}_b\rho^m}{\rho^m}\widetilde{\nabla}^b R\right) = -\frac{w}{1+w}\widetilde{\nabla}_a\left(\frac{D_b^m\mathcal{R}^b}{a^2}\right)\,, \end{aligned}$$

which is second order due to the product of D_b^m and $\mathcal{R}^b$. The latter necessitates the use of the trace equation (4.5), which we multiply by the metric tensor to isolate R:

$$\begin{aligned}g^{ab}R_{ab} - \frac{1}{2}g^{ab}g_{ab}R &= g^{ab}\frac{1}{f'}T^m_{ab} + g^{ab}T^R_{ab} \\ \therefore R - 2R &= \frac{1}{f'}T^m + T^R\,. \end{aligned}$$

The terms on the right-hand side can be found by multiplying the energy-momentum equation by the metric tensor,

$$\begin{aligned}g^{ab}T^m_{ab} &= g^{ab}\rho^m u_a u_b + g^{ab}p^m h_{ab} + 2g^{ab}q^m{}_{(a}u_{b)} + \pi^m_{ab} \\ \therefore T^m &= \rho^m u_a u^a + g^{ab}h_{ab}p^m \\ &= -\rho^m + 3p^m\,, \end{aligned}$$

where q^m_a and π^m_{ab} vanish due to the assumption of standard matter as a perfect fluid. The result above also applies to T^R, as q^R_a and π^R_{ab} are inhomogeneous and anisotropic quantities that consequently vanish on a Friedmann background. This allows R to be written as

$$\begin{aligned}R &= \frac{1}{f'}\left(\rho^m - 3p^m\right) + \left(\rho^R - 3p^R\right) \\ &= \frac{\rho^m}{f'}(1-3w) + \frac{1}{f'}\left(2Rf' - 2f - 3\Theta f''\dot{R} + 3f''\widetilde{\nabla}^2 R + 2f''\dot{u}_b\widetilde{\nabla}^b R - 3f''\ddot{R} - 3f'''\dot{R}^2\right)\,, \end{aligned}$$

where we discarded the last term ($2f'''\widetilde{\nabla}^a R\widetilde{\nabla}_a R$) due to being second order. Now we make $3f''\ddot{R}$ the subject of the equation and take the projected covariant derivative, giving

$$3f''\widetilde{\nabla}_a\ddot{R} = \left(-3\ddot{R}f''' + Rf'' - f' - 3\Theta\dot{R}f''' + 3f'''\widetilde{\nabla}^2 R + 2f'''\dot{u}_b\widetilde{\nabla}^b R - 3f^{(4)}\dot{R}^2 + 6f''\frac{K}{a^2}\right)\frac{\mathcal{R}_a}{a}$$
$$+\rho^m(1-3w)\frac{D_a^m}{a} - 3f''\dot{R}\frac{Z_a}{a} + \left(-3\Theta f'' - 6f'''\dot{R}\right)\frac{\mathfrak{R}_a}{a} + 3f''\widetilde{\nabla}^2\left(\frac{\mathcal{R}_a}{a}\right) + 2f''\widetilde{\nabla}_a\left(\dot{u}_b\widetilde{\nabla}^b R\right), \tag{4.70}$$

where we used the identity (B.2) to rewrite the $a\widetilde{\nabla}_a\widetilde{\nabla}^2 R$ term. We can now write $\dot{Z}_a$ as

$$\dot{Z}_a = \left(\frac{f''\dot{R}}{f'} - \frac{2\Theta}{3}\right)Z_a + \left(\frac{(w-1)\rho^m}{f'(1+w)} - \frac{w(f+2\Theta f''\dot{R})}{f'(w+1)} + \frac{wR}{2(w+1)} - \frac{w}{(w+1)}\frac{3K}{a^2}\right)D_a^m$$
$$+\left(\frac{f''\Theta}{f'}\right)\mathfrak{R}_a - \frac{w}{w+1}\widetilde{\nabla}^2 D_a^m + \left(-\frac{f''}{f'}\right)\widetilde{\nabla}^2\mathcal{R}_a$$
$$+\left(\frac{1}{2} - \frac{ff''}{2f'^2} - \left(\frac{f''}{f'}\right)^2\Theta\dot{R} + \frac{f''\rho^m}{f'^2} - \frac{2f''}{f'}\frac{K}{a^2} + \frac{\Theta\dot{R}f'''}{f'}\right)\mathcal{R}_a\,. \tag{4.71}$$

The Ricci scalar and momentum in (4.66) have the following evolution equations:

$$\dot{\mathcal{R}}_a = \mathfrak{R}_a - \frac{w}{(w+1)}\dot{R}D_a^m\,, \tag{4.72}$$

$$\dot{\mathfrak{R}}_a = \left(-\frac{\ddot{R}f'''}{f''} + \frac{R}{3} - \frac{f'}{3f''} - \frac{\Theta\dot{R}f'''}{f''} - \frac{f^{(4)}\dot{R}^2}{f''} + 2\frac{K}{a^2}\right)\mathcal{R}_a$$
$$+\left(\frac{\rho^m(1-3w)}{3f''} - \frac{w}{(1+w)}\ddot{R}\right)D_a^m - \dot{R}Z_a + \left(-\Theta - \frac{2f'''}{f''}\dot{R}\right)\mathfrak{R}_a + \widetilde{\nabla}^2\mathcal{R}_a\,. \tag{4.73}$$

The equations (4.67) and (4.71)–(4.73) describe the general evolution of density fluctuations and other scalar quantities on the FLRW background. However, the clustering of matter is generally described by only the scalar components of these variables. Since we are focusing on density perturbations, we need only extract the scalar components of the density gradient. To do this, we can use the methods introduced in §3.2.2. We define the scalar quantities

$$\Delta_m = \frac{a^2}{\rho_m}\widetilde{\nabla}^2\rho_m\,,\quad Z = 3a^2\widetilde{\nabla}^2 H\,,\quad C = a^2\widetilde{\nabla}^2\widetilde{R}\,,$$
$$\mathcal{R} = a^2\widetilde{\nabla}^2 R\,,\quad \mathfrak{R} = a^2\widetilde{\nabla}^2\dot{R}\,, \tag{4.74}$$

where $\widetilde{\nabla}^2 \equiv \widetilde{\nabla}_a\widetilde{\nabla}^a$ and $\Delta_m = a\widetilde{\nabla}^a\mathcal{D}_a^m$. Equations (4.67) and (4.71)–(4.73) therefore become

$$\dot{\Delta}_m = w\Theta\Delta - (1+w)Z\,, \tag{4.75}$$

$$\dot{Z} = \left(\frac{f''\dot{R}}{f'} - \frac{2\Theta}{3}\right)Z + \left(\frac{(w-1)\rho^m}{f'(1+w)} - \frac{w(f+2\Theta f''\dot{R})}{f'(w+1)} + \frac{wR}{2(w+1)} - \frac{w}{(w+1)}\frac{5K}{S^2}\right)\Delta^m$$
$$+\left(\frac{f''\Theta}{f'}\right)\mathfrak{R} - \frac{w}{w+1}\widetilde{\nabla}^2\Delta^m + \left(-\frac{f''}{f'}\right)\left(\widetilde{\nabla}^2\mathcal{R}\right)$$
$$+\left(\frac{1}{2} - \frac{ff''}{2f'^2} - \left(\frac{f''}{f'}\right)^2\Theta\dot{R} + \frac{f''\rho^m}{f'^2} - \frac{4f''}{f'}\frac{K}{S^2} + \frac{\Theta\dot{R}f'''}{f'}\right)\mathcal{R}\,, \tag{4.76}$$

$$\dot{\mathcal{R}} = \mathfrak{R} - \frac{w}{(w+1)}\dot{R}\Delta^m\,, \tag{4.77}$$

$$\dot{\mathfrak{R}} = \left(-\frac{\ddot{R}f'''}{f''} + \frac{R}{3} - \frac{f'}{3f''} - \frac{\Theta\dot{R}f'''}{f''} - \frac{f^{(4)}\dot{R}^2}{f''} + 4\frac{K}{S^2}\right)\mathcal{R} - \dot{R}Z + \left(-\Theta - \frac{2f'''}{f''}\dot{R}\right)\mathfrak{R} + \widetilde{\nabla}^2\mathcal{R} + \left(\frac{\rho^m(1-3w)}{3f''} - \frac{w}{(1+w)}\ddot{R}\right)\Delta^m\,. \tag{4.78}$$

We now perform the harmonic decomposition introduced in (3.56), using the eigenfunctions of the Laplace-Beltrami operator defined in Ellis, Bruni, and Dunsby [109, 189]. This allows us to expand all first-order quantities as

$$X(t,\mathbf{x}) = \sum X^{(k)}(t)Q^{(k)}(\mathbf{x})\,, \tag{4.79}$$

where $\sum$ is the summation over a discrete index or integration over a continuous one. We can derive a pair of coupled second-order equations that describes the evolution of the k-th mode density perturbations [187].

$$\begin{aligned}\ddot{\Delta}^k_m &- \left((3w-2)\,H + \frac{f''\dot{R}}{f'}\right)\dot{\Delta}^k_m + \left(w\frac{k^2}{a^2} + \frac{(w-1)\rho_m}{f'} - w\frac{f}{f'}\right)\Delta^k_m \\ &= \left(-\frac{3H(1+w)f''}{f'}\right)\dot{\mathcal{R}}^k \\ &+ \left(\frac{(1+w)}{2}\left[-1 + \frac{ff''}{f'^2} - \frac{2\rho_m f''}{f'^2} - \frac{2f''}{f'}\frac{k^2}{a^2} + 6H\dot{R}\left(\frac{f''}{f'}\right)^2 - \frac{6Hf'''\dot{R}}{f'}\right]\right)\mathcal{R}^k\,,\end{aligned} \tag{4.80a}$$

$$\begin{aligned}\ddot{\mathcal{R}}^k &+ \left(3H + \frac{2f'''\dot{R}}{f''}\right)\dot{\mathcal{R}}^k + \left(-\frac{R}{3} + \frac{k^2}{a^2} + \frac{f'}{3f''} + \frac{3Hf'''\dot{R}}{f''} + \frac{f^{(4)}\dot{R}^2}{f''} + \frac{f'''\ddot{R}}{f''}\right)\mathcal{R}^k \\ &= \left(\frac{(1-w)\dot{R}}{(1+w)}\right)\dot{\Delta}^k_m \\ &- \left(\frac{(3w-1)\rho_m}{3f''} + \frac{2w}{w+1}\left[\ddot{R} + \dot{R}\left(3H + \frac{f'''\dot{R}}{f''}\right)\right]\right)\Delta^k_m\,.\end{aligned} \tag{4.80b}$$

To simplify this system, we can use the *quasi-static (QS) approximation* employed in [190, 191, 192]. There are actually two assumptions present here: the first is the QS approximation

$$|\dot{X}| \lesssim H|X|\,; \tag{4.81}$$

the second is the sub-horizon or *sub-Hubble approximation*

$$|\nabla^2 X| \gg a^2H^2|X|\,, \tag{4.82}$$

where $X = \mathcal{R}^k, \dot{\mathcal{R}}^k$. The former is the assumption that time derivatives of perturbations are suppressed compared to their spatial derivatives, while the latter allows us to take the limit $k \gg aH$; which is necessary when ignoring time derivatives. In $f(R)$ gravity this translates to

$k^2/(aH)^2 \gg f^{(n)}/f'$ for $n = 2, 3, 4$. The QS approximation applied to Δ_m gives

$$\ddot{\Delta} + \left[\frac{3(1-w)f''\dot{R}\left(-f''\left(6Hf''\dot{R} + f - 2\rho_m\right) + f'^2 + 2\frac{k^2}{a^2}f'f''\right)}{2f'^2\left(f''\left(3\frac{k^2}{a^2} - R\right) + f' + 3\left(f^{(4)}\dot{R}^2 + 3f^{(3)}H\dot{R} + f^{(3)}\ddot{R}\right)\right)} - \frac{f''\dot{R}}{f'} + H(2-3w)\right]\dot{\Delta}$$
$$+\left[\frac{\left(6w\left(f^{(3)}\dot{R}^2 + 3Hf''\dot{R} + f''\ddot{R}\right) + \rho_m\left(3w^2 + 2w - 1\right)\right)\left(f''\left(6Hf''\dot{R} + f - 2\rho_m\right) - f'^2 - 2\frac{k^2}{a^2}f'f''\right)}{2f'^2\left(f''\left(3\frac{k^2}{a^2} - R\right) + f' + 3\left(f^{(4)}\dot{R}^2 + 3f^{(3)}H\dot{R} + f^{(3)}\ddot{R}\right)\right)}\right.$$
$$\left. + \frac{k^2 w}{a^2} - \frac{wf}{f'} + \frac{\rho_m(w-1)}{f'}\right]\Delta = 0\,, \tag{4.83}$$

which we will refer to as the *semi-quasi-static approximation*. Further simplifying the system with the sub-Hubble assumption leads to

$$\ddot{\Delta}^k_m + (2-3w)H\dot{\Delta}^k_m = \left[w\left(\frac{f}{f'} - \frac{k^2}{a^2}\right) + \frac{\rho_m}{2f'}\left(\frac{(3w^2+1) + 2(3w^2 - w + 2)\frac{f''}{f'}\frac{k^2}{a^2}}{1 + 3\frac{f''}{f'}\frac{k^2}{a^2}}\right)\right]\Delta^k_m\,, \tag{4.84}$$

which we will refer to as the *full QS approximation*. For the dust-only case ($w = 0$), this reduces to

$$\ddot{\Delta} + (2-3w)H\dot{\Delta} - \frac{\rho_m}{2f'}\left(\frac{1 + 4\frac{f''}{f'}\frac{k^2}{a^2}}{1 + 3\frac{f''}{f'}\frac{k^2}{a^2}}\right)\Delta = 0\,. \tag{4.85}$$

The accuracy of this approximation in the synchronous and conformal gauges has been thoroughly examined in [193, 191, 194]. It has been shown to perform well in certain model-dependent, sub-horizon scenarios, particularly for models where the background evolution closely mimics ΛCDM , as demonstrated by De la Cruz-Dombriz et al. (2008) [195]. However, concerns persist regarding whether the omission of both higher-order derivatives of f and time derivatives of the curvature perturbation is an oversimplification, as highlighted by Bean et al. (2007) [196], especially in models that significantly diverge from ΛCDM. To address this issue, it is advisable to apply the quasi-static method both with and without higher-order derivatives of f. All these approaches retain an explicit dependence on $f(R)$, meaning they can only be employed once a specific model is chosen. Moreover, in order to establish a connection between perturbation studies and results obtained from dynamical systems analyses, it is necessary to express all coefficients in terms of the selected dynamical systems variables. In the case of the full covariant system, this requires derivatives of $f(R)$ up to fourth order, further constraining the range of $f(R)$ models that can be explored using this framework.

4.3.3 Model-independent system construction

The model-dependence of (4.80) can be avoided by employing the cosmographic approach outlined in §4.2.1. Time derivatives of the Hubble parameter (2.54) and the Ricci scalar (1.52) can be expressed in terms of the cosmographic parameters (2.59):

$$\frac{\dot{H}}{H^2} = -q - 1\,, \tag{4.86a}$$
$$\frac{\ddot{H}}{H^3} = 3q + j + 2\,, \tag{4.86b}$$
$$\frac{\dddot{H}}{H^4} = -3q^2 - 12q - 4j + s - 6\,, \tag{4.86c}$$
$$\frac{\ddddot{H}}{H^5} = 30q^2 + 60q + 10qj + 20j - 5s + l + 24\,; \tag{4.86d}$$

$$R = 6H^2\left(\frac{\dot{H}}{H^2} + 2\right) = 6H^2(1-q)\,, \tag{4.87a}$$

$$\dot{R} = 6H^3\left(\frac{\ddot{H}}{H^3} + 4\frac{\dot{H}}{H^2}\right) = 6H^3\left(j - q - 2\right)\,, \tag{4.87b}$$

$$\ddot{R} = 6H^4\left(s + q^2 + 8q + 6\right)\,, \tag{4.87c}$$

$$\dddot{R} = 6H^5\left(l - s - 2qj - 8j - 18q^2 - 48q - 24\right)\,. \tag{4.87d}$$

Taking higher-order time derivatives of the Friedmann equation (4.8) yields higher-order R derivatives, for which we can substitute the cosmographic parameters. f can then be expressed exclusively in terms of the dynamical systems variables and R derivatives, and thus the cosmographic parameters using (4.87). For example, by taking the first time derivative of (4.8) we can express f''' as

$$\begin{aligned} f''' &= \frac{f'H^2}{\dot{R}^2}\left(x - 3(1+w)\Omega + 2(q+1) - \frac{\ddot{R}x}{H\dot{R}}\right), \\ &= \frac{f'H^2}{\dot{R}^2}\left(x - 3(1+w)\Omega + 2(q+1) - \frac{(s+q^2+8q+6)x}{(j-q-2)}\right), \end{aligned}$$

and by taking the second derivative we can obtain an expression for $f^{(4)}$:

$$\begin{aligned} f^{(4)} &= \frac{f'H^3}{\dot{R}^3}\left[6(y+\Omega-2) - x(y-3) - 3\frac{\ddot{R}}{H\dot{R}}\left(x - 3\Omega - 2y + 4\right) - x\frac{\dddot{R}}{H^2\dot{R}} + 3x\left(\frac{\ddot{R}}{H\dot{R}}\right)^2 - \frac{\dot{R}}{3H^3}\right] \\ &= \frac{f'H^3}{\dot{R}^3}\left[3x(1+q) - 3\Omega + x\frac{K^2}{H^2a^2} + \frac{(s+q^2+8q+6)}{(j-q-2)}(-5x + 6\Omega - 4q - 4)\right. \\ &\quad \left. - x\frac{(l-s-2qj-j-18q^2-48q-24)}{(j-q-2)} + 2x\left(\frac{(s+q^2+8q+6)}{(j-q-2)}\right)^2\right]. \end{aligned}$$

In this form we can easily see the excess Hubble parameters in the coefficients. To eliminate these, we redefine the time and Ricci variables:

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = \frac{1}{H}\frac{\mathrm{d}}{\mathrm{d}t} \implies \frac{\mathrm{d}}{\mathrm{d}t} = H\frac{\mathrm{d}}{\mathrm{d}\tau} \implies \frac{\mathrm{d}^2}{\mathrm{d}t^2} = \dot{H}\frac{\mathrm{d}}{\mathrm{d}\tau} + H^2\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\,. \tag{4.88}$$

Setting $\hat{\mathcal{R}} \equiv \dfrac{\mathcal{R}}{H^2}$ gives

$$\begin{aligned} \frac{\mathrm{d}\hat{\mathcal{R}}}{\mathrm{d}\tau} &= \frac{1}{H^2}\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\tau} - 2\frac{\dot{H}}{H^2}\hat{\mathcal{R}} \\ \frac{\mathrm{d}^2\hat{\mathcal{R}}}{\mathrm{d}\tau^2} &= \frac{1}{H^2}\frac{\mathrm{d}^2\mathcal{R}}{\mathrm{d}\tau^2} - 2\frac{\dot{H}}{H^2}\frac{\mathrm{d}\hat{\mathcal{R}}}{\mathrm{d}\tau} - 2\frac{\ddot{H}}{H^3}\hat{\mathcal{R}}\,. \end{aligned}$$

This allows us to eliminate those remaining Hs. We then change the dependence again to redshift. Starting with the definition of z (2.29),

$$z = \frac{a(t_0) - a(t_i)}{a(t_i)}\,, \tag{4.89}$$

we normalise the scale factor today, $a(t_0) = 1$, to get

$$z = \frac{1}{a} - 1\,, \quad \frac{\mathrm{d}z}{\mathrm{d}\tau} = \frac{1}{H}$$
$$\therefore \frac{\mathrm{d}z}{\mathrm{d}\tau} = \frac{a}{\dot{a}}\left(-\frac{\dot{a}}{a^2}\right) = -\frac{1}{a} = -\frac{1}{z+1} \implies \mathrm{d}z = -\frac{1}{z+1}\,\mathrm{d}\tau$$
$$\therefore \frac{\mathrm{d}}{\mathrm{d}\tau} = -(z+1)\frac{\mathrm{d}}{\mathrm{d}z}$$
$$\therefore \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} = (z+1)\frac{\mathrm{d}}{\mathrm{d}z} + (z+1)^2\frac{\mathrm{d}^2}{\mathrm{d}z^2}\,.$$

Lastly we define the wave number $\hat{k}$ such that

$$\hat{k} = \frac{k}{a_0 H_0}\,, \tag{4.90}$$

where $h = \dfrac{H}{H_0}$. The dimensionless quantity $\hat{k}$ can be thought of as the comoving wavenumber, which scales with the expansion of the Universe. It essentially indicates how the physical size of the perturbation changes relative to the current size of the Universe and the rate of expansion. A larger value of $\hat{k}$ corresponds to smaller physical scales in the Universe, while a smaller value corresponds to larger physical scales. We therefore have

$$\frac{k}{aH} = \frac{\hat{k}(1+z)}{h} \equiv \zeta(z)\,. \tag{4.91}$$

This introduces an h dependence, for which we require another differential equation that will need to be solved alongside the equations for the cosmographic parameters (4.24):

$$\frac{\mathrm{d}h}{\mathrm{d}z} = \frac{h}{1+z}(1+q)\,. \tag{4.92}$$

This means we can finally replace the H variables with the cosmographic parameters and express our system (4.80) in terms of redshift. This results in a pair of second-order equations that describe the redshift evolution of the k-th mode of density perturbations in $f(R)$ gravity in a fully model-independent framework:

$$\begin{aligned} &h^2(1+z)^2\frac{\mathrm{d}^2\Delta_m^k}{\mathrm{d}z^2} + h^2(1+z)(q+x+3w)\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} + h^2\left(w\left[6(q+x)+\zeta^2(z)\right] - 3(1+w)\Omega\right)\Delta_m^k \\ &\quad = (1+z)\left(\frac{(1+w)x}{2(j-q-2)}\right)\frac{\mathrm{d}\hat{\mathcal{R}}^k}{\mathrm{d}z} - \frac{\left[\zeta^2(z)+3q(x-1)+3j-6\right](1+w)}{6(j-q-2)}\hat{\mathcal{R}}^k\,, \end{aligned} \tag{4.93a}$$

$$\frac{\mathrm{d}^2\hat{\mathcal{R}}^k}{\mathrm{d}z^2} + \left(\frac{1+A}{1+z}\right)\frac{\mathrm{d}\hat{\mathcal{R}}^k}{\mathrm{d}z} + \left(\frac{\hat{k}^2}{h^2}(1+z)^2 + B\right)\hat{\mathcal{R}}^k = -\frac{h^2 D}{(1+z)}\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} + \frac{h^2 C}{(1+z)^2}\Delta_m^k\,, \tag{4.93b}$$

where the coefficients A–D can be found in Appendix C.1 and the dimensionless quantities $\hat{\mathcal{R}}^k = \mathcal{R}^k/H_0^2$, $\hat{k} = k/a_0H_0$ and $h = \frac{H}{H_0}$ have been used. We can apply the same procedure to the semi (4.83) and full QS (4.84) perturbation equations. The corresponding equations are, respectively,

$$\begin{aligned} &(1+z)^2\frac{\mathrm{d}^2\Delta_m^k}{\mathrm{d}z^2} + (1+z)\left(q+x+3w+\frac{(w-1)[\zeta^2(z)+3q(x-1)+3j-6]}{E+\zeta^2(z)}\right)\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} \\ &+ \left(w\left[\zeta^2(z)+6q+6x\right] - 3(1+w)\Omega - \frac{[\zeta^2(z)x+3q(x-1)+3j-6]F}{x[E+\zeta^2(z)]}\right)\Delta_m^k = 0\,, \end{aligned} \tag{4.94}$$

where E, F can be found in Appendix C.2; and

$$(1+z)^2\frac{\mathrm{d}^2\Delta_m^k}{\mathrm{d}z^2}+(1+z)(q+3w)\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} \tag{4.95}$$
$$=\left[\Omega\left(\frac{3(3w^2+4w+1)(j-q-2)+(3w^2+5w+2)x\zeta^2(z)}{2(j-q-2)+x\zeta^2(z)}\right)-w\left[6(x+q)+\zeta^2(z)\right]\right]\Delta_m^k\,.$$

In the GR limit $f(R)=R$, the system (4.93) – as well as its approximations (4.94) and (4.95) – reduce to

$$\ddot{\Delta}_m^k-(3w-2)\,H\dot{\Delta}_m^k+\left[w\frac{k^2}{a^2}+\left(\frac{-1+2w-3w^2}{2}\right)\rho_m\right]\Delta_m^k=0\,, \tag{4.96a}$$

$$\mathcal{R}^k=(1-3w)\rho_m\Delta_m^k\,, \tag{4.96b}$$

which, for a dust-only case, gives the well-known equation

$$\ddot{\Delta}_m^k+2H\dot{\Delta}_m^k-\frac{1}{2}\rho_m\Delta_m^k=0\,. \tag{4.97}$$

The key differences between the $f(R)$ perturbation equations and those in GR are their fourth-order structure and, more importantly, the scale-dependence of density perturbations, which arises for any equation of state. If such scale-dependence were detected, it would provide strong evidence for a deviation from ΛCDM.

4.3.4 Growth function for matter perturbations

An alternative method to discriminate between modified gravity and GR models uses the *growth function* S and growth index γ:

$$S\equiv\frac{\mathrm{d}\ln\delta}{\mathrm{d}\ln a}=\tilde{\Omega}_m{}^{\gamma}\,, \tag{4.98}$$

where $\tilde{\Omega}_m$ is the standard matter density parameter,

$$\tilde{\Omega}_m\equiv\frac{\rho_m}{3H^2}=f'\Omega\,, \tag{4.99}$$

and $\tilde{\Omega}_m(z=0)=0.3$. In the covariant formalism, in terms of redshift, this is equivalent to

$$S\equiv-(1+z)\frac{\mathrm{d}\ln\Delta}{\mathrm{d}z}\,, \tag{4.100}$$

from which the growth index can be calculated using

$$\gamma=\frac{\ln\left(-(1+z)\frac{\mathrm{d}\ln\Delta}{\mathrm{d}z}\right)}{\ln\tilde{\Omega}}\,. \tag{4.101}$$

The full equation governing the growth of matter perturbations on sub-horizon scales takes the form

$$\ddot{\Delta}_m+2H\dot{\Delta}_m-4\pi G_{\text{eff}}\rho_m\Delta_m=0\,, \tag{4.102}$$

where G_{eff} is the gravitational constant appearing in the action. This equation describes how small matter density fluctuations in the early universe grow over time under gravitational influence as the Universe expands, and contains a friction term $2H\dot{\delta}$ which incorporates the effect of cosmic expansion. In a rapidly expanding universe, i.e. below the Jeans length (3.63), the fluctuations oscillate at decreasing amplitudes. Below the Jeans length, gravity more effectively

amplifies the density contrast and the fluctuations undergo a power-law growth. When considering modified gravity theories we replace G with $G_{\rm eff}$, which is dependent on the specific model being used [197] and measured using a Cavendish-type experiment[5] to a very high precision [38]:

$$G_{\rm eff} = \frac{G_*}{F}\frac{1+4\frac{k^2}{a^2}\frac{F'}{F}}{1+3\frac{k^2}{a^2}\frac{F'}{F}}\,. \tag{4.103}$$

It is convenient to rewrite (4.102) with the growth index to obtain

$$\frac{\mathrm{d}S}{\mathrm{d}\eta} + f^2 + \frac{1}{2}\left(1 - \frac{\mathrm{d}\ln\Omega_m}{\mathrm{d}\eta}\right)S = \frac{3}{2}\frac{G_{\rm eff}}{G_*}\Omega_m\,. \tag{4.104}$$

From §4.2, we have $\frac{\mathrm{d}\Omega}{\mathrm{d}\ln a} = 3w\Omega(1-\Omega)$, allowing us to write the equation for S in terms of Ω:

$$3w\Omega_m(1-\Omega_m)\frac{\mathrm{d}S}{\mathrm{d}\Omega_m} + S^2 + \frac{1}{2}\left[1-3w(1-\Omega_m)\right]S = \frac{3}{2}\frac{G_{\rm eff}}{G_*}\Omega_m\,. \tag{4.105}$$

We can then convert to redshift dependence, giving

$$-(1+z)\frac{\mathrm{d}S}{\mathrm{d}z} + S^2 + \left[\frac{1}{2} - \frac{3}{2}w(1-\Omega_m)\right]S = \frac{3}{2}\frac{\Omega_m}{F}\frac{3(-q+j-2)+2x(1+z)^2\frac{\hat{k}^2}{h^2}}{2(-q+j-2)+x(1+z)^2\frac{\hat{k}^2}{h^2}}\,. \tag{4.106}$$

When first introduced, the growth index was assumed to be constant at low redshifts [198, 199]. Although the time dependence is evident, this approximation remains valid for models within GR that have constant or smoothly varying equations of state, characterised by $|\gamma_0' \equiv \frac{\mathrm{d}\gamma}{\mathrm{d}z}| \lesssim 0.02$ [197]. Specifically, the ΛCDM model exhibits $\gamma_0 \approx 0.55$ with minimal variation for $z < 0.55$ [200]. Significant deviations from these values, pronounced time variation at low redshifts, or scale dependence of γ_0 and γ_0' could serve as compelling evidence for modifications to GR. For a ΛCDM universe, the growth factor γ can be defined relative to S via the approximation

$$S \cong \Omega_m(z)^{\gamma(z)}\,, \tag{4.107}$$

which takes into account the redshift dependency. Substituting $\gamma(z)$ and rewriting as a function of redshift, with $\gamma' \equiv \frac{\mathrm{d}\gamma}{\mathrm{d}z}$, gives

$$-(1+z)\Omega_m^\gamma \ln\Omega_m\gamma' + \Omega_m^{2\gamma} + \frac{1}{2}\left[1+3\left(2\gamma-1\right)w_{\rm eff}\right]\Omega_m^\gamma = \frac{3}{2}\frac{G_{\rm eff}}{G_*}\Omega_m\,. \tag{4.108}$$

Both γ and γ' can be constrained from a number of observables, including rich clusters [201]. These are some of the largest virialised objects in the Universe, so their abundance is closely related to the linear mass power spectrum $P(k)$. Cluster mass can be inferred from X-ray temperatures, allowing application of Press-Schechter theory [202] and giving a constraint on the parameters σ_8 and Ω_m^γ. Redshift space distortions allow us to infer the value of $f\sigma_8$ and subsequently γ [203], and the Planck collaboration combined CMB data with large-scale structure observations to constrain γ [24]. The cosmological parameters can also be extracted from redshift data using tomographic weak lensing [204]. Statistical methods, such as using parametric models [205] or Bayesian inference with MCMC analysis can also be used [206].

[5] Devised in 1797, the Cavendish experiment used a torsion balance to measure the gravitational attraction between known masses. Modern experiments use more sensitive torsion balances, vacuum environments, and laser measurements to determine G.

Now we'd like to express (4.102) using our dynamical variables so we can plot it with the solutions we calculated in §4.3. The last term is

$$4\pi G_{\rm eff}\rho_m\delta_m = 4\pi\frac{G_*}{F}\left(\frac{1+4\frac{k^2}{a^2}\frac{F'}{F}}{1+3\frac{k^2}{a^2}\frac{F'}{F}}\right)\left(\frac{3H^2\Omega_m}{8\pi G_*}\right)\delta_m$$
$$= \frac{3}{2}H^2\Omega\left(\frac{\dot{R}+4xH\frac{k^2}{a^2}}{\dot{R}+3xH\frac{k^2}{a^2}}\right)\delta_m\,,$$

where we recalled that $x \equiv \frac{\dot{R}F'}{HF}$. Equation (4.102) as a function of redshift is therefore

$$(z+1)^2\frac{\mathrm{d}^2\Delta}{\mathrm{d}z^2} - (z+1)\left(1+\frac{\dot{H}}{H^2}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}z} - \frac{3}{2}\Omega\left(\frac{\dot{R}+4xH\frac{k^2}{a^2}}{\dot{R}+3xH\frac{k^2}{a^2}}\right)\Delta_m = 0\,. \tag{4.109}$$

Now we define

$$\mathcal{D} \equiv \frac{\mathrm{d}\ln\Delta}{\mathrm{d}z} = \frac{\dot{\Delta}}{H\Delta(z+1)} \implies \frac{\mathrm{d}\mathcal{D}}{\mathrm{d}z} = \frac{1}{\Delta}\frac{\mathrm{d}^2\Delta}{\mathrm{d}z^2} - \mathcal{D}^2\,,$$

and using definitions (4.86), (4.87) and (4.92) we obtain

$$(1+z)^2\frac{\mathrm{d}\mathcal{D}}{\mathrm{d}z} + (1+z)^2\mathcal{D}^2 + q(1+z)\mathcal{D} - \Omega\left(\frac{3(j-q-2)+2x(1+z)^2\frac{\hat{k}^2}{h^2}}{2(j-q-2)+x(1+z)^2\frac{\hat{k}^2}{h^2}}\right)\,.$$

4.4 Application to ΛCDM background cosmology

We can now apply our dynamical system formulation to the ΛCDM model, for which we constrain the jerk parameter to $j = 1$. This enables us to solve the background exact and quasi-static evolution equations for our $f(R)$ theory at discrete wavenumbers and compare them to ΛCDM.

4.4.1 Dynamical system in ΛCDM

Taking the system to be in flat space means that $K = 0$ and we can eliminate the dimensionless variable K from our system in (4.14). This allows us to simplify the constraint equation (4.31) to

$$y = 1 - q\,. \tag{4.110}$$

As mentioned previously, expressing $\Gamma(R)$ in terms of the cosmographic parameters allows us to formulate a model-independent dynamical system. We achieve this by differentiating the dimensionless variables – bar K – and the cosmographic parameters with respect to redshift z, obtaining

$$\frac{\mathrm{d}x}{\mathrm{d}z} = \frac{1}{(z+1)}\Big[3(1+w)\Omega - (2+q-x)x - 2q - 2\Big]\,, \tag{4.111a}$$
$$\frac{\mathrm{d}\Omega}{\mathrm{d}z} = \frac{\Omega}{(1+z)}\Big[x+1+3w-2q\Big]\,, \tag{4.111b}$$
$$\frac{\mathrm{d}h}{\mathrm{d}z} = \frac{h}{(1+z)}\Big[1+q\Big]\,, \tag{4.111c}$$
$$\frac{\mathrm{d}q}{\mathrm{d}z} = \frac{1}{(1+z)}\Big[j-q-2q^2\Big]\,, \tag{4.111d}$$
$$\frac{\mathrm{d}j}{\mathrm{d}z} = -\frac{1}{(1+z)}\Big[s+j(2+3q)\Big]\,, \tag{4.111e}$$
$$\frac{\mathrm{d}s}{\mathrm{d}z} = -\frac{1}{(1+z)}\Big[s(3+4q)+l\Big]\,, \tag{4.111f}$$

where $h \equiv H/H_0$, and we used (4.110) to eliminate y and z. This system can be decoupled further provided $l = l(h, q, j, s)$, as the last four equations are completely kinematic in nature, together forming another closed dynamical system. The first two equations contribute the gravitational dynamics. This formulation allows us to study the phase space of any $f(R)$ theory that can reproduce a cosmic evolution for $l = l(q, j)$ with explicitly reconstructing the specific form of $f(R)$, as in [178, 207, 179]. Recall that avoiding ghost and tachyonic instabilities requires $f' > 0$ and $f'' \geq 0$, so in flat space we have the following constraint on the phase space from (4.36):

$$\frac{x}{-q + j - 2} \geq 0\,. \tag{4.112}$$

The viable portion of the phase space is the region which satisfies this constraint, but this is not necessarily sufficient to avoid instabilities. We must explicitly calculate f', f'' along a trajectory of interest and ensure that, in addition to (4.112), the dynamical variable Ω must also satisfy $\Omega \equiv \frac{\rho}{3f'H^2} \geq 0$. This ensures that both f' and f'' are positive, and thus the physical viability of the region.

4.4.2 The cosmographic condition $j = 1$

It is generally accepted that the ΛCDM model exhibits a cosmic evolution corresponding to $j = 1$ [91], and cosmic chronometers and supernova data support these predictions to within 3σ [208, 209]. When certain deviations from the standard ΛCDM parameters (e.g. matter density Ω_{m0} or dark energy density $\Omega_{\Lambda 0}$) are assumed and constrained through data fitting, the resulting 'reconstructed' cosmographic quantities still match predictions. This suggests that even with small variations, the overall behaviour remains similar to ΛCDM. The redshift evolution of the jerk parameter [208] is

$$j(z) = \frac{H(z)(1+z)^2 \frac{\mathrm{d}^2 H(z)}{\mathrm{d}z^2} + \left[(1+z)\frac{\mathrm{d}H(z)}{\mathrm{d}z} - H(z)\right]^2}{H^2(z)}\,. \tag{4.113}$$

Solving for $j(z) = 1$ we obtain

$$h^2(z) \equiv \frac{H^2(z)}{H_0^2} = \mathcal{C}_1(1+z)^3 + \mathcal{C}_2\,, \tag{4.114}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are two arbitrary constants satisfying $\mathcal{C}_1 + \mathcal{C}_2 = 1$. For $\mathcal{C}_1, \mathcal{C}_2 \neq 0$, the family of solutions (4.114) specifies a ΛCDM evolution history, with two clear asymptotic limits: an effective CDM limit at asymptotic past $z \to \infty$ where $h^2 \sim (1+z)^3$; and an effective Λ limit at asymptotic future $z \to -1$ where $h^2 \sim$ constant. The particular solution, specified by $\mathcal{C}_1 = \Omega_{m0}$ and $\mathcal{C}_2 = \Omega_{\Lambda 0}$, corresponds to the particular ΛCDM evolution of our universe.

The condition $j(z) \approx 1$ does not uniquely distinguish ΛCDM as the underlying model. This 'cosmographic degeneracy' means that multiple cosmological models could produce similar observational signatures that align with ΛCDM, despite having different underlying dynamics or physical assumptions [210, 211, 212]. We can increase the specificity of the conditions by using the statefinder parameters $\{r, s\}$ introduced in [213]. The first parameter r is just the jerk parameter – which alone cannot specify a model – while the second parameter s is defined as

$$s \equiv \frac{r-1}{3(q-\frac{1}{2})} = 1 + w - \frac{1}{3}\frac{\dot{w}}{wH}\,, \tag{4.115}$$

making it a linear combination of both the deceleration and jerk parameters. Crucially, s does not explicitly depend on the dark energy density or matter density; therefore degeneracies present

in r are resolved by the pair $\{r, s\}$. If the role of dark energy is represented by a cosmological constant $w = -1$ as in ΛCDM, the value of r remains locked at $r = 1$ throughout the matter-dominated epoch, regardless of the value of Ω_m. A trivial fixed point for ΛCDM is thus $\{r, s\} = \{1, 0\}$. Alam et al [214] showed that the statefinder pair can easily distinguish between ΛCDM ($w = -1$), quiessence (constant w), and kinessence (time-varying w) models.

A physical interpretation of the constants $\mathcal{C}_1, \mathcal{C}_2$ requires us to specify the model [208]. For ΛCDM, dark matter and dark energy evolve independently (as $(1+z)^3$ and $(1+z)$, respectively), allowing us to identify $\mathcal{C}_1$ with Ω_{m0} and $\mathcal{C}_2$ with $\Omega_{\Lambda 0}$. If, on the other hand, we took $w_{\rm DE} \neq 1$ while still maintaining the condition $j(z) = 1$, the constants would be $\{\mathcal{C}_1, \mathcal{C}_2\} = \{1 + w_{\rm DE} - w_{\rm DE}\Omega_{m0}, -w_{\rm DE}\Omega_{\rm DE0}\}$ [215]. The relationships $\mathcal{C}_1 = \mathcal{C}_1(\Omega_{\rm DE0}, \Omega_{m0})$ and $\mathcal{C}_2 = \mathcal{C}_2(\Omega_{\rm DE0}, \Omega_{m0})$ are therefore sensitive to the choice of model. We can conclude that the condition $j = 1$ is sufficient for reconstructing a similar cosmic evolution to ΛCDM, but not the explicit dynamics. The values of $\mathcal{C}_1$ and $\mathcal{C}_2$ do not necessarily represent the abundances of dark matter and dark energy at present [210].

In the context of $f(R)$ gravity, we can rewrite the field equations (4.6b) for a late-time $f(R)$ model with no dark-sector interaction[6]:

$$3H^2 = \rho_{\rm tot} = \rho + \rho_{\rm DE}\,, \tag{4.116a}$$

$$-\left(2\dot{H} + 3H^2\right) = p_{\rm tot} = p_{\rm DE}\,, \tag{4.116b}$$

where ρ is the energy density of the non-relativistic matter, the energy density and pressure of the dark matter are defined as

$$\rho_{\rm DE} = \frac{1}{2}(RF - f) - 3H\dot{F} + 3H^2(1 - F)\,, \tag{4.117a}$$

$$p_{\rm DE} = \ddot{F} + 2H\dot{F} - \frac{1}{2}(RF - f) - (2\dot{H} + 3H^2)(1 - F)\,. \tag{4.117b}$$

The dark energy equation of state $w_{\rm DE} = \frac{p_{\rm DE}}{\rho_{\rm DE}}$ is

$$w_{\rm DE} = \frac{\ddot{F} + 2H\dot{F} - \frac{1}{2}(RF - f) - (2\dot{H} + 3H^2)(1 - F)}{\frac{1}{2}(RF - f) - 3H\dot{F} + 3H^2(1 - F)}\,. \tag{4.118}$$

Using the field equations (4.116a), this can be written as

$$w_{\rm DE} = -\frac{2\dot{H} + 3H^2}{3H^2 - \rho} = \frac{H^2 - \frac{R}{3}}{3H^2 - \rho}\,. \tag{4.119}$$

In this form, the non-relativistic matter and the dark energy are separately conserved

$$\dot{\rho} + 3H\rho = 0\,, \tag{4.120a}$$

$$\dot{\rho}_{\rm DE} + 3H\rho_{\rm DE}(1 + w_{\rm DE}) = 0\,. \tag{4.120b}$$

We divide by H^2 and substitute the definitions of the dynamical variables and cosmographic parameters to obtain

$$w_{\rm DE} = \frac{2q - 1}{3 - 3\tilde{\Omega}}\,, \tag{4.121}$$

where we used $y = 1 - q$ from (4.32) in flat space and $\tilde{\Omega}$ – defined in (4.99) – obeys the dynamical equation

$$\frac{{\rm d}\tilde{\Omega}}{{\rm d}\eta} = -\tilde{\Omega}(1 - 2q)\,. \tag{4.122}$$

[6] No coupling is present between dark matter and dark energy components.

By demanding $w_{\rm DE} = -1$ we obtain

$$2q - 3\tilde{\Omega} = -2\,. \tag{4.123}$$

If we take the derivative, using equations (4.122) and (4.24), we arrive at the condition $j = 1$. We can then conclude that if the curvature part of the total energy density $\rho_{\rm DE}$ mimics a cosmological constant ($w_{\rm DE} = -1$), it yields a cosmology with $j(z) = 1$, as expected.

We now consider a dynamical dark energy model with $w_{\rm DE} = -\alpha(\eta)$, which we substitute into (4.121) to obtain

$$j - 1 = \left(q - \frac{1}{2}\right)\left(3 - 3\alpha - \frac{1}{\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}\eta}\right)\,. \tag{4.124}$$

One possible solution is $q = \frac{1}{2}$, but this cannot remain constant throughout the expansion history; another possibility is that $\alpha = 1$ and thus $w_{\rm DE} = -1$, which we have already covered. A third solution is that $\alpha(\eta)$ satisfies

$$\frac{\mathrm{d}\alpha}{\mathrm{d}\eta} = 3\alpha(1-\alpha)\,, \tag{4.125}$$

which, when solved, gives

$$\alpha(\eta) = \frac{\alpha_0 e^{3\eta}}{\alpha_0 e^{3\eta} + (1-\alpha_0)}\,, \tag{4.126}$$

where $\alpha_0 = \alpha(0)$. This implies a dark energy equation of state

$$w_{\rm DE}(\eta) = -\frac{w_{\rm DE}(0)e^{3\eta}}{w_{\rm DE}(0)e^{3\eta} - (1 + w_{\rm DE}(0))}\,. \tag{4.127}$$

Excepting the special case $w_{\rm DE}(0) = -1$, (4.127) corresponds to a dynamical dark energy model which is compatible with the condition $j(z) = 1$. We can expand on this by including the constants $\mathcal{C}_1$ and $\mathcal{C}_2$ in (4.120) to obtain

$$\rho_{\rm DE} \propto \frac{1}{\alpha} = \tilde{\mathcal{C}}_1 e^{3\eta} + \tilde{\mathcal{C}}_2 = \tilde{\mathcal{C}}_1(1+z)^3 + \tilde{\mathcal{C}}_2\,, \tag{4.128}$$

where the tilde distinguishes the constants from those in ΛCDM, allowing us to write (using (4.116a))

$$h^2(z) = \left(\Omega_{m0} + \frac{\tilde{\mathcal{C}}_1}{3H_0^2}\right)(1+z)^3 + \frac{\tilde{\mathcal{C}}_2}{3H_0^2}\,. \tag{4.129}$$

In other words, even though (4.127) describes a dynamical dark energy model, the cosmic evolution remains similar to ΛCDM. This similarity arises from the existence of a well-defined effective CDM limit in the asymptotic past and an effective Λ limit in the asymptotic future. Since our subsequent analysis relies on the condition $j = 1$, these models are naturally encompassed within our results.

4.4.3 Background evolution with $j = 1$

In this section we examine the background cosmic evolution and growth of matter perturbations for an $f(R)$ theory that mimics the evolutionary history of ΛCDM. As discussed in §4.4.2, this can be achieved by fixing $j = 1$, which subsequently fixes the snap and lerk parameters:

$$s = \frac{\mathrm{d}j}{\mathrm{d}\eta} - j(2+3q) = -(2+3q)\,, \tag{4.130}$$

$$l = \frac{\mathrm{d}s}{\mathrm{d}\eta} - s(3+4q) = 6q^2 + 14q + 9\,. \tag{4.131}$$

For a dust equation of state ($w = 0$), fixing $j = 1$ also closes the dynamical system (4.111):

$$\frac{\mathrm{d}x}{\mathrm{d}z} = \frac{1}{(z+1)}\Big[3\Omega - (2+q-x)x - 2q - 2\Big]\,, \tag{4.132a}$$

$$\frac{\mathrm{d}\Omega}{\mathrm{d}z} = \frac{\Omega}{(1+z)}\Big[x + 1 - 2q\Big]\,, \tag{4.132b}$$

$$\frac{\mathrm{d}h}{\mathrm{d}z} = \frac{h}{(1+z)}\Big[1+q\Big]\,, \tag{4.132c}$$

$$\frac{\mathrm{d}q}{\mathrm{d}z} = \frac{1}{(1+z)}\Big[1 - q - 2q^2\Big]\,. \tag{4.132d}$$

The background variables' initial conditions are established deep within the matter-dominated era, with Ω and q set to align with ΛCDM values. Specifically, we assume $\Omega_0^{\Lambda\mathrm{CDM}} = \tilde{\Omega}(z = 0) = 0.3$ and $q_0 = q(z = 0) = -0.55$. The dynamical system is highly sensitive to initial conditions, often requiring precision up to the fourth decimal place. As a result, specifying initial conditions at $z = 0$ without sufficient accuracy may lead to an incorrect evolution, skipping the matter-dominated phase and instead producing a scalaron-dominated regime for $z \gtrsim 2$. To ensure consistency with ΛCDM evolution, initial conditions must be set deep within the matter-dominated era. For $j = 1$, the viability condition (4.112) becomes

$$\frac{x}{1+q} \leq 0\,. \tag{4.133}$$

Since $q > -1$, we must have $x \leq 0$ for all redshifts, where $x \to 0$ represents the GR limit. The value of $|x|$ thus describes the deviation from GR and should be small in the matter-dominated era. The value of x_{in} is chosen such that $x < 0$ for the entirety of its evolution up to $z = 0$ and results in a value of $\Omega_0 \approx 0.3$. Figure 4.2 shows the effect of this value on the x_0 and Ω_0. We select the following values that satisfy these conditions:

$$x_{\mathrm{in}} = -0.001\,, \quad \Omega_{\mathrm{in}} = 0.993243\,, \quad q_{\mathrm{in}} = 0.489865\,. \tag{4.134}$$

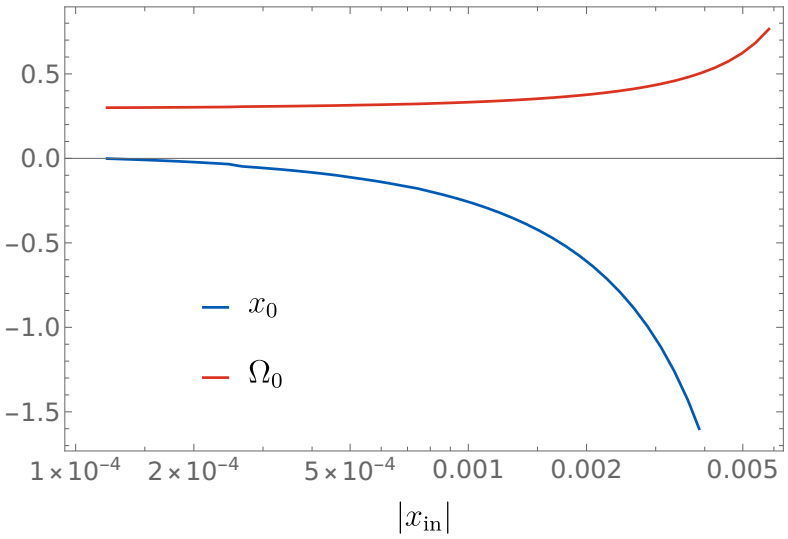
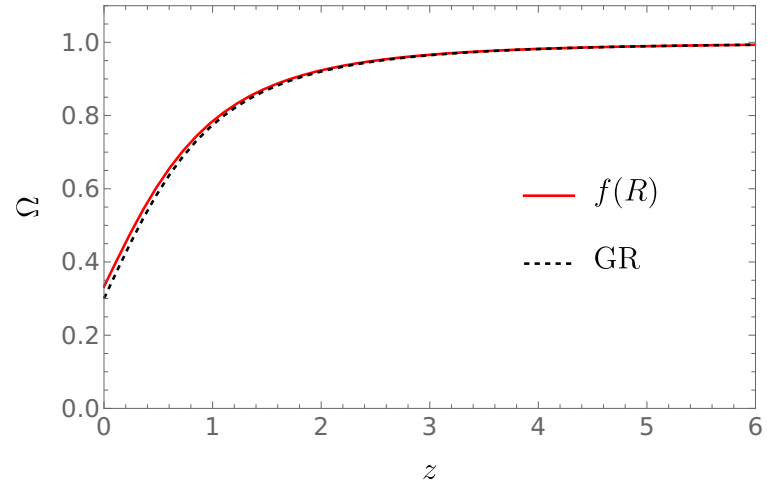
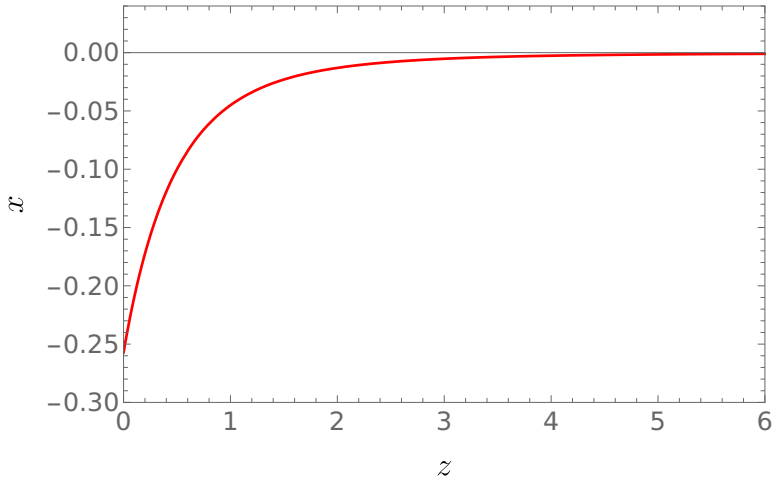


Figure 4.2: The values of x_0 and Ω_0 depending the chosen value of $|x_{\mathrm{in}}|$.

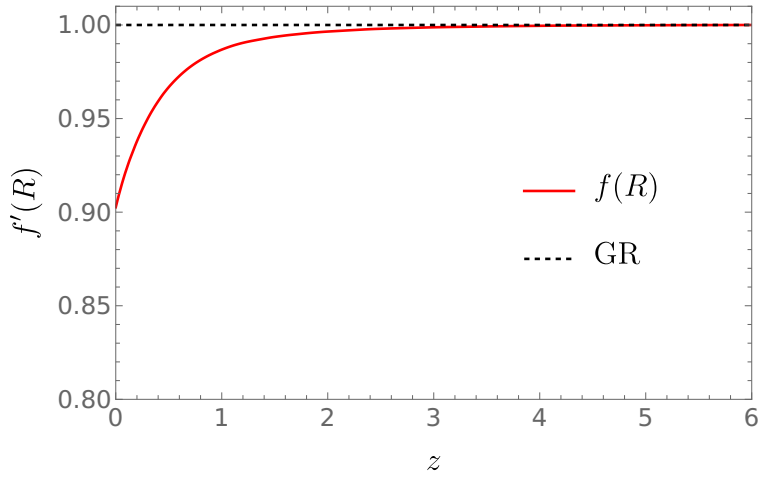
These initial values yield the background evolutions in figure 4.3. In figure 4.3a we can see that the $f(R)$ background variable Ω follows a very similar trajectory to ΛCDM. The trajectory of $f'(R) = \Omega_m/\Omega$ in figure 4.3c meets all the viability conditions we have discussed: $x < 0$ for all time (as seen in figure 4.3b); $f' > 0$; and (4.133) is satisfied, ensuring that $f'' > 0$.



(a)

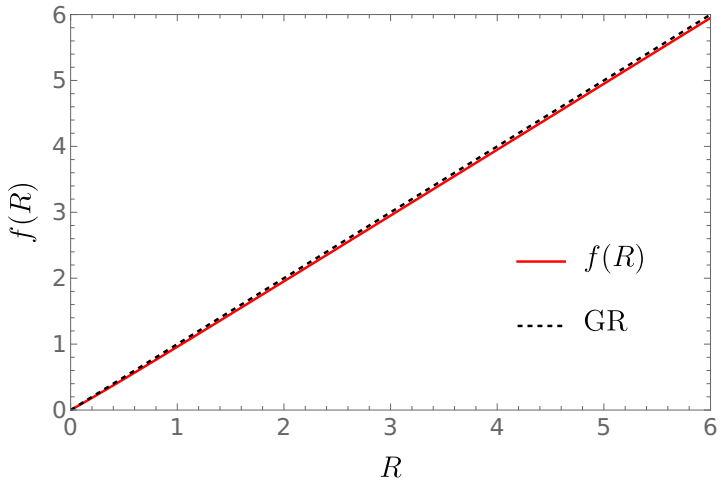


(b)

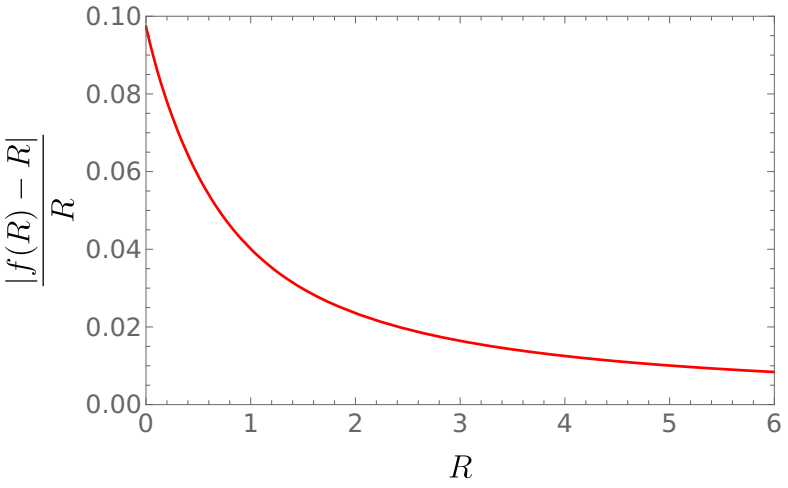


(c)

Figure 4.3: Evolution of the background variables (a) Ω and (b) x with initial conditions (4.134), and (c) $f'(R) = \Omega_m/\Omega$. The trajectories are shown in red for $f(R)$ and black for GR.



(a)



(b)

Figure 4.4: (a) $f(R)$ (red), obtained from numerical integration of $f'(R)$, compared with GR (black dashed) as a function of R. (b) Percentage deviation of $f(R)$ from GR as a function of R.

We now analyse matter perturbations in a dust-only $f(R)$ model with a ΛCDM background. A fundamental feature of modified gravity theories is the transition between the 'GR regime' and the modified regime. In $f(R)$ models, this transition is governed by the parameter

$$M^2 \equiv \frac{f'}{3f''} = \frac{-2(1+q)}{x}\,, \tag{4.135}$$

which represents the squared mass of the scalaron in the regime where $M^2 \gg R$ [216]. When $M \ll k/(aH)$, characteristic of the $f(R)$ regime, the scalaron mass is low, leading to a finite-range 'fifth force' that affects perturbations [217]. On the other hand, in the GR regime where $M \gg k/(aH)$, the scalaron becomes massive, suppressing the fifth force and making the system behave as in GR. The transition between these regimes is both time- and scale-dependent, causing different k-modes to shift from the GR regime to the $f(R)$ regime at different times. These differential deviations occur when $M = \frac{k}{aH}$ for any perturbation k-mode, corresponding to a redshift z_c related to M via

$$M = \sqrt{\frac{-2(1+q(z_c))}{x(z_c)}} = \frac{k}{aH}\,. \tag{4.136}$$

Figure 4.5 illustrates the redshift at which this transition occurs, showing that smaller scales (larger k) transition earlier, while larger scales (smaller k) either transition near $z = 0$ or remain in the GR regime. These modes are indistinguishable from ΛCDM in the latter case, as they have always remained within the GR regime. The transition redshift explicitly depends on

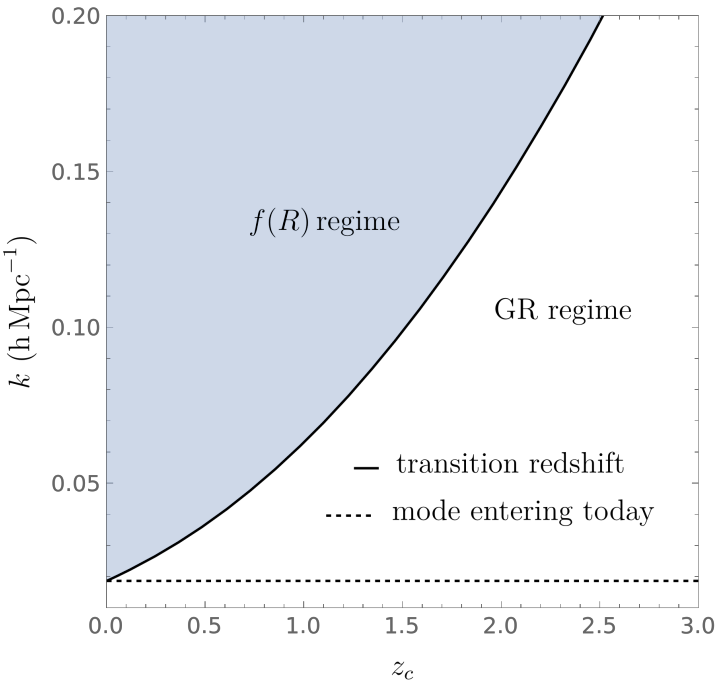


Figure 4.5: Transition redshift as a function of scale for matter perturbations, indicating the shift from the GR regime to the $f(R)$ regime. The shaded region marks the $f(R)$ regime, where the fifth force is active, while the unshaded area corresponds to the GR regime, where this force is suppressed. The dashed line is the perturbation mode $k = 0.019\ h\ \mathrm{Mpc}^{-1}$, which enters the $f(R)$ regime at $z = 0$. Larger scales (smaller k) remain in the GR regime today.

the evolution of x, so varying the initial condition $x_{\rm in}$ affects which modes have entered the $f(R)$ regime by the present time. As demonstrated in figure 4.6, for larger values of $|x_{\rm in}|$, such as $|x_{\rm in}| \sim 0.003$, all galaxy power spectrum-relevant modes ($0.01\,h\,\mathrm{Mpc}^{-1} \leq k \leq 0.2\,h\,\mathrm{Mpc}^{-1}$)

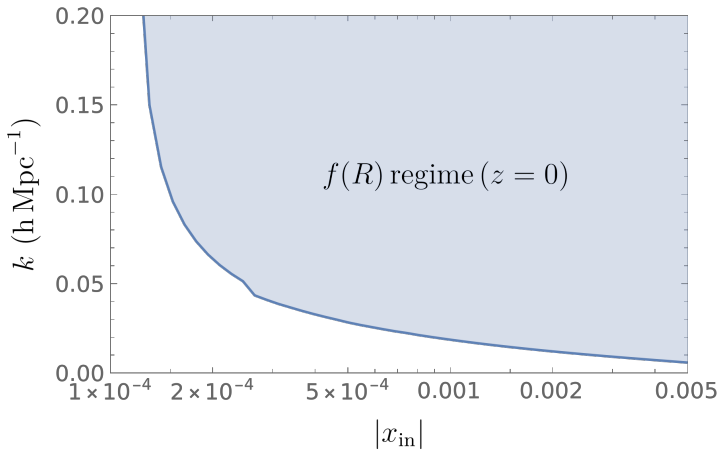


Figure 4.6: Range of perturbation modes transitioning from the GR regime to the $f(R)$ regime, shown as a function of the initial value $|x_{\rm in}|$. The shaded region indicates modes that entered the $f(R)$ regime at or before $z = 0$. All relevant perturbation modes remain in the GR regime for a background evolution closely aligned with GR ($|x_{\rm in}| \sim 10^{-4}$ during the matter-dominated era).

have entered the $f(R)$ regime by today. Conversely, for very small values of $|x_{\rm in}|$, these modes remain in the GR regime, making them identical to ΛCDM. The choice of $x_{\rm in} = -0.001$ therefore ensures that most relevant perturbation modes have transitioned to the $f(R)$ regime, allowing the $f(R)$ modification to be effective. For the background, we set $j = 1$ in the exact (4.93a), semi-QS (4.94) and fully QS (4.95) equations, which simplifies them to:

Exact

$$h^2\frac{\mathrm{d}^2\Delta_m^k}{\mathrm{d}z^2} + h^2\frac{(q+x)}{(1+z)}\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} + 3h^2\frac{\Omega}{(1+z)^2}\Delta_m^k = -\frac{x}{2(q+1)(1+z)}\frac{\mathrm{d}\hat{\mathcal{R}}^k}{\mathrm{d}z} + \left(\frac{\frac{\hat{k}^2}{h^2}(1+z)^2x + 3q(x-1) - 3}{6(q+1)(1+z)^2}\right)\hat{\mathcal{R}}^k\,, \tag{4.137a}$$

$$\frac{\mathrm{d}^2\hat{\mathcal{R}}^k}{\mathrm{d}z^2} - \frac{(q(x+4)+11x-6\Omega+4)}{x(1+z)}\frac{\mathrm{d}\hat{\mathcal{R}}^k}{\mathrm{d}z} + \left(\frac{\hat{k}^2}{h^2} + \frac{2q(q(x+2)+6x-3\Omega+10)-27\Omega+16+34x}{x(1+z)^2}\right)\hat{\mathcal{R}}^k = \frac{6h^2(q+1)}{(1+z)}\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} - \frac{6h^2(q+1)\Omega}{x(1+z)^2}\Delta_m^k\,; \tag{4.137b}$$

Semi-QS

$$\frac{\mathrm{d}^2\Delta_m^k}{\mathrm{d}z^2} + \left(\frac{q+x(1-\tilde{C})}{(1+z)}\right)\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} = \Omega\left(\frac{3-\tilde{C}}{(1+z)^2}\right)\Delta_m^k\,, \tag{4.138}$$

where

$$\tilde{C} = \frac{\left(\frac{\hat{k}^2(z+1)^2}{h^2}x - 3(1+q) + 3qx\right)}{\left(\frac{\hat{k}^2(z+1)^2}{h^2}x + (34+2q^2+12q)x + 4(6q^2+5q+4) - (6q+27)\Omega\right)}\,; \tag{4.139}$$

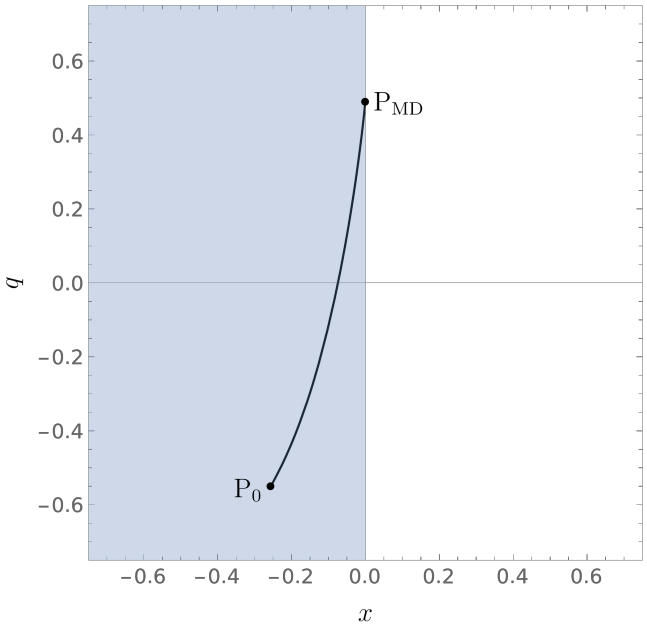


Figure 4.7: Parametric plot of trajectory with initial conditions (4.134), starting in the matter-dominated era $P_{\rm MD} = \{-0.001, 0.489865\}$ and evolving to the present values $P_0 = \{-0.25715, -0.55\}$.

Full QS

$$(1+z)^2\frac{\mathrm{d}^2\Delta_m^k}{\mathrm{d}z^2} + (1+z)q\frac{\mathrm{d}\Delta_m^k}{\mathrm{d}z} = \Omega\left(\frac{2\frac{\hat{k}^2}{h^2}x - 3\frac{(q+1)}{(1+z)^2}}{\frac{\hat{k}^2}{h^2}x - 2\frac{(q+1)}{(1+z)^2}}\right)\Delta_m^k\,. \tag{4.140}$$

Following [188], the scale-independent initial conditions in the CMB era at $z_{in} = 2000$ are set to

$$\Delta_m^k|_{in} = \hat{\mathcal{R}}^k|_{in} = 10^{-5}\,, \quad (\Delta_m^k)'|_{in} = (\hat{\mathcal{R}}^k)'|_{in} = 0\,. \tag{4.141}$$

The matter perturbation evolution corresponding to the background trajectory shown in figure 4.7 is illustrated in Figures 4.8 and 4.9 for different methods. In figure 4.8, Δ_m is normalised at $z = 0$, facilitating a comparison with similar plots in model-dependent studies [38], which analysed perturbation evolution in Starobinsky's model. However, this normalisation does not clearly highlight the scale dependence of Δ_m at low redshifts. To address this, figure 4.9 presents Δ_m normalised at $z = 2000$, offering a clearer depiction of scale dependence at very low redshifts ($z \approx 0$). Across all relevant modes, the full covariant and semi-QS methods exhibit minimal scale dependence at low redshifts, while the full QS method shows more pronounced scale dependence.

Figures 4.10, 4.11, and 4.12 display the growth function S (4.100) and the growth index γ (4.101) for the full covariant, semi-QS, and full QS perturbation methods, respectively. The scales are split to illustrate general trends between large and small scales. All methods (shown in the right column) produce similar results on smaller scales, with a minor scale dependence of S at $z = 0$ that becomes more apparent around $z \sim 1.5$. According to figure 4.5, these scales have been in the $f(R)$ regime since $z \simeq 1$, allowing the fifth force to influence the growth, leading to significant deviations from ΛCDM. The trend reverses for larger scales, where scale dependence is more prominent at $z = 0$. In figure 4.12, the growth function S and growth index γ for the full QS method significantly differ from the other methods, with the growth function only slightly deviating from ΛCDM. The growth index γ does not match ΛCDM values precisely, but its suppression is weaker, resulting in slower growth compared to the other methods. This underscores the full QS method's inability to distinguish $f(R)$ from ΛCDM at these scales. Figure

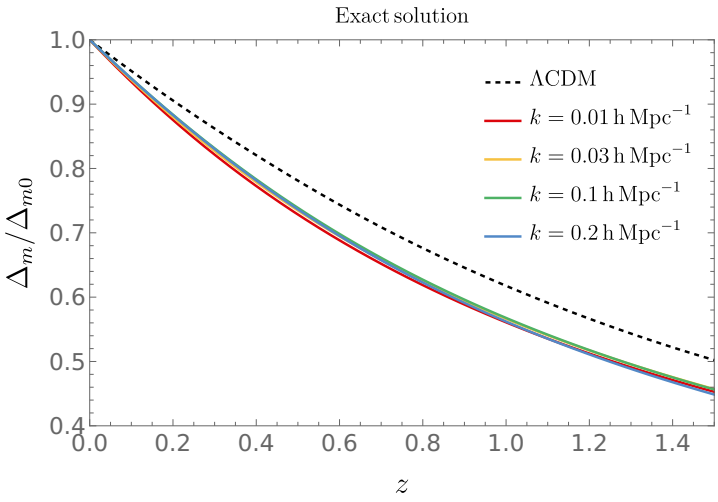
4.13 quantifies the differences between methods, plotting the percentage deviations of semi-QS and full QS solutions from the exact solution for both large and small scales. We can see that i), the deviation between QS-based and exact solutions is greater for larger scales and smaller for smaller scales; and ii) for smaller scales, the full QS approximation aligns more closely with the exact solution compared to the semi-QS approximation. In contrast, the semi-QS approximation performs better for larger scales.

This scale-dependent behaviour highlights the importance of carefully applying the QS approximation to $f(R)$ gravity. Despite methodological differences, the results align broadly with previous model-specific or constrained $f(R)$ analyses [218, 219, 220, 197, 38, 221]. They exhibit pronounced scale dependence at smaller scales, a larger growth rate, and significant deviations from ΛCDM in the growth index γ at low redshifts. Additionally, there is notable time dependence in γ, even at low redshifts. $f(R)$ models typically predict a growth index approximated as $\gamma \simeq \gamma_0 + \gamma_0' z$ at low redshifts, with γ_0 and γ_0' inferred from observables such as galaxy clusters, redshift-space distortions, and weak lensing [203, 24, 107]. Current constraints, however, are insufficient to rule out many dark energy and modified gravity models.

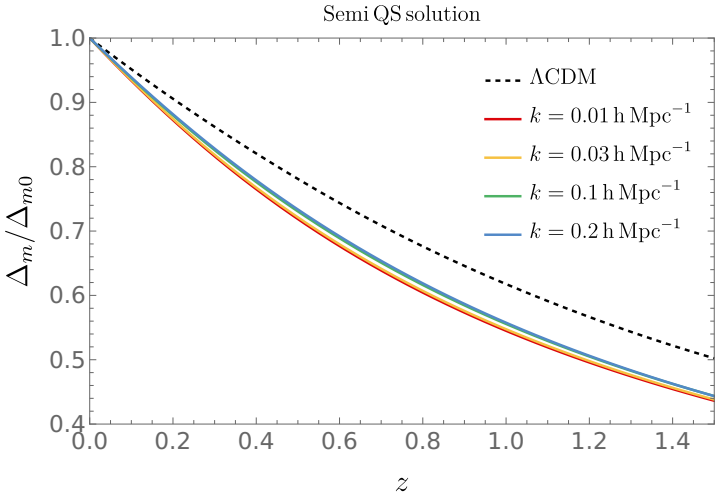
To directly compare with prior $f(R)$ analyses, figure 4.14 shows the scale dependence of γ_0 and γ_0' for each method. All methods converge for $k \gtrsim 0.1$, where both QS approximations are valid. For small scales, the semi-QS approximation yields γ_0 values farther from the exact and full QS solutions. At large scales, the full QS solution for γ_0 aligns more closely with ΛCDM, underscoring its inadequacy for distinguishing $f(R)$ from ΛCDM. In contrast, γ_0' in figure 4.14b shows more significant deviations across all scales for the semi-QS method compared to the full QS method. Focusing on the exact solutions, these model-independent results agree with prior model-dependent analyses, providing $0.3 \lesssim \gamma_0 \lesssim 0.4$ and $-0.2 \lesssim \gamma_0' \lesssim -0.07$ [197, 38]. Both parameters exhibit a dispersion of at least ~ 0.1 across scales, which could serve as a robust indicator of $f(R)$ gravity if detected by future high-precision observations.

Our results[7] show that this method, both with and without QS approximations, effectively captures the overall trends observed in prior model-specific or constrained $f(R)$ studies, such as those in [197, 200, 222]. The density contrast Δ_m, growth function Ω_m^γ, and growth index γ exhibit significant scale dependence, which becomes more pronounced at smaller scales for redshifts $z \gtrsim 0.5$. Furthermore, γ varies notably with time, taking lower values that indicate substantial deviations from ΛCDM. Notably, we determined the current values of the growth index and its derivative (γ_0 and γ_0') for ΛCDM -mimicking $f(R)$ models without assuming a specific functional form of $f(R)$. Since explicitly reconstructed $f(R)$ models that replicate ΛCDM are often expressed in terms of hypergeometric functions [207, 180], computing these quantities and their variations would have been highly complex without this approach. By comparing the exact, semi-QS, and full-QS methods, we identified limitations of the full-QS approximation on large scales ($k \lesssim 0.08\,h\,\mathrm{Mpc}^{-1}$), where perturbations have not remained in the $f(R)$ regime for extended periods and thus tend to follow ΛCDM more closely. At smaller scales, the differences between the methods are minimal. Regardless of the chosen approach, these findings strongly support many studies on $f(R)$ structure growth while avoiding the constraints of predefining $f(R)$ and the computational challenges of reconstruction methods. Even when solving the exact perturbation equations, this method remains both computationally efficient and straightforward, facilitating predictions that can be tested against future high-precision weak lensing surveys. Additionally, this framework is highly flexible, allowing for extensions to other modified gravity theories, dynamical dark energy models, and alternative background expansion histories.

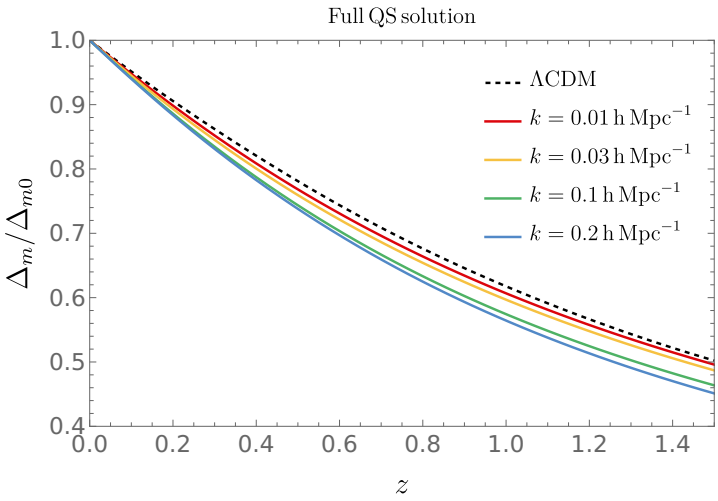
[7] See MacDevette et al. (2025) [37].



(a)

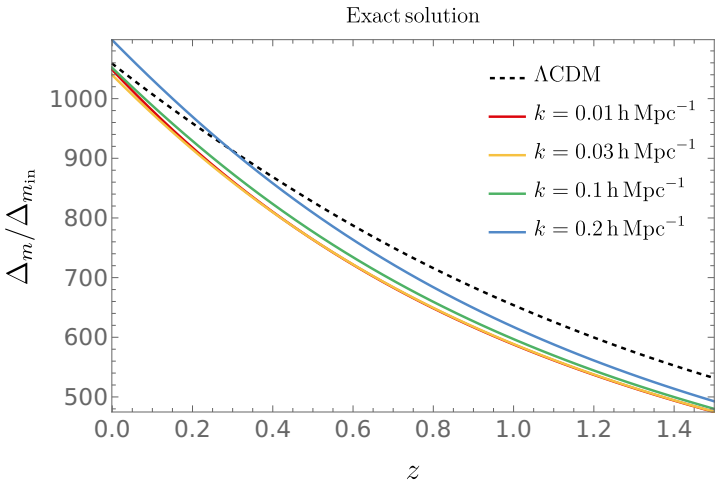


(b)

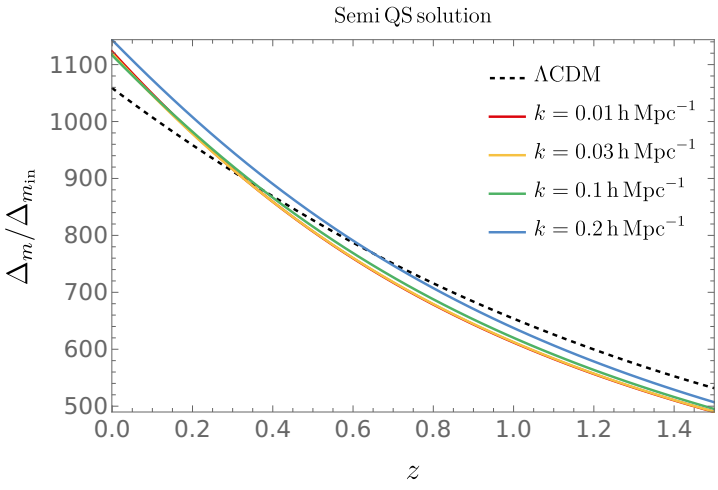


(c)

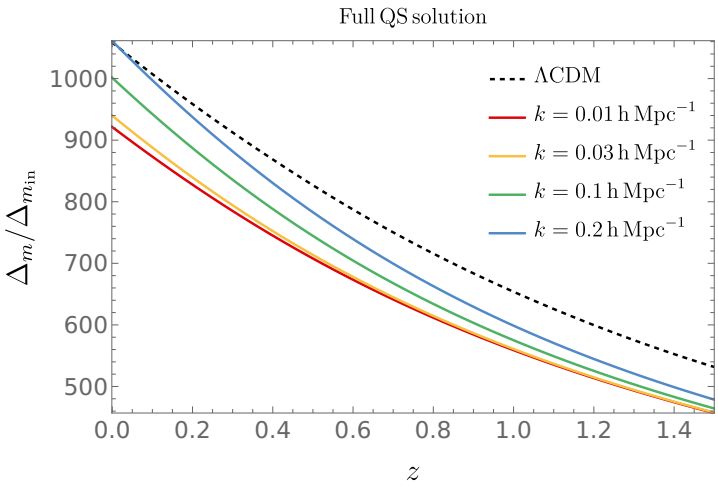
Figure 4.8: Evolution of the matter density perturbation Δ_m^k on sub-horizon scales for different values of k normalised at $z = 0$ using (a) the exact covariant method, (b) the semi-QS approximation, and (c) the full QS approximation. The evolution of the perturbations in the ΛCDM background with the same initial conditions are indicated by the black dashed lines.



(a)



(b)



(c)

Figure 4.9: Evolution of the matter density perturbation Δ_m^k on sub-horizon scales for different values of k normalised at $z = 2000$ using (a) the exact covariant method, (b) the semi-QS approximation, and (c) the full QS approximation. The evolution of the perturbations in the ΛCDM background with the same initial conditions are indicated by the black dashed lines.

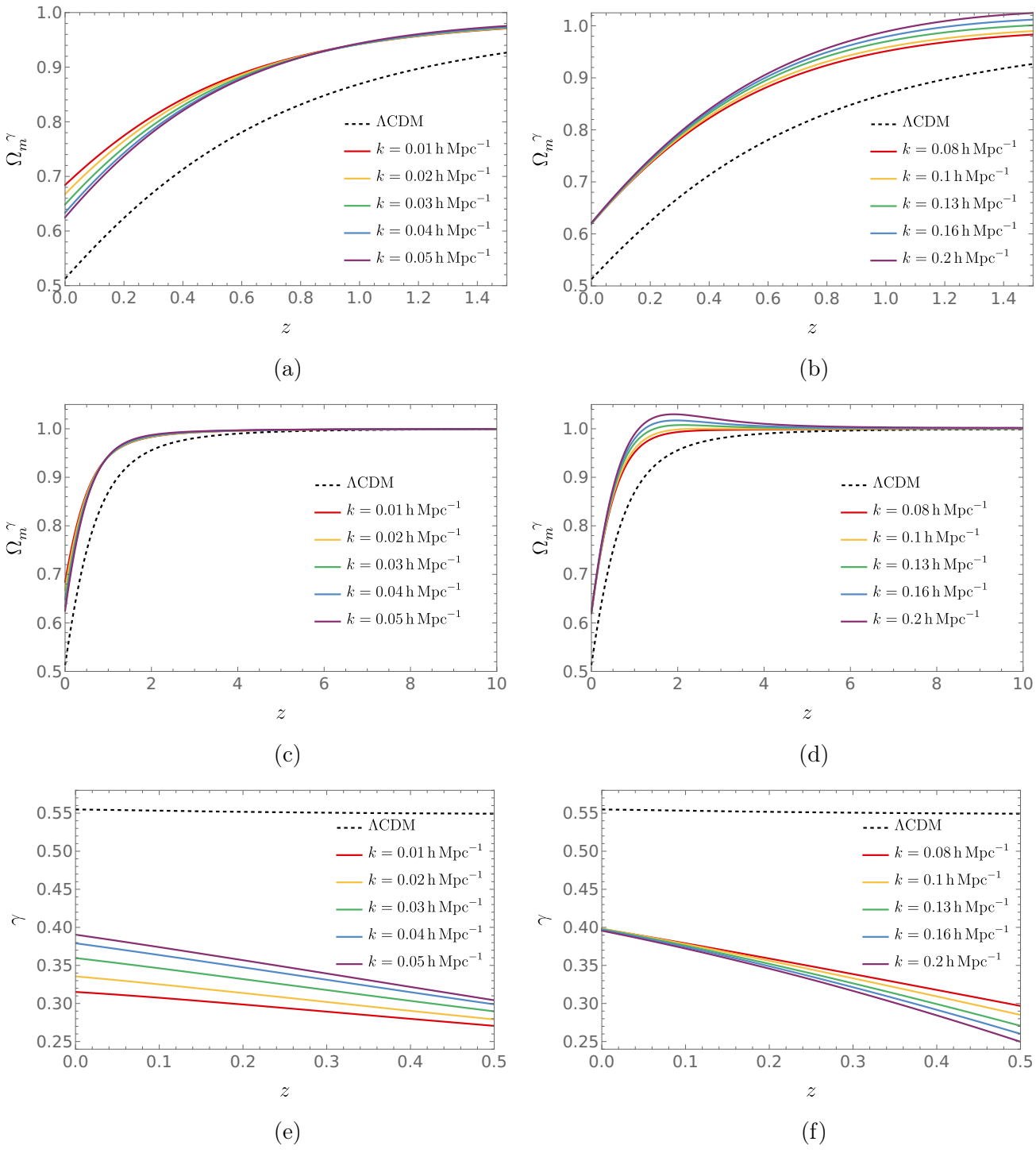


Figure 4.10: (a-d) Growth function S and (e-f) growth index γ for the exact covariant perturbation method. The left column is large scale; the right column is small scale. Figures (a,b) are for a small redshift range ($0 < z < 1.5$) and the figures (c,d) are for a larger range ($0 < z < 10$), clearly showing that in the far past the scale dependence is suppressed as the growth functions for all modes converge to the ΛCDM value.

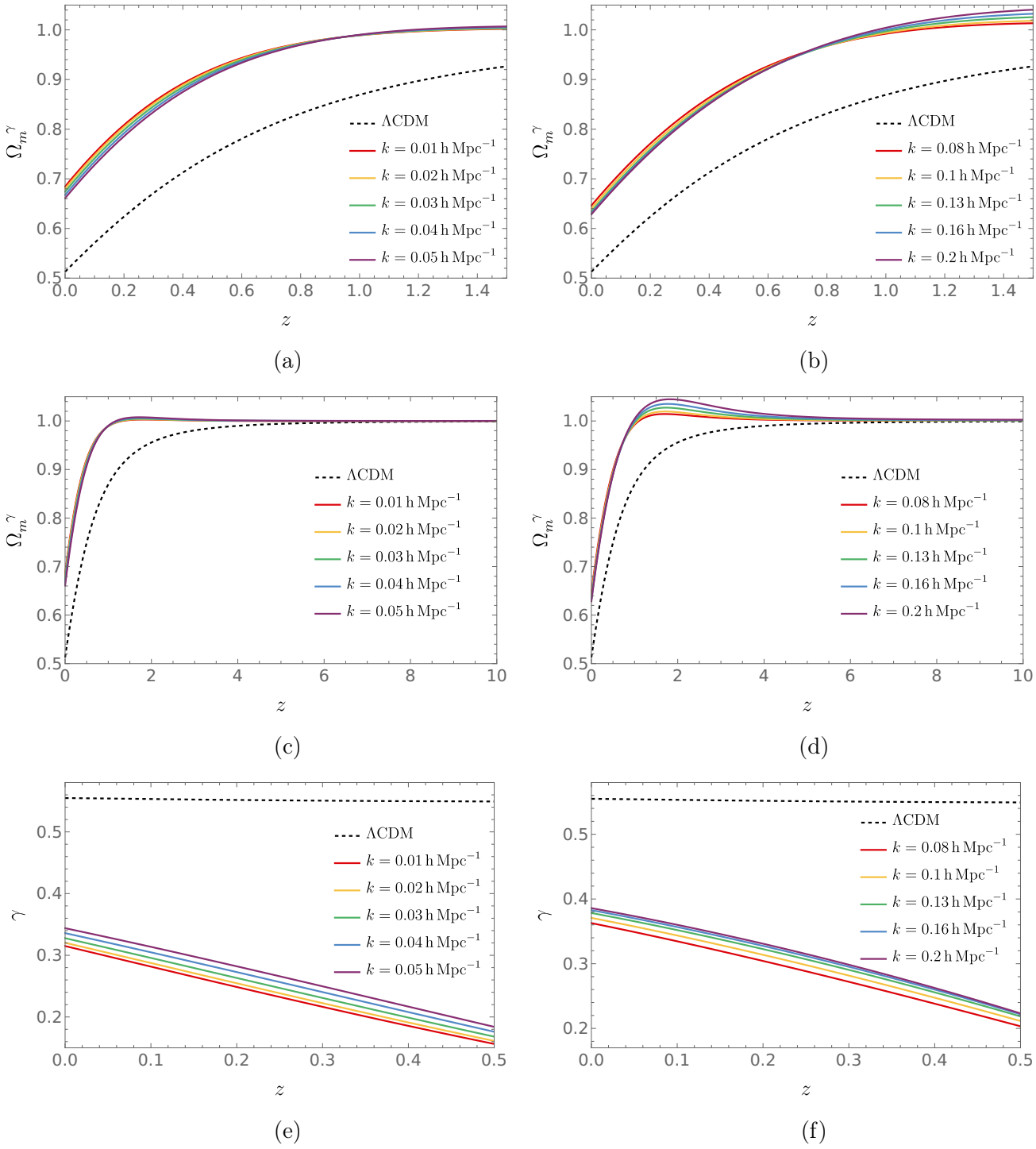


Figure 4.11: (a-d) Growth function S and (e-f) growth index γ for the semi-QS covariant perturbation method. The left column is large scale; the right column is small scale. Figures (a,b) are for a small redshift range ($0 < z < 1.5$) and the figures (c,d) are for a larger range ($0 < z < 10$).

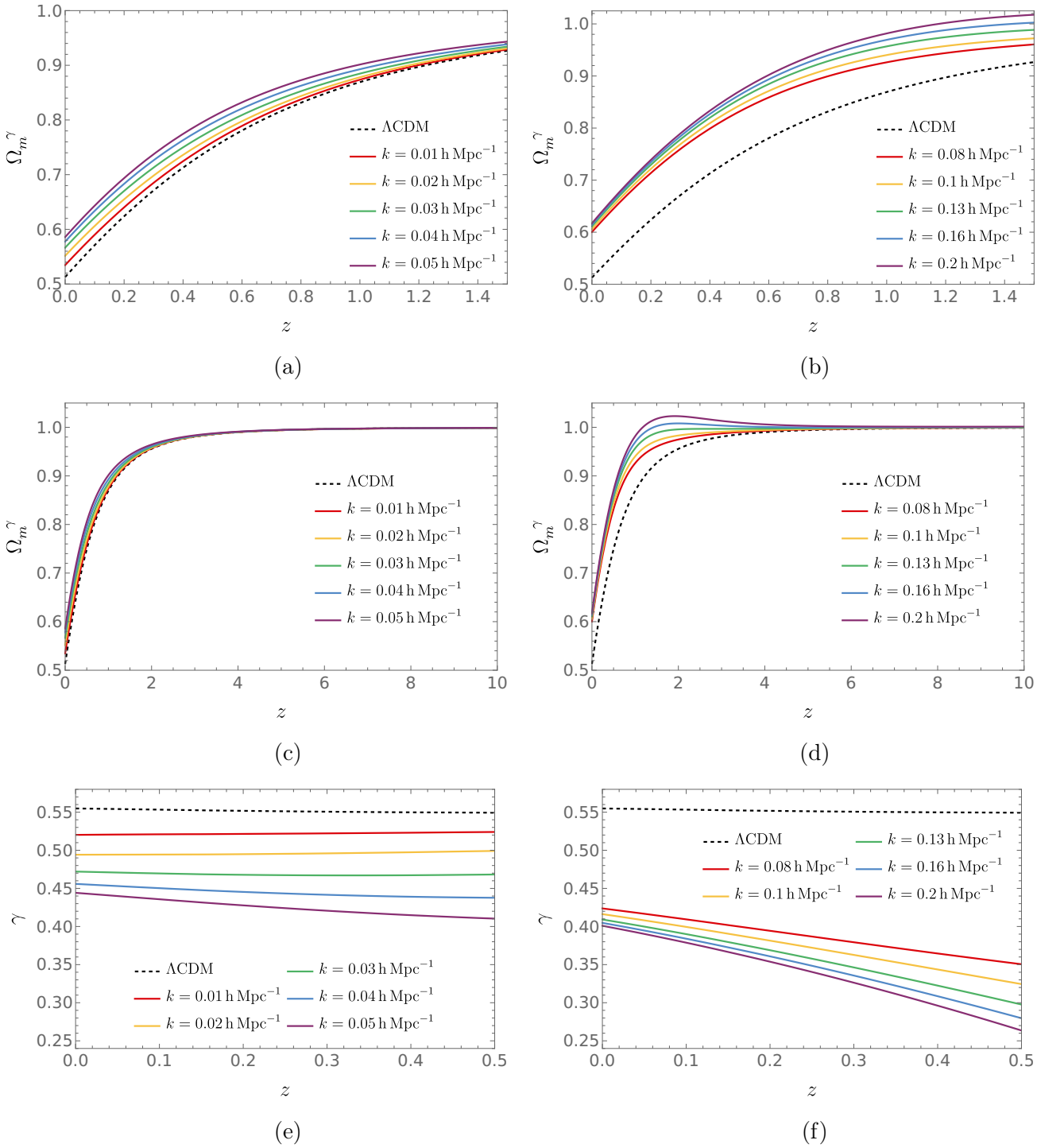
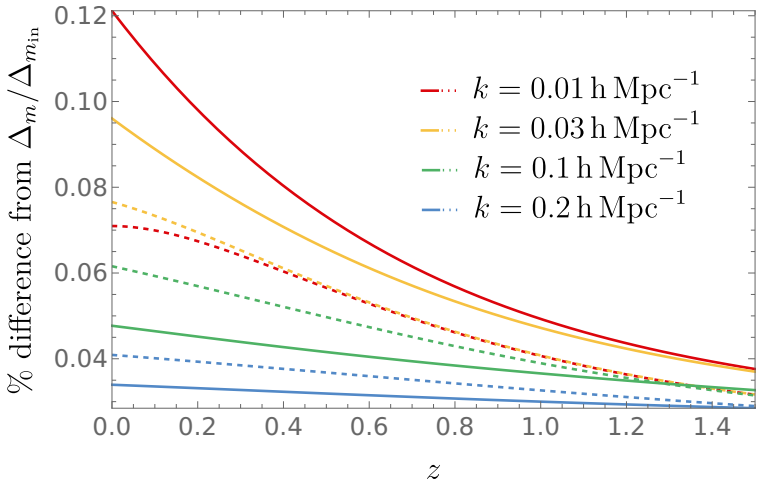
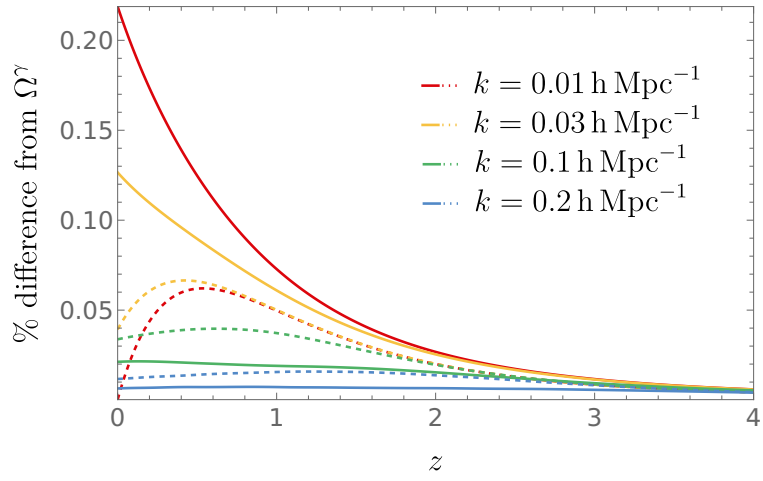


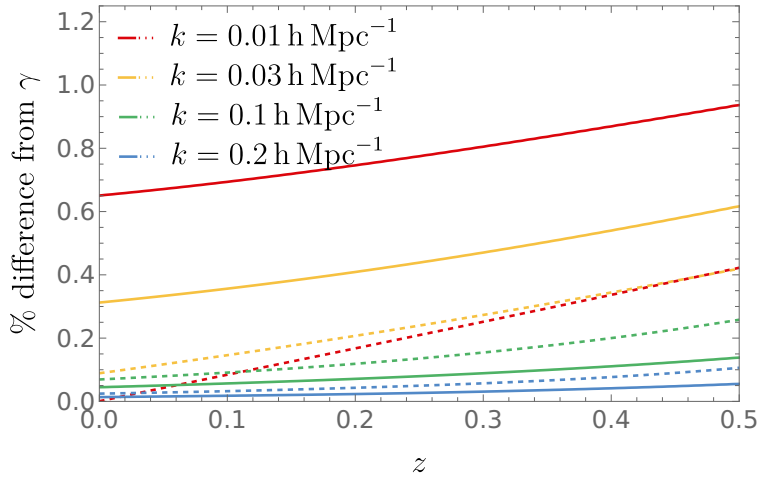
Figure 4.12: (a-d) Growth function S and (e-f) growth index γ for the fully quasi-static perturbation method. The left column is large scale; the right column is small scale. Figures (a,b) are for a small redshift range ($0 < z < 1.5$) and the figures (c,d) are for a larger range ($0 < z < 10$).



(a)

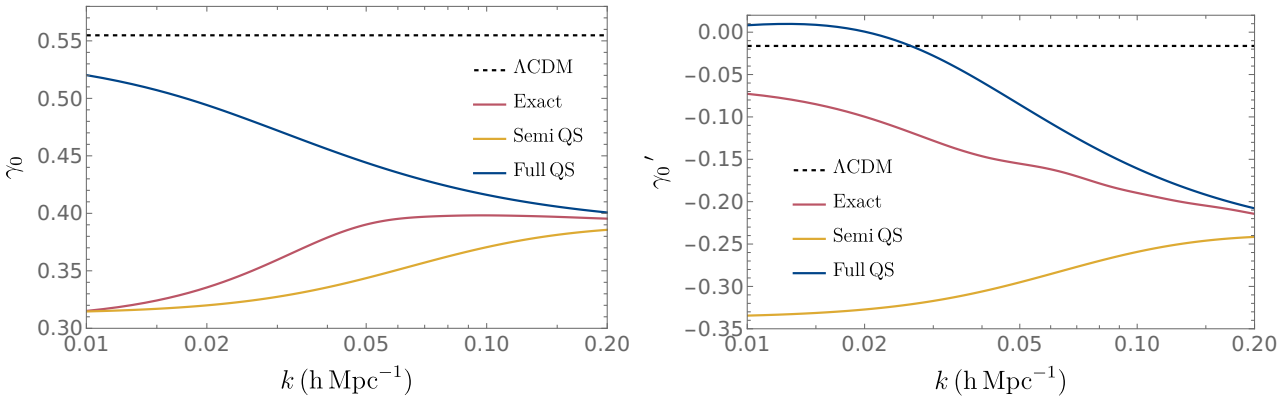


(b)



(c)

Figure 4.13: Percentage deviation between exact solutions and semi QS (dashed) and full QS (solid) solutions for (a) the normalised perturbation $\Delta_m/\Delta_{m_{\rm in}}$, (b) the growth function Ω^γ, and (c) the growth index γ.



(a) (b)

Figure 4.14: Scale dependence of (a) γ_0 and (b) γ_0' for the exact (pink), semi-QS (yellow), and fully QS (blue) methods. The black dashed line represents the value in ΛCDM.

The appeal of modified gravity theories is their ability to address the late-time acceleration issues of ΛCDM by modifying the underlying theory of general relativity. Our objective here was to analyse $f(R)$ gravity in the context of large-scale structure formation and determine if its evolution was viable. We introduced a novel dynamical systems framework for studying $f(R)$ cosmology, avoiding the need for a reconstruction programme by leveraging cosmographic parameters to formulate the theory in a model-independent way. By employing these parameters as dynamical variables, this method imposes algebraic constraints on the phase space that can be fixed by observational data. Unlike previous formulations, which require specifying a functional form for $f(R)$ to close the dynamical equations, this approach presents an autonomous dynamical system for $f(R)$ gravity that does not depend on a specific model.

Examining the phase space properties and the growth of matter perturbations reveals that models closely resembling ΛCDM are less susceptible to future Dolgov-Kawasaki instabilities. Phase space analysis also enables the study of prominent $f(R)$ theories like those by Hu and Sawicki [174], offering broader insights into their background dynamics and parameter spaces. Projections of the phase portrait on the planes $q = -1$, $q = 1/2$, $\Omega = 0$, and $\Omega = 1$ show that trajectories which exit the regions where $F, F' > 0$ more rapidly are associated with cosmic evolutions that are more likely to encounter the Dolgov-Kawasaki instability, rendering them less physically viable. The trajectories that move toward the accelerating submanifold $q = -1$ represent plausible evolutionary paths transitioning from a decelerating, matter-dominated phase to a late-time accelerated expansion.

The findings emphasise the importance of initial conditions. For instance, matching ΛCDM cosmology parameters at $z = 0$ can lead to high redshift deviations. However, initialising the model to align with ΛCDM at high redshifts and evolving toward $z = 0$ produces viable cosmological evolutions. The dynamical system is highly sensitive to initial conditions, often requiring precision to the fourth decimal place or beyond. If the initial conditions are inaccurately set to $z = 0$, the system may fail to reproduce the matter-dominated era and instead transition prematurely to a scalaron-dominated phase for $z \gtrsim 2$. To achieve an evolution consistent with ΛCDM, it is crucial to specify the initial conditions deep within the matter-dominated era, ensuring a proper trajectory through the cosmological history. This approach also identifies the potential for curvature singularities, where cosmographic parameters diverge at finite redshifts, emphasising the need for careful parameter selection.

By plotting the redshift at which the transition to the $f(R)$ regime occurs, we showed that smaller scales transition earlier, while larger scales either transition near $z = 0$ or remain in the GR regime. In the latter case, these modes are indistinguishable from ΛCDM, as they never leave the GR regime. The redshift of this transition depends explicitly on the evolution of x, meaning the choice of the initial condition $x_{\rm in}$ directly impacts which modes have entered the $f(R)$ regime by the present time. For larger values of $|x_{\rm in}|$, such as $|x_{\rm in}| \sim 0.003$, all galaxy power spectrum-relevant modes ($0.01\,h\,\mathrm{Mpc}^{-1} \leq k \leq 0.2\,h\,\mathrm{Mpc}^{-1}$) transition into the $f(R)$ regime by today. Conversely, for smaller values of $|x_{\rm in}|$, these modes remain within the GR regime, rendering them observationally identical to ΛCDM. Setting $x_{\rm in} = -0.001$ ensures that most relevant perturbation modes transition into the $f(R)$ regime, enabling the $f(R)$ modifications to have an observable effect.

We analysed the matter perturbation evolution for the background trajectory using different methods, highlighting distinct scale dependencies and their implications for $f(R)$ cosmology. In one approach, Δ_m was normalised at $z = 0$ to align with prior model-dependent studies, such as those of Starobinsky's model. However, this normalisation obscures scale dependence at lower redshifts, prompting the use of an alternative normalisation at $z = 2000$, which better captures scale dependence at $z \approx 0$. Across all scales of interest, the full covariant and semi-QS methods show minimal scale dependence at low redshifts, while the full QS method exhibits more pronounced differences. Notably, smaller scales display subtle scale dependence in the growth rate at $z = 0$, becoming more significant around $z \sim 1.5$, as these scales transition into the $f(R)$ regime by $z \simeq 1$. Larger scales, however, show stronger scale dependence at $z = 0$, with the full QS method struggling to differentiate $f(R)$ from ΛCDM at these scales. Quantitative comparisons revealed that QS-based solutions deviate more from exact solutions at larger scales, with the full QS approximation aligning better at smaller scales, while the semi-QS method performs more reliably at larger scales. These findings underscore the necessity of applying QS approximations carefully, as highlighted by model-specific studies.

The growth index γ and its evolution offer critical insights. Our methods yield γ_0 values between $0.3 \lesssim \gamma_0 \lesssim 0.4$ and γ_0' values in the range $-0.2 \lesssim \gamma_0' \lesssim -0.07$, consistent with previous studies. Both parameters show significant dispersion across scales, with deviations in γ_0 more prominent for semi-QS approximations at small scales and in γ_0' at all scales. This scale-dependent behaviour and time variability in γ could serve as distinctive markers of $f(R)$ gravity in future precision cosmological observations.

4.5 Summary

This approach promises a more comprehensive exploration of $f(R)$ gravity's parameter space and its role in large-scale structure formation. It has the potential to be used for comparing inflationary models producing scale-invariant power spectra, and extending dynamical analyses to other modified gravity theories such as scalar-tensor models. However, as mentioned previously, the high sensitivity of the parameters means that we do encounter the issue of fine-tuning. It would be advantageous to refine the method and include symmetries or mechanisms that naturally lead to observed parameter values without extensive tuning. We could also employ Bayesian techniques to explore the parameter space systematically and evaluate how well the model performs with fewer fine-tuned parameters, introducing priors that penalise unlikely or overly precise parameter combinations. This would be complemented by sensitivity analyses to identify which parameters significantly impact the observations.

Chapter 5

Conclusion

This thesis has explored the theoretical underpinnings, cosmological applications, and challenges of classical gravity, and endeavoured to investigate modified gravity as an alternative to the standard model. Starting from the foundations of general relativity, we developed the tools required to model the Universe's large-scale evolution, and used these to examine both the successes and shortcomings of the ΛCDM model. A key aim was to investigate whether $f(R)$ gravity can explain cosmic acceleration and structure formation without invoking dark energy or exotic matter.

We began with the foundations of general relativity, the cornerstone of modern gravitational theory. After introducing the curvature of spacetime via differential geometry, we derived the Einstein field equations from the Einstein-Hilbert action. The cosmological principle of large-scale homogeneity and isotropy, supported by observations such as the cosmic microwave background, provided the bedrock for describing the Universe's geometry with the FLRW metric. This metric allowed us to derive key equations, such as the Friedmann, Raychaudhuri, and continuity equations, which describe the Universe's expansion history. Working with these enabled us to examine various cosmic models, including matter-, radiation-, and vacuum-dominated universes. The cosmological constant Λ, inflation, and the interplay of dark matter and dark energy were central to addressing the challenges of the ΛCDM model. However, these idealised models do not account for the Universe's observed inhomogeneities.

To address this limitation we turned to the growth of large-scale structure, focusing on how initial density perturbations evolved into the cosmic web of galaxies and clusters. Perturbation theory provided the tools to describe these overdensities and their evolution, and we examined both the Newtonian and relativistic approaches. From this we derived the Jeans instability criterion and discussed the importance of the matter power spectrum in understanding density fluctuations. These analyses illuminated the mechanisms driving the amplification of perturbations and the resulting structure formation.

Finally, we examined modified gravity, specifically $f(R)$ theories, as a means of addressing unresolved questions in ΛCDM. A novel dynamical systems framework was developed to model $f(R)$ cosmology without requiring a specific functional form for $f(R)$. This method enabled a model-independent exploration of the phase space, using cosmographic parameters constrained by observational data. We found that a narrow subset of phase trajectories achieved the conditions necessary for avoiding instabilities and viable cosmic evolutions consistent with observed large-scale structures. The transition of perturbation modes from general relativity to the $f(R)$ regime was found to be highly sensitive to initial conditions, requiring precise parameter tuning to reproduce a ΛCDM-like evolution. Observables like the growth index γ and its scale-dependent deviations offer concrete avenues for distinguishing $f(R)$ gravity from ΛCDM using new cosmological data.

While promising, $f(R)$ theories are not without their challenges. The viability of the solutions is highly sensitive to initial conditions and requires significant fine-tuning to reproduce observational data. Bayesian techniques and sensitivity analyses could systematically explore the parameter space and potentially reduce this reliance on precise initial conditions. Additionally, the perturbation analysis was restricted to the linear regime; this may not capture the full range of physical behaviours, particularly at large scales or during transition epochs. These caveats highlight the need for a more complete numerical treatment and tighter observational constraints. Future work could involve combining the dynamical framework laid out in the thesis with numerical simulations and parameter inference techniques, such as MCMC, to better constrain viable forms of modified gravity in light of upcoming high-precision surveys.

Modified gravity offers a unified explanation for the observed evolution of our universe, accounting for various phenomena without invoking exotic matter. It is not without its flaws, and requires rigorous observational testing to constrain its myriad models, but with these elegant mathematical tools we may be able to circumvent the need for the ever-elusive dark matter and dark energy. Regardless of the final verdict, solving this mystery would bring clarity to one of the most puzzling dilemmas in modern physics.

Bibliography

[1] Arkani-Hamed, N., Motl, L., Nicolis, A. & Vafa, C. The string landscape, black holes and gravity as the weakest force. *Journal of High Energy Physics* **2007**, 060 (2007). URL https://dx.doi.org/10.1088/1126-6708/2007/06/060. (Cited on page 4.)

[2] Aristoteles & Guthrie, W. K. C. *Aristotle. 6: On the Heavens / with an engl. transl. by W. K. C. Guthrie.* No. 338 in The Loeb classical library (Harvard Univ. Press, Cambridge, Mass., 2006), repr edn. (Cited on page 4.)

[3] Galilei, G. *Discorsi e dimostrazioni matematiche intorno a due nuove scienze attenenti alla mecanica ed i movimenti locali* (Gli Elsevirii, 1638). Google-Books-ID: zB5aAAAAcAAJ. (Cited on page 4.)

[4] Newton, I., Halley, E., Streater, J., Wheldon & Wesley, f. o. M. & Riviere & Son, b. *Philosophiae naturalis principia mathematica* (Londini : Jussu Societatis Regiae ac typis Josephi Streater. Prostat apud plures bibliopolas, 1687). URL http://archive.org/details/philosophiaenatu00newt_0. (Cited on page 4.)

[5] Einstein, A. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften* 844–847 (1915). URL https://articles.adsabs.harvard.edu/pdf/1915SPAW.....844E. (Cited on pages 4 and 12.)

[6] Einstein, A. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik* **354**, 769–822 (1916). URL https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19163540702. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19163540702. (Cited on pages 5 and 12.)

[7] Einstein, A. Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes. *Annalen der Physik* **340**, 898–908 (1911). URL https://ui.adsabs.harvard.edu/abs/1911AnP...340..898E. ADS Bibcode: 1911AnP...340..898E. (Cited on page 5.)

[8] Pound, R. V. & Rebka, G. A. Gravitational Red-Shift in Nuclear Resonance. *Physical Review Letters* **3**, 439–441 (1959). URL https://link.aps.org/doi/10.1103/PhysRevLett.3.439. Publisher: American Physical Society. (Cited on pages 5 and 10.)

[9] Hubble, E. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences* **15**, 168–173 (1929). URL https://www.pnas.org/doi/full/10.1073/pnas.15.3.168. Publisher: Proceedings of the National Academy of Sciences. (Cited on pages 5 and 37.)

[10] Hafele, J. C. & Keating, R. E. Around-the-World Atomic Clocks: Predicted Relativistic Time Gains. *Science* **177**, 166–168 (1972). URL https://www.science.org/doi/10.1126/science.177.4044.166. Publisher: American Association for the Advancement of Science. (Cited on page 5.)

[11] Abbott, T. M. C. *et al.* The Dark Energy Survey Data Release 2. *The Astrophysical Journal Supplement Series* **255**, 20 (2021). URL https://iopscience.iop.org/article/10.3847/1538-4365/ac00b3. (Cited on pages 5 and 49.)

[12] Friedman, A. Über die Krümmung des Raumes. *Zeitschrift für Physik* **10**, 377–386 (1922). URL https://doi.org/10.1007/BF01332580. (Cited on page 5.)

[13] Friedmann, A. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift für Physik* **21**, 326–332 (1924). URL https://doi.org/10.1007/BF01328280. (Cited on page 5.)

[14] Perlmutter, S. *et al.* Measurements of Ω and Λ from 42 High-Redshift Supernovae. *The Astrophysical Journal* **517**, 565–586 (1999). URL https://iopscience.iop.org/article/10.1086/307221. (Cited on pages 5, 38, and 53.)

[15] Hawking, S. W., Penrose, R. & Bondi, H. The singularities of gravitational collapse and cosmology. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **314**, 529–548 (1997). URL https://royalsocietypublishing.org/doi/10.1098/rspa.1970.0021. Publisher: Royal Society. (Cited on pages 5 and 51.)

[16] Guth, A. H. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D* **23**, 347–356 (1981). URL https://link.aps.org/doi/10.1103/PhysRevD.23.347. Publisher: American Physical Society. (Cited on pages 5, 42, and 46.)

[17] Hawking, S. W. & Ellis, G. F. R. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, 1973). URL https://www.cambridge.org/core/books/large-scale-structure-of-spacetime/1E6B961EC9878EDDBBD6AC0AF031CC93. (Cited on page 5.)

[18] Linde, A. D. A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B* **108**, 389–393 (1982). URL https://www.sciencedirect.com/science/article/pii/0370269382912199. (Cited on pages 5, 42, and 46.)

[19] Burgess, C. P. & Moore, G. D. *The standard model: a primer* (Cambridge Univ. Press, Cambridge, 2013), revised 1. ed edn. (Cited on page 5.)

[20] Carlip, S. Quantum gravity: a progress report. *Reports on Progress in Physics* **64**, 885 (2001). URL https://dx.doi.org/10.1088/0034-4885/64/8/301. (Cited on page 5.)

[21] Liddle, A. R. & Lyth, D. H. *Cosmological Inflation and Large-Scale Structure* (Cambridge University Press, Cambridge, 2000). URL https://www.cambridge.org/core/books/cosmological-inflation-and-largescale-structure/52695A7D6FD3BE61F02BDA896EE2C733. (Cited on pages 5, 46, and 47.)

[22] Zwicky, F. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta* **6**, 110–127 (1933). URL https://ui.adsabs.harvard.edu/abs/1933AcHPh...6..110Z. ADS Bibcode: 1933AcHPh...6..110Z. (Cited on page 5.)

[23] Rubin, V. C. & Ford, W. K., Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *The Astrophysical Journal* **159**, 379 (1970). URL https://ui.adsabs.harvard.edu/abs/1970ApJ...159..379R. Publisher: IOP ADS Bibcode: 1970ApJ...159..379R. (Cited on page 5.)

[24] Aghanim, N. *et al.* Planck 2018 results - I. Overview and the cosmological legacy of Planck. *Astronomy & Astrophysics* **641**, A1 (2020). URL https://www.aanda.org/articles/aa/abs/2020/09/aa33880-18/aa33880-18.html. Publisher: EDP Sciences. (Cited on pages 5, 6, 33, 39, 55, 97, and 108.)

[25] Hinshaw, G. *et al.* Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *The Astrophysical Journal Supplement Series* **208**, 19 (2013). URL https://dx.doi.org/10.1088/0067-0049/208/2/19. Publisher: The American Astronomical Society. (Cited on page 5.)

[26] Capozziello, S. & De Laurentis, M. Extended Theories of Gravity. *Physics Reports* **509**, 167 (2011). (Cited on pages 6, 7, 25, and 26.)

[27] Dicke, R. H. Dirac's Cosmology and Mach's Principle. *Nature* **192**, 440–441 (1961). URL https://www.nature.com/articles/192440a0. Publisher: Nature Publishing Group. (Cited on pages 6 and 55.)

[28] Brans, C. & Dicke, R. H. Mach's Principle and a Relativistic Theory of Gravitation. *Physical Review* **124**, 925–935 (1961). URL https://link.aps.org/doi/10.1103/PhysRev.124.925. Publisher: American Physical Society. (Cited on pages 6 and 25.)

[29] Al Mamon, A., Paliathanasis, A. & Saha, S. An extended analysis for a generalized Chaplygin gas model. *The European Physical Journal C* **82**, 232 (2022). URL https://doi.org/10.1140/epjc/s10052-022-10185-4. (Cited on page 6.)

[30] Cardone, V. F., Troisi, A. & Capozziello, S. Unified dark energy models: A phenomenological approach. *Physical Review D* **69**, 083517 (2004). URL https://link.aps.org/doi/10.1103/PhysRevD.69.083517. Publisher: American Physical Society. (Cited on page 6.)

[31] Saridakis, E. N. *et al.* (eds.) *Modified Gravity and Cosmology: An Update by the CANTATA Network* (Springer International Publishing, Cham, 2021). URL https://link.springer.com/10.1007/978-3-030-83715-0. (Cited on pages 6, 25, 26, and 79.)

[32] Rovelli, C. *Quantum gravity*. Cambridge monographs on mathematical physics (Cambridge University Press, Cambridge, UK; New York, 2004). (Cited on page 7.)

[33] Capozziello, S. & Francaviglia, M. Extended Theories of Gravity and their Cosmological and Astrophysical Applications. *General Relativity and Gravitation* **40**, 357–420 (2008). URL http://arxiv.org/abs/0706.1146. ArXiv:0706.1146 [astro-ph, physics:gr-qc]. (Cited on page 7.)

[34] Nojiri, S., Odintsov, S. D. & Oikonomou, V. K. Modified gravity theories on a nutshell: Inflation, bounce and late-time evolution. *Physics Reports* **692**, 1–104 (2017). URL https://www.sciencedirect.com/science/article/pii/S0370157317301527. (Cited on page 7.)

[35] Yashiki, M., Sakai, N. & Saito, R. Local-gravity test of unified models of inflation and dark energy in $f(R)$ gravity. *Physical Review D* **102**, 043504 (2020). URL https://link.aps.org/doi/10.1103/PhysRevD.102.043504. Publisher: American Physical Society. (Cited on page 7.)

[36] Chakraborty, S., MacDevette, K. & Dunsby, P. A model independent approach to the study of $f(R)$ cosmologies with expansion histories close to ΛCDM. *Physical Review D* **103**, 124040 (2021). URL http://arxiv.org/abs/2103.02274. ArXiv:2103.02274 [astro-ph, physics:gr-qc]. (Cited on pages 7, 76, 80, and 84.)

[37] MacDevette, K., Worsley, J., Dunsby, P. & Chakraborty, S. A model-independent approach to the study of structure growth in $f(R)$ gravity. *Monthly Notices of the Royal Astronomical Society* **537**, 2471–2495 (2025). URL https://doi.org/10.1093/mnras/staf168. (Cited on pages 7, 80, and 108.)

[38] Gannouji, R., Moraes, B. & Polarski, D. The growth of matter perturbations in $f(R)$ models. *Journal of Cosmology and Astroparticle Physics* **2009**, 034–034 (2009). URL https://iopscience.iop.org/article/10.1088/1475-7516/2009/02/034. (Cited on pages 7, 58, 97, 107, and 108.)

[39] Einstein, A. Zur Elektrodynamik bewegter Körper. *Annalen der Physik* **322**, 891–921 (1905). URL https://ui.adsabs.harvard.edu/abs/1905AnP...322..891E. ADS Bibcode: 1905AnP...322..891E. (Cited on page 9.)

[40] Schutz, B. F. *A first course in general relativity* (Cambridge University Press, Cambridge; New York, 2009), 2nd ed edn. (Cited on pages 10, 11, 18, and 20.)

[41] Roos, M. Expansion of the Universe - Standard Big Bang Model. *UNESCO Encyclopedia of Life Supporting Systems / Earth and Atmospheric Sciences* **5**, 1 (2008). (Cited on page 11.)

[42] Peter, P. & Uzan, J.-P. *Primordial cosmology* (Oxford University Press, Oxford; New York, 2009), english ed edn. OCLC: ocn313665248. (Cited on pages 11, 12, 13, 15, 16, 19, 30, and 62.)

[43] Carroll, S. M. *Spacetime and geometry: an introduction to general relativity* (Pearson Education, Harlow, 2014), pearson new international edn. (Cited on pages 11, 14, 15, 16, 17, 18, 19, 23, 30, 36, 37, 43, and 60.)

[44] Hobson, M. P., Efstathiou, G. & Lasenby, A. N. *General relativity: an introduction for physicists* (Cambridge University Press, Cambridge, UK; New York, 2006). OCLC: ocm61757089. (Cited on pages 12, 19, 20, 28, 30, 40, 44, 46, 48, 72, and 73.)

[45] Emam, M. H. *Covariant Physics: From Classical Mechanics to General Relativity and Beyond* (Oxford University Press, 2021), 1 edn. URL https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198864899.001.0001/oso-9780198864899. (Cited on pages 12, 13, 17, 18, 29, and 51.)

[46] Hamilton, W. R. & Beaufort, F. XV. On a general method in dynamics; by which the study of the motions of all free systems of attracting or repelling points is reduced to the search and differentiation of one central relation, or characteristic function. *Philosophical Transactions of the Royal Society of London* **124**, 247–308 (1997). URL https://royalsocietypublishing.org/doi/10.1098/rstl.1834.0017. Publisher: Royal Society. (Cited on page 21.)

[47] De Felice, A. & Tsujikawa, S. $f(R)$ Theories. *Living Reviews in Relativity* **13**, 3 (2010). (Cited on pages 22, 78, and 79.)

[48] Blau, M. Lecture Notes on General Relativity 992. (Cited on pages 22, 50, and 77.)

[49] Weinberg, S. *Gravitation And Cosmology: Principles And Applications Of The General Theory Of Relativity* (Wiley, 1972). URL http://archive.org/details/WeinbergS.GravitationAndCosmology..PrinciplesAndApplicationsOfTheGeneralTheoryOf. (Cited on page 22.)

[50] Gleick, J. *Isaac Newton* (Pantheon Books, New York, 2003), 1st ed edn. (Cited on page 25.)

[51] Chan, M. H. & Lee, C. M. The solar system test for the general modified gravity theories. *Monthly Notices of the Royal Astronomical Society* **518**, 6238–6242 (2022). URL http://arxiv.org/abs/2212.01019. ArXiv:2212.01019 [astro-ph, physics:gr-qc]. (Cited on page 26.)

[52] Belgacem, E. *et al.* Testing modified gravity at cosmological distances with LISA standard sirens. *Journal of Cosmology and Astroparticle Physics* **2019**, 024 (2019). URL https://dx.doi.org/10.1088/1475-7516/2019/07/024. (Cited on page 26.)

[53] Miao, X., Xu, H., Shao, L., Liu, C. & Ma, B.-Q. Stringent Tests of Gravity with Highly Relativistic Binary Pulsars in the Era of LISA and SKA. *The Astrophysical Journal* **921**, 114 (2021). URL https://dx.doi.org/10.3847/1538-4357/ac1d48. Publisher: The American Astronomical Society. (Cited on page 26.)

[54] Simpson, F. *et al.* CFHTLenS: testing the laws of gravity with tomographic weak lensing and redshift-space distortions. *Monthly Notices of the Royal Astronomical Society* **429**, 2249–2263 (2013). URL https://doi.org/10.1093/mnras/sts493. (Cited on page 26.)

[55] Li, J. & Zhao, G.-B. Cosmological Tests of Gravity with the Latest Observations. *The Astrophysical Journal* **871**, 196 (2019). URL https://dx.doi.org/10.3847/1538-4357/aaf869. Publisher: The American Astronomical Society. (Cited on page 26.)

[56] Capoferri, M. & Vassiliev, D. Spacetime diffeomorphisms as matter fields. *Journal of Mathematical Physics* **61**, 111508 (2020). URL https://pubs.aip.org/jmp/article/61/11/111508/234801/Spacetime-diffeomorphisms-as-matter-fields. (Cited on page 26.)

[57] Abdalla, M. C. B., Nojiri, S. & Odintsov, S. D. Consistent modified gravity: dark energy, acceleration and the absence of cosmic doomsday. *Classical and Quantum Gravity* **22**, L35 (2005). URL https://dx.doi.org/10.1088/0264-9381/22/5/L01. (Cited on page 26.)

[58] Nojiri, S. & Odintsov, S. D. Unified cosmic history in modified gravity: from $f(R)$ theory to Lorentz non-invariant models. *Physics Report* **505**, 59 (2011). (Cited on page 26.)

[59] Briscese, F., Elizalde, E., Nojiri, S. & Odintsov, S. D. Phantom scalar dark energy as modified gravity: Understanding the origin of the Big Rip singularity. *Physics Letters B* **646**, 105–111 (2007). URL https://www.sciencedirect.com/science/article/pii/S0370269307000925. (Cited on page 26.)

[60] Ellis, G. F. R., Maartens, R. & MacCallum, M. A. H. *Relativistic cosmology* (Cambridge University Press, Cambridge, 2012), first edn. (Cited on pages 28 and 62.)

[61] Uzan, J.-P., Clarkson, C. & Ellis, G. F. R. Time drift of cosmological redshifts as a test of the Copernican principle. *Physical Review Letters* **100**, 191303 (2008). URL http://arxiv.org/abs/0801.0068. ArXiv:0801.0068 [astro-ph, physics:gr-qc]. (Cited on page 28.)

[62] Sandage, A. The Change of Redshift and Apparent Luminosity of Galaxies due to the Deceleration of Selected Expanding Universes. *The Astrophysical Journal* **136**, 319 (1962). URL https://ui.adsabs.harvard.edu/abs/1962ApJ...136..319S. Publisher: IOP ADS Bibcode: 1962ApJ...136..319S. (Cited on page 28.)

[63] McVittie, G. C. Appendix to The Change of Redshift and Apparent Luminosity of Galaxies due to the Deceleration of Selected Expanding Universes. *The Astrophysical Journal* **136**, 334 (1962). URL https://ui.adsabs.harvard.edu/abs/1962ApJ...136..334M. Publisher: IOP ADS Bibcode: 1962ApJ...136..334M. (Cited on page 28.)

[64] Amendola, L. & Tsujikawa, S. *Dark Energy: Theory and Observations* (Cambridge University Press, Cambridge, 2010). URL https://www.cambridge.org/core/books/dark-energy/EC55E8BF946C34D61B758273D8286618. (Cited on pages 28, 76, and 79.)

[65] Rocha, B. A. R. & Martins, C. J. A. P. Redshift drift cosmography with ELT and SKAO measurements. *Monthly Notices of the Royal Astronomical Society* **518**, 2853–2869 (2023). URL https://doi.org/10.1093/mnras/stac3240. (Cited on page 28.)

[66] Coles, P. & Lucchin, F. *Cosmology: the origin and evolution of cosmic structure* (John Wiley, Chichester, Eng, 2002), 2nd ed edn. (Cited on pages 28, 37, 41, 42, 46, 51, 54, 59, 73, and 74.)

[67] Bennett, C. L. *et al.* First-Year Wilkinson Microwave Anisotropy Probe(WMAP) Observations: Preliminary Maps and Basic Results. *The Astrophysical Journal Supplement Series* **148**, 1 (2003). URL https://iopscience.iop.org/article/10.1086/377253/meta. Publisher: IOP Publishing. (Cited on page 28.)

[68] Tegmark, M. *et al.* The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey. *The Astrophysical Journal* **606**, 702 (2004). URL https://iopscience.iop.org/article/10.1086/382125/meta. Publisher: IOP Publishing. (Cited on page 28.)

[69] Cole, S. *et al.* The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *Monthly Notices of the Royal Astronomical Society* **362**, 505–534 (2005). URL https://doi.org/10.1111/j.1365-2966.2005.09318.x. (Cited on page 28.)

[70] Penzias, A. A. & Wilson, R. W. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *The Astrophysical Journal* **142**, 419–421 (1965). URL https://ui.adsabs.harvard.edu/abs/1965ApJ...142..419P. Publisher: IOP ADS Bibcode: 1965ApJ...142..419P. (Cited on page 28.)

[71] Dicke, R. H., Peebles, P. J. E., Roll, P. G. & Wilkinson, D. T. Cosmic Black-Body Radiation. *The Astrophysical Journal* **142**, 414–419 (1965). URL https://ui.adsabs.harvard.edu/abs/1965ApJ...142..414D. Publisher: IOP ADS Bibcode: 1965ApJ...142..414D. (Cited on page 28.)

[72] Fixsen, D. J. The temperature of the cosmic microwave background. *The Astrophysical Journal* **707**, 916 (2009). URL https://dx.doi.org/10.1088/0004-637X/707/2/916. Publisher: The American Astronomical Society. (Cited on page 28.)

[73] Mather, J. C. *et al.* A Preliminary Measurement of the Cosmic Microwave Background Spectrum by the Cosmic Background Explorer (COBE) Satellite. *The Astrophysical Journal* **354**, L37 (1990). URL https://ui.adsabs.harvard.edu/abs/1990ApJ...354L..37M. Publisher: IOP ADS Bibcode: 1990ApJ...354L..37M. (Cited on page 29.)

[74] Smoot, G. F. *et al.* Structure in the COBE Differential Microwave Radiometer First-Year Maps. *The Astrophysical Journal* **396**, L1 (1992). URL https://ui.adsabs.harvard.edu/abs/1992ApJ...396L...1S. Publisher: IOP ADS Bibcode: 1992ApJ...396L...1S. (Cited on page 29.)

[75] Giacconi, R., Gursky, H., Paolini, F. R. & Rossi, B. B. Evidence for X-Rays From Sources Outside the Solar System. *Physical Review Letters* **9**, 439–443 (1962). URL https://link.aps.org/doi/10.1103/PhysRevLett.9.439. Publisher: American Physical Society. (Cited on page 29.)

[76] Weisskopf, M. C. *et al.* An Overview of the Performance and Scientific Results from the Chandra X-Ray Observatory. *Publications of the Astronomical Society of the Pacific* **114**, 1 (2002). URL https://dx.doi.org/10.1086/338108. Publisher: The University of Chicago Press. (Cited on page 29.)

[77] Skrutskie, M. F. *et al.* The Two Micron All Sky Survey (2MASS). *The Astronomical Journal* **131**, 1163 (2006). URL https://dx.doi.org/10.1086/498708. (Cited on page 29.)

[78] Loveday, J. The APM Bright Galaxy Catalogue. *Monthly Notices of the Royal Astronomical Society* **278**, 1025–1048 (1996). URL https://doi.org/10.1093/mnras/278.4.1025. (Cited on page 29.)

[79] Overzier, R. A., Röttgering, H. J. A., Rengelink, R. B. & Wilman, R. J. The spatial clustering of radio sources in NVSS and FIRST; implications for galaxy clustering evolution. *Astronomy & Astrophysics* **405**, 53–72 (2003). URL https://www.aanda.org/articles/aa/abs/2003/25/aa3020/aa3020.html. Number: 1 Publisher: EDP Sciences. (Cited on page 29.)

[80] Zehavi, I. *et al.* Galaxy clustering in the completed SDSS redshift survey: the dependence on color and luminosity. *The Astrophysical Journal* **736**, 59 (2011). URL https://dx.doi.org/10.1088/0004-637X/736/1/59. Publisher: The American Astronomical Society. (Cited on page 29.)

[81] Colless, M. *et al.* The 2dF Galaxy Redshift Survey: spectra and redshifts. *Monthly Notices of the Royal Astronomical Society* **328**, 1039–1063 (2001). URL https://doi.org/10.1046/j.1365-8711.2001.04902.x. (Cited on page 29.)

[82] Newman, J. A. *et al.* The DEEP2 galaxy redshift survey: design, observations, data reduction, and redshifts. *The Astrophysical Journal Supplement Series* **208**, 5 (2013). URL https://dx.doi.org/10.1088/0067-0049/208/1/5. Publisher: The American Astronomical Society. (Cited on page 29.)

[83] Lilly, S. J. *et al.* zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering $0 < z < 3$ in the COSMOS Field. *The Astrophysical Journal Supplement Series* **172**, 70 (2007). URL https://dx.doi.org/10.1086/516589. (Cited on page 29.)

[84] Driver, S. P. *et al.* GAMA: towards a physical understanding of galaxy formation. *Astronomy & Geophysics* **50**, 5.12–5.19 (2009). URL https://doi.org/10.1111/j.1468-4004.2009.50512.x. (Cited on page 29.)

[85] Jones, D. H. *et al.* The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures. *Monthly Notices of the Royal Astronomical Society* **399**, 683–698 (2009). URL https://doi.org/10.1111/j.1365-2966.2009.15338.x. (Cited on page 29.)

[86] Hubble, E. P. *Realm of the Nebulae* (1936). URL https://ui.adsabs.harvard.edu/abs/1936rene.book.....H. Publication Title: Realm of the Nebulae ADS Bibcode: 1936rene.book.....H. (Cited on page 29.)

[87] Hogg, D. W. *et al.* Cosmic Homogeneity Demonstrated with Luminous Red Galaxies. *The Astrophysical Journal* **624**, 54 (2005). URL https://iopscience.iop.org/article/10.1086/429084/meta. Publisher: IOP Publishing. (Cited on page 29.)

[88] Ellis, G. F. R. Issues in the Philosophy of Cosmology (2006). URL http://arxiv.org/abs/astro-ph/0602280. ArXiv:astro-ph/0602280. (Cited on page 29.)

[89] Moresco, M. *et al.* A 6% measurement of the Hubble parameter at $z \sim 0.45$: direct evidence of the epoch of cosmic re-acceleration. *Journal of Cosmology and Astroparticle Physics* **2016**, 014–014 (2016). URL https://iopscience.iop.org/article/10.1088/1475-7516/2016/05/014. (Cited on page 37.)

[90] Farooq, O., Madiyar, F. R., Crandall, S. & Ratra, B. Hubble parameter measurement constraints on the redshift of the deceleration–acceleration transition, dynamical dark energy, and space curvature. *The Astrophysical Journal* **835**, 26 (2017). URL https://iopscience.iop.org/article/10.3847/1538-4357/835/1/26. (Cited on page 37.)

[91] Dunajski, M. & Gibbons, G. Cosmic jerk, snap and beyond. *Classical and Quantum Gravity* **25**, 235012 (2008). URL https://iopscience.iop.org/article/10.1088/0264-9381/25/23/235012. (Cited on pages 38, 83, and 99.)

[92] Dunsby, P. K. S. & Luongo, O. On the theory and applications of modern cosmography. *International Journal of Geometric Methods in Modern Physics* **13**, 1630002 (2016). URL http://arxiv.org/abs/1511.06532. ArXiv:1511.06532 [astro-ph, physics:gr-qc, physics:math-ph]. (Cited on page 38.)

[93] Ryden, B. *Introduction to cosmology* (Cambridge university press, Cambridge, 2017), 2nd ed edn. (Cited on pages 38, 41, 51, 54, and 72.)

[94] Riess, A. G. *et al.* Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *The Astronomical Journal* **116**, 1009–1038 (1998). URL http://arxiv.org/abs/astro-ph/9805201. ArXiv:astro-ph/9805201. (Cited on pages 38 and 53.)

[95] Capozziello, S., Cardone, V. F. & Salzano, V. Cosmography of $f(R)$ gravity. *Physical Review D* **78**, 063504 (2008). URL http://arxiv.org/abs/0802.1583. ArXiv:0802.1583 [astro-ph, physics:gr-qc]. (Cited on page 38.)

[96] Riess, A. G. *et al.* Type Ia Supernova Discoveries at z > 1 from the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution. *The Astrophysical Journal* **607**, 665 (2004). URL https://iopscience.iop.org/article/10.1086/383612/meta. Publisher: IOP Publishing. (Cited on page 38.)

[97] Visser, M. Jerk, snap and the cosmological equation of state. *Classical and Quantum Gravity* **21**, 2603 (2004). URL https://dx.doi.org/10.1088/0264-9381/21/11/006. (Cited on page 38.)

[98] Rapetti, D., Allen, S. W., Amin, M. A. & Blandford, R. D. A kinematical approach to dark energy studies. *Monthly Notices of the Royal Astronomical Society* **375**, 1510–1520 (2007). URL https://doi.org/10.1111/j.1365-2966.2006.11419.x. (Cited on page 38.)

[99] DESI Collaboration *et al.* DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations (2024). URL http://arxiv.org/abs/2404.03002. ArXiv:2404.03002 [astro-ph]. (Cited on pages 39, 42, and 56.)

[100] de Sitter, W. On Einstein's Theory of Gravitation and its Astronomical Consequences. Third Paper.. *Monthly Notices of the Royal Astronomical Society* **78**, 3–28 (1917). URL https://doi.org/10.1093/mnras/78.1.3. (Cited on page 41.)

[101] Ellis, G. F. R. & Maartens, R. The emergent universe: inflationary cosmology with no singularity. *Classical and Quantum Gravity* **21**, 223 (2003). URL https://dx.doi.org/10.1088/0264-9381/21/1/015. (Cited on page 41.)

[102] Einstein, A. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften* 142–152 (1917). URL https://ui.adsabs.harvard.edu/abs/1917SPAW.....142E. ADS Bibcode: 1917SPAW.....142E. (Cited on page 42.)

[103] Kerszberg, P. *The Einstein - de Sitter controversy of 1916 - 1917 and the rise of relativistic cosmology*. (1989). URL https://ui.adsabs.harvard.edu/abs/1989ehgr.conf..325K. Conference Name: Einstein and the History of General Relativity Pages: 325-366 ADS Bibcode: 1989ehgr.conf..325K. (Cited on page 42.)

[104] VandenBerg, D. A., Bolte, M. & Stetson, P. B. The age of the galactic globular cluster system. *Annual Review of Astronomy and Astrophysics* **34**, 461–510 (1996). URL https://www.annualreviews.org/content/journals/10.1146/annurev.astro.34.1.461. Publisher: Annual Reviews. (Cited on page 42.)

[105] Spergel, D. N., Bolte, M. & Freedman, W. The age of the universe. *Proceedings of the National Academy of Sciences* **94**, 6579–6584 (1997). URL https://www.pnas.org/doi/10.1073/pnas.94.13.6579. Publisher: Proceedings of the National Academy of Sciences. (Cited on page 42.)

[106] Peebles, P. J. E. Tests of cosmological models constrained by inflation. *The Astrophysical Journal* **284**, 439 (1984). URL http://adsabs.harvard.edu/doi/10.1086/162425. (Cited on page 42.)

[107] Aghanim, N. *et al.* Planck 2018 results - VI. Cosmological parameters. *Astronomy & Astrophysics* **641**, A6 (2020). URL https://www.aanda.org/articles/aa/abs/2020/09/aa33910-18/aa33910-18.html. Publisher: EDP Sciences. (Cited on pages 42, 49, 55, 72, and 108.)

[108] Martel, H. Galaxy Formation in $\Lambda > 0$ Friedmann Models: Consequences for the Number Counts versus Redshift Test. *The Astrophysical Journal* **421**, L67 (1994). URL https://ui.adsabs.harvard.edu/abs/1994ApJ...421L..67M. Publisher: IOP ADS Bibcode: 1994ApJ...421L..67M. (Cited on page 42.)

[109] Ellis, G. F. R. & Bruni, M. Covariant and gauge-invariant approach to cosmological density fluctuations. *Physical Review D* **40**, 1804–1818 (1989). URL https://link.aps.org/doi/10.1103/PhysRevD.40.1804. (Cited on pages 42, 59, 60, 61, 62, 63, 66, 69, 89, and 92.)

[110] Blumenthal, G. R., Faber, S. M., Primack, J. R. & Rees, M. J. Formation of galaxies and large-scale structure with cold dark matter. *Nature* **311**, 517–525 (1984). URL https://www.nature.com/articles/311517a0. Publisher: Nature Publishing Group. (Cited on pages 42 and 48.)

[111] Davis, M., Efstathiou, G., Frenk, C. S. & White, S. D. M. The evolution of large-scale structure in a universe dominated by cold dark matter. *The Astrophysical Journal* **292**, 371–394

(1985). URL https://ui.adsabs.harvard.edu/abs/1985ApJ...292..371D. Publisher: IOP ADS Bibcode: 1985ApJ...292..371D. (Cited on page 42.)

[112] Starobinsky, A. A. A new type of isotropic cosmological models without singularity. *Physics Letters B* **91**, 99–102 (1980). URL https://www.sciencedirect.com/science/article/pii/037026938090670X. (Cited on pages 42 and 79.)

[113] Ellis, G. F. R. & Rothman, T. Lost horizons. *American Journal of Physics* **61**, 883–893 (1993). URL https://aapt.scitation.org/doi/10.1119/1.17400. Publisher: American Association of Physics Teachers. (Cited on pages 44 and 52.)

[114] Baumann, D. *Cosmology Part III Mathematical Tripos* (DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, 2014). (Cited on pages 44 and 62.)

[115] Debono, I. & Smoot, G. General Relativity and Cosmology: Unsolved Questions and Future Directions. *Universe* **2** (2016). (Cited on pages 45 and 53.)

[116] Weinberg, S. *Cosmology* (Oxford university press, New York, 2008). (Cited on page 48.)

[117] Kolb, E. W. & Turner, M. S. The early Universe. *Nature* **294**, 521–526 (1981). URL https://www.nature.com/articles/294521a0. Publisher: Nature Publishing Group. (Cited on page 51.)

[118] Follin, B., Knox, L., Millea, M. & Pan, Z. First Detection of the Acoustic Oscillation Phase Shift Expected from the Cosmic Neutrino Background. *Physical Review Letters* **115**, 091301 (2015). URL https://link.aps.org/doi/10.1103/PhysRevLett.115.091301. Publisher: American Physical Society. (Cited on page 51.)

[119] Ma, C.-P. & Bertschinger, E. Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. *The Astrophysical Journal* **455**, 7 (1995). URL http://adsabs.harvard.edu/doi/10.1086/176550. (Cited on page 52.)

[120] Schneider, P. *Extragalactic astronomy and cosmology: an introduction* (Springer, Berlin; New York, 2006). OCLC: ocm74270145. (Cited on pages 53 and 54.)

[121] Steinhardt, P. J. 7 cosmological challenges for the 21st century. In Fitch, V. L., Marlow, D. R. & Dementi, M. A. (eds.) *Critical Problems in Physics*, 123–146 (Princeton University Press, 2021). URL https://www.degruyter.com/document/doi/10.1515/9780691227498-008/html. (Cited on page 55.)

[122] Susskind, L. The anthropic landscape of string theory. In Carr, B. (ed.) *Universe or Multiverse?*, 247–266 (Cambridge University Press, Cambridge, 2007). URL https://www.cambridge.org/core/books/universe-or-multiverse/anthropic-landscape-of-string-theory/60C574844DB1E631356BD24AED0CD052. (Cited on page 55.)

[123] Adler, R. J., Casey, B. & Jacob, O. C. Vacuum catastrophe: An elementary exposition of the cosmological constant problem. *American Journal of Physics* **63**, 620–626 (1995). URL https://doi.org/10.1119/1.17850. (Cited on page 55.)

[124] Tada, Y. & Terada, T. Quintessential interpretation of the evolving dark energy in light of DESI observations. *Physical Review D* **109**, L121305 (2024). URL https://link.aps.org/doi/10.1103/PhysRevD.109.L121305. Publisher: American Physical Society. (Cited on page 55.)

[125] Riess, A. G. *et al.* A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km/s/Mpc Uncertainty from the Hubble Space Telescope and the SH0ES Team. *The Astrophysical Journal Letters* **934**, L7 (2022). URL https://dx.doi.org/10.3847/2041-8213/ac5c5b. Publisher: The American Astronomical Society. (Cited on page 55.)

[126] Freedman, W. L. Measurements of the Hubble Constant: Tensions in Perspective. *The Astrophysical Journal* **919**, 16 (2021). URL https://dx.doi.org/10.3847/1538-4357/ac0e95. Publisher: The American Astronomical Society. (Cited on page 55.)

[127] Li, S. & Beaton, R. L. The Tip of the Red Giant Branch Distance Ladder and the Hubble Constant (2024). URL http://arxiv.org/abs/2403.17048. ArXiv:2403.17048. (Cited on page 55.)

[128] Perivolaropoulos, L. & Skara, F. Challenges for ΛCDM: An update. *New Astronomy Reviews* **95**, 101659 (2022). URL https://www.sciencedirect.com/science/article/pii/S1387647322000185. (Cited on page 55.)

[129] Dark Energy Survey Collaboration 1 *et al.* Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing. *Physical Review D* **98**, 043526 (2018). URL https://link.aps.org/doi/10.1103/PhysRevD.98.043526. Publisher: American Physical Society. (Cited on page 55.)

[130] Akrami, Y. *et al.* Planck 2018 results - VII. Isotropy and statistics of the CMB. *Astronomy & Astrophysics* **641**, A7 (2020). URL https://www.aanda.org/articles/aa/abs/2020/09/aa35201-19/aa35201-19.html. Publisher: EDP Sciences. (Cited on page 55.)

[131] Bull, P. *et al.* Beyond ΛCDM: Problems, solutions, and the road ahead. *Physics of the Dark Universe* **12**, 56–99 (2016). URL https://www.sciencedirect.com/science/article/pii/S2212686416300097. (Cited on page 55.)

[132] Klypin, A., Kravtsov, A. V., Valenzuela, O. & Prada, F. Where Are the Missing Galactic Satellites? *The Astrophysical Journal* **522**, 82 (1999). URL https://iopscience.iop.org/article/10.1086/307643/meta. Publisher: IOP Publishing. (Cited on page 56.)

[133] Moore, B. *et al.* Dark Matter Substructure within Galactic Halos. *The Astrophysical Journal* **524**, L19 (1999). URL https://iopscience.iop.org/article/10.1086/312287/meta. Publisher: IOP Publishing. (Cited on page 56.)

[134] Boylan-Kolchin, M., Bullock, J. S. & Kaplinghat, M. Too big to fail? The puzzling darkness of massive Milky Way subhaloes. *Monthly Notices of the Royal Astronomical Society: Letters* **415**, L40–L44 (2011). URL https://doi.org/10.1111/j.1745-3933.2011.01074.x. (Cited on page 56.)

[135] Ostriker, J. P., Choi, E., Chow, A. & Guha, K. Mind the Gap: Is the Too Big to Fail Problem Resolved? *The Astrophysical Journal* **885**, 97 (2019). URL https://dx.doi.org/10.3847/1538-4357/ab3288. Publisher: The American Astronomical Society. (Cited on page 56.)

[136] Zavala, J. *et al.* The velocity function in the local environment from ΛCDM and ΛWDM constrained simulations. *The Astrophysical Journal* **700**, 1779 (2009). URL https://dx.doi.org/10.1088/0004-637X/700/2/1779. Publisher: The American Astronomical Society. (Cited on page 56.)

[137] Tikhonov, A. V. & Klypin, A. The emptiness of voids: yet another overabundance problem for the Λ cold dark matter model. *Monthly Notices of the Royal Astronomical Society* **395**, 1915–1924 (2009). URL https://ieeexplore.ieee.org/document/8222872. Conference Name: Monthly Notices of the Royal Astronomical Society. (Cited on page 56.)

[138] Dubinski, J. & Carlberg, R. G. The Structure of Cold Dark Matter Halos. *The Astrophysical Journal* **378**, 496 (1991). URL https://ui.adsabs.harvard.edu/abs/1991ApJ...378..496D. Publisher: IOP ADS Bibcode: 1991ApJ...378..496D. (Cited on page 56.)

[139] Walker, M. G. & Peñarrubia, J. A method for measuring (slopes of) the mass profiles of dwarf spheroidal galaxies. *The Astrophysical Journal* **742**, 20 (2011). URL https://dx.doi.org/10.1088/0004-637X/742/1/20. Publisher: The American Astronomical Society. (Cited on page 56.)

[140] Genina, A. *et al.* The core–cusp problem: a matter of perspective. *Monthly Notices of the Royal Astronomical Society* **474**, 1398–1411 (2018). URL https://doi.org/10.1093/mnras/stx2855. (Cited on page 56.)

[141] Kendall, E. & Easther, R. The core-cusp problem revisited: ULDM vs. CDM. *Publications of the Astronomical Society of Australia* **37**, e009 (2020). URL https://www.cambridge.org/core/journals/publications-of-the-astronomical-society-of-australia/article/corecusp-problem-revisited-uldm-vs-cdm/AA927FE820DA05854AED112B296E2789. (Cited on page 56.)

[142] Peebles, P. J. E. The Void Phenomenon. *The Astrophysical Journal* **557**, 495 (2001). URL https://dx.doi.org/10.1086/322254. (Cited on page 56.)

[143] Moss, A., Zibin, J. P. & Scott, D. Precision cosmology defeats void models for acceleration. *Physical Review D* **83**, 103515 (2011). URL https://link.aps.org/doi/10.1103/PhysRevD.83.103515. Publisher: American Physical Society. (Cited on page 56.)

[144] Tinker, J. L. & Conroy, C. The void phenomenon explained. *The Astrophysical Journal* **691**, 633 (2009). URL https://dx.doi.org/10.1088/0004-637X/691/1/633. Publisher: The American Astronomical Society. (Cited on page 56.)

[145] Pawlowski, M. S. It's time for some plane speaking. *Nature Astronomy* **5**, 1185–1187 (2021). URL https://www.nature.com/articles/s41550-021-01452-7. Publisher: Nature Publishing Group. (Cited on page 56.)

[146] Boylan-Kolchin, M. Planes of satellites are not a problem for (just) ΛCDM. *Nature Astronomy* **5**, 1188–1190 (2021). URL https://www.nature.com/articles/s41550-021-01467-0. Publisher: Nature Publishing Group. (Cited on page 56.)

[147] Sawala, T. *et al.* The Milky Way's plane of satellites is consistent with ΛCDM. *Nature Astronomy* **7**, 481–491 (2023). URL https://www.nature.com/articles/s41550-022-01856-z. Publisher: Nature Publishing Group. (Cited on page 56.)

[148] Sales, L. V. & Navarro, J. F. Planes of satellites no longer in tension with ΛCDM. *Nature Astronomy* **7**, 376–377 (2023). URL https://www.nature.com/articles/s41550-023-01924-y. Publisher: Nature Publishing Group. (Cited on page 56.)

[149] McGaugh, S. S. Novel Test of Modified Newtonian Dynamics with Gas Rich Galaxies. *Physical Review Letters* **106**, 121303 (2011). URL https://link.aps.org/doi/10.1103/PhysRevLett.106.121303. Publisher: American Physical Society. (Cited on page 56.)

[150] Famaey, B. & McGaugh, S. S. Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions. *Living Reviews in Relativity* **15**, 10 (2012). URL https://doi.org/10.12942/lrr-2012-10. (Cited on page 57.)

[151] Peebles, P. J. E. *Principles of Physical Cosmology* (1993). URL https://ui.adsabs.harvard.edu/abs/1993ppc..book.....P. Publication Title: Principles of physical cosmology ADS Bibcode: 1993ppc..book.....P. (Cited on page 58.)

[152] Lifshitz, E. M. On the gravitational stability of the expanding universe. *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* **16**, 587–602 (1946). URL https://ui.adsabs.harvard.edu/abs/1946ZhETF..16..587L. ADS Bibcode: 1946ZhETF..16..587L. (Cited on pages 58 and 61.)

[153] Sachs, R. K. & Wolfe, A. M. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *The Astrophysical Journal* **147**, 73 (1967). URL https://ui.adsabs.harvard.edu/abs/1967ApJ...147...73S. Publisher: IOP ADS Bibcode: 1967ApJ...147...73S. (Cited on pages 58 and 59.)

[154] Stewart, J. M. Perturbations of Friedmann-Robertson-Walker cosmological models. *Classical and Quantum Gravity* **7**, 1169–1180 (1990). URL https://iopscience.iop.org/article/10.1088/0264-9381/7/7/013. (Cited on pages 59, 60, 63, and 88.)

[155] Bertschinger, E. Cosmological Perturbation Theory and Structure Formation (2000). URL http://arxiv.org/abs/astro-ph/0101009. ArXiv:astro-ph/0101009. (Cited on page 59.)

[156] Sasaki, M. & Tanaka, T. Super-Horizon Scale Dynamics of Multi-Scalar Inflation. *Progress of Theoretical Physics* **99**, 763–781 (1998). URL http://arxiv.org/abs/gr-qc/9801017. ArXiv:gr-qc/9801017. (Cited on page 59.)

[157] Wands, D., Malik, K. A., Lyth, D. H. & Liddle, A. R. A new approach to the evolution of cosmological perturbations on large scales. *Physical Review D* **62**, 043527 (2000). URL http://arxiv.org/abs/astro-ph/0003278. ArXiv:astro-ph/0003278. (Cited on page 59.)

[158] Kurki-Suonio, H. *Cosmological Perturbation Theory I* (University of Helsinki, 2022). URL https://www.mv.helsinki.fi/home/hkurkisu/CosPer.pdf. (Cited on page 61.)

[159] Bardeen, J. M. Gauge-invariant cosmological perturbations. *Physical Review D* **22**, 1882–1905 (1980). URL https://link.aps.org/doi/10.1103/PhysRevD.22.1882. (Cited on pages 61, 62, 68, 69, and 70.)

[160] Kodama, H. & Sasaki, M. Cosmological Perturbation Theory. *Progress of Theoretical Physics Supplement* **78**, 1–166 (1984). URL https://doi.org/10.1143/PTPS.78.1. (Cited on page 62.)

[161] Clifton, T., Gallagher, C. S., Goldberg, S. & Malik, K. A. Viable Gauge Choices in Cosmologies with Non-Linear Structures. *Physical Review D* **101**, 063530 (2020). URL http://arxiv.org/abs/2001.00394. ArXiv:2001.00394 [gr-qc]. (Cited on page 62.)

[162] Clarkson, C. Local gauge-invariance at any order in cosmological perturbation theory (2011). URL http://arxiv.org/abs/1108.4513. ArXiv:1108.4513 [astro-ph]. (Cited on pages 62 and 63.)

[163] Dunsby, P. K. S. & Ellis, G. F. R. Perturbations of Cosmological Backgrounds. In Iyer, B. R. & Bhawal, B. (eds.) *Black Holes, Gravitational Radiation and the Universe: Essays in Honor of C.V. Vishveshwara*, 493–508 (Springer Netherlands, Dordrecht, 1999). URL https://doi.org/10.1007/978-94-017-0934-7_29. (Cited on page 62.)

[164] Hawking, S. W. Perturbations of an Expanding Universe. *The Astrophysical Journal* **145**, 544 (1966). URL http://adsabs.harvard.edu/doi/10.1086/148793. (Cited on pages 63 and 70.)

[165] Olson, D. W. Density perturbations in cosmological models. *Physical Review D* **14**, 327–331 (1976). URL https://link.aps.org/doi/10.1103/PhysRevD.14.327. Publisher: American Physical Society. (Cited on page 63.)

[166] Ellis, G. F. R. & van Elst, H. Cosmological models (Cargèse lectures 1998) (2008). URL http://arxiv.org/abs/gr-qc/9812046. ArXiv:gr-qc/9812046. (Cited on pages 63 and 71.)

[167] Ellis, G. F. R. Relativistic cosmology. *Cargèse Lectures in Physics* **6**, 1–60 (1973). (Cited on page 63.)

[168] Clarkson, C. & Osano, B. Locally extracting scalar, vector and tensor modes in cosmological perturbation theory. *Classical and Quantum Gravity - CLASS QUANTUM GRAVITY* **29** (2012). (Cited on page 64.)

[169] Ellis, G. F. R., Hwang, J. & Bruni, M. Covariant and gauge-independent perfect-fluid Robertson-Walker perturbations. *Physical Review D* **40**, 1819–1826 (1989). URL https://link.aps.org/doi/10.1103/PhysRevD.40.1819. (Cited on page 68.)

[170] Liddle, A. R. *An Introduction to Cosmological Inflation* (eprint: arXiv:astro-ph/9901124, 1999). URL https://ui.adsabs.harvard.edu/abs/1999hepc.conf..260L. Conference Name: High Energy Physics and Cosmology, 1998 Summer School Pages: 260 ADS Bibcode: 1999hepc.conf..260L. (Cited on page 70.)

[171] Hu, W. & Eisenstein, D. J. Structure of structure formation theories. *Physical Review D* **59**, 083509 (1999). URL https://link.aps.org/doi/10.1103/PhysRevD.59.083509. Publisher: American Physical Society. (Cited on page 70.)

[172] Norman, M. L. Simulating Galaxy Clusters (2010). URL http://arxiv.org/abs/1005.1100. ArXiv:1005.1100 version: 1. (Cited on page 73.)

[173] Peacock, J. A. Large-scale surveys and cosmic structure (2003). URL http://arxiv.org/abs/astro-ph/0309240. ArXiv:astro-ph/0309240. (Cited on page 74.)

[174] Hu, W. & Sawicki, I. Models of $f(R)$ Cosmic Acceleration that Evade Solar-System Tests. *Physical Review D* **76**, 064004 (2007). URL http://arxiv.org/abs/0705.1158. ArXiv:0705.1158 [astro-ph, physics:gr-qc, physics:hep-th]. (Cited on pages 76 and 115.)

[175] Starobinsky, A. A. Disappearing cosmological constant in $f(R)$ gravity. *JETP Letters* **86**, 157–163 (2007). URL http://arxiv.org/abs/0706.2041. ArXiv:0706.2041 [astro-ph, physics:gr-qc, physics:hep-th]. (Cited on page 76.)

[176] Olmo, G. J. Palatini Approach to Modified Gravity: $f(R)$ Theories and Beyond. *International Journal of Modern Physics D* **20**, 413–462 (2011). URL https://www.worldscientific.com/doi/abs/10.1142/S0218271811018925. Publisher: World Scientific Publishing Co. (Cited on page 78.)

[177] Dolgov, A. D. & Kawasaki, M. Can modified gravity explain accelerated cosmic expansion? *Physics Letters B* **573**, 1–4 (2003). URL http://arxiv.org/abs/astro-ph/0307285. ArXiv:astro-ph/0307285. (Cited on page 79.)

[178] Nojiri, S., Odintsov, S. D. & Saez-Gomez, D. Cosmological reconstruction of realistic modified $f(R)$ gravities. *Physics Letters B* **681**, 74–80 (2009). URL http://arxiv.org/abs/0908.1269. ArXiv:0908.1269 [astro-ph, physics:gr-qc, physics:hep-th]. (Cited on pages 80 and 99.)

[179] Carloni, S., Goswami, R. & Dunsby, P. K. S. A new approach to reconstruction methods in $f(R)$ gravity. *Classical and Quantum Gravity* **29**, 135012 (2012). URL https://iopscience.iop.org/article/10.1088/0264-9381/29/13/135012. (Cited on pages 80 and 99.)

[180] He, J.-h. & Wang, B. Revisiting $f(R)$ gravity models that reproduce ΛCDM expansion. *Physical Review D* **87**, 023508 (2013). URL https://link.aps.org/doi/10.1103/PhysRevD.87.023508. Publisher: American Physical Society. (Cited on pages 80 and 108.)

[181] Wainwright, J. & Ellis, G. F. R. (eds.) *Dynamical Systems in Cosmology* (Cambridge University Press, Cambridge, 1997). URL https://www.cambridge.org/core/books/dynamical-systems-in-cosmology/70E346CFD5114030B7B62079C7A58289. (Cited on page 80.)

[182] Coley, A. A. *Dynamical Systems and Cosmology*, vol. 291 of *Astrophysics and Space Science Library* (Springer Netherlands, Dordrecht, 2003). URL http://link.springer.com/10.1007/978-94-017-0327-7. (Cited on page 80.)

[183] Bahamonde, S. *et al.* Dynamical systems applied to cosmology: dark energy and modified gravity. *Physics Reports* **775-777**, 1–122 (2018). URL http://arxiv.org/abs/1712.03107. ArXiv:1712.03107 [astro-ph, physics:gr-qc, physics:hep-th]. (Cited on page 80.)

[184] Carloni, S., Dunsby, P. K. S., Capozziello, S. & Troisi, A. Cosmological dynamics of R^n gravity. *Classical and Quantum Gravity* **22**, 4839–4868 (2005). URL http://arxiv.org/abs/gr-qc/0410046. ArXiv:gr-qc/0410046. (Cited on page 80.)

[185] Goliath, M. & Ellis, G. F. R. Homogeneous cosmologies with a cosmological constant. *Physical Review D* **60**, 023502 (1999). URL https://link.aps.org/doi/10.1103/PhysRevD.60.023502. Publisher: American Physical Society. (Cited on page 80.)

[186] Carloni, S. A new approach to the analysis of the phase space of $f(R)$-gravity. *Journal of Cosmology and Astroparticle Physics* **2015**, 013 (2015). URL https://dx.doi.org/10.1088/1475-7516/2015/09/013. (Cited on page 81.)

[187] Carloni, S., Dunsby, P. K. S. & Troisi, A. The evolution of density perturbations in $f(R)$ gravity. *Physical Review D* **77**, 024024 (2008). URL http://arxiv.org/abs/0707.0106. ArXiv:0707.0106 [astro-ph, physics:gr-qc, physics:hep-th]. (Cited on pages 89 and 92.)

[188] Abebe, A., Dombriz, A. d. l. C. & Dunsby, P. K. S. Large Scale Structure Constraints for a Class of $f(R)$ Theories of Gravity. *Physical Review D* **88**, 044050 (2013). URL http://arxiv.org/abs/1304.3462. ArXiv:1304.3462 [astro-ph, physics:gr-qc]. (Cited on pages 89 and 107.)

[189] Bruni, M., Dunsby, P. K. S. & Ellis, G. F. R. Cosmological perturbations and the physical meaning of gauge-invariant variables. *The Astrophysical Journal* **395**, 34 (1992). URL http://adsabs.harvard.edu/doi/10.1086/171629. (Cited on page 92.)

[190] Silvestri, A., Pogosian, L. & Buniy, R. V. Practical approach to cosmological perturbations in modified gravity. *Physical Review D* **87**, 104015 (2013). URL https://link.aps.org/doi/10.1103/PhysRevD.87.104015. Publisher: American Physical Society. (Cited on page 92.)

[191] Noller, J., von Braun-Bates, F. & Ferreira, P. G. Relativistic scalar fields and the quasistatic approximation in theories of modified gravity. *Physical Review D* **89**, 023521 (2014). URL https://link.aps.org/doi/10.1103/PhysRevD.89.023521. Publisher: American Physical Society. (Cited on pages 92 and 93.)

[192] Chiu, M.-C., Taylor, A., Shu, C. & Tu, H. Cosmological perturbations and quasistatic assumption in $f(R)$ theories. *Physical Review D* **92**, 103514 (2015). URL https://link.aps.org/doi/10.1103/PhysRevD.92.103514. Publisher: American Physical Society. (Cited on page 92.)

[193] Hojjati, A., Pogosian, L., Silvestri, A. & Talbot, S. Practical solutions for perturbed $f(R)$ gravity. *Physical Review D* **86**, 123503 (2012). URL https://link.aps.org/doi/10.1103/PhysRevD.86.123503. Publisher: American Physical Society. (Cited on page 93.)

[194] Sawicki, I. & Bellini, E. Limits of quasistatic approximation in modified-gravity cosmologies. *Physical Review D* **92**, 084061 (2015). URL https://link.aps.org/doi/10.1103/PhysRevD.92.084061. Publisher: American Physical Society. (Cited on page 93.)

[195] de la Cruz-Dombriz, A., Dobado, A. & Maroto, A. L. On the evolution of density perturbations in $f(R)$ theories of gravity. *Physical Review D* **77**, 123515 (2008). (Cited on page 93.)

[196] Bean, R., Bernat, D., Pogosian, L., Silvestri, A. & Trodden, M. Dynamics of linear perturbations in $f(R)$ gravity. *Physical Review D* **75**, 064020 (2007). URL https://link.aps.org/doi/10.1103/PhysRevD.75.064020. Publisher: American Physical Society. (Cited on page 93.)

[197] Polarski, D. & Gannouji, R. On the growth of linear perturbations. *Physics Letters B* **660**, 439–443 (2008). URL https://linkinghub.elsevier.com/retrieve/pii/S0370269308000567. (Cited on pages 97 and 108.)

[198] Peebles, P. J. E. *The large-scale structure of the universe* (1980). URL https://ui.adsabs.harvard.edu/abs/1980lssu.book.....P. Publication Title: Large-Scale Structure of the Universe by Phillip James Edwin Peebles. Princeton University Press ADS Bibcode: 1980lssu.book.....P. (Cited on page 97.)

[199] Lightman, A. P. & Schechter, P. L. The Omega Dependence of Peculiar Velocities Induced by Spherical Density Perturbations. *The Astrophysical Journal Supplement Series* **74**, 831 (1990). URL https://ui.adsabs.harvard.edu/abs/1990ApJS...74..831L. Publisher: IOP ADS Bibcode: 1990ApJS...74..831L. (Cited on page 97.)

[200] Gannouji, R. & Polarski, D. The growth of matter perturbations in some scalar-tensor DE models. *Journal of Cosmology and Astroparticle Physics* **2008**, 018 (2008). URL http://arxiv.org/abs/0802.4196. ArXiv:0802.4196 [astro-ph, physics:gr-qc]. (Cited on pages 97 and 108.)

[201] Wang, L. & Steinhardt, P. J. Cluster Abundance Constraints on Quintessence Models. *The Astrophysical Journal* **508**, 483–490 (1998). URL http://arxiv.org/abs/astro-ph/9804015. ArXiv:astro-ph/9804015. (Cited on page 97.)

[202] Press, W. H. & Schechter, P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *The Astrophysical Journal* **187**, 425–438 (1974). URL https://ui.adsabs.harvard.edu/abs/1974ApJ...187..425P. Publisher: IOP ADS Bibcode: 1974ApJ...187..425P. (Cited on page 97.)

[203] Reid, B. A. *et al.* The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: measurements of the growth of structure and expansion rate at $z = 0.57$ from anisotropic clustering. *Monthly Notices of the Royal Astronomical Society* **426**, 2719–2737 (2012). URL https://doi.org/10.1111/j.1365-2966.2012.21779.x. (Cited on pages 97 and 108.)

[204] Hu, W. Power Spectrum Tomography with Weak Lensing. *The Astrophysical Journal* **522**, L21 (1999). URL https://dx.doi.org/10.1086/312210. (Cited on page 97.)

[205] Zhao, G.-B. *et al.* Probing modifications of general relativity using current cosmological observations. *Physical Review D* **81**, 103510 (2010). URL https://link.aps.org/doi/10.1103/PhysRevD.81.103510. Publisher: American Physical Society. (Cited on page 97.)

[206] Jasche, J., Kitaura, F. S., Wandelt, B. D. & Enßlin, T. A. Bayesian power-spectrum inference for large-scale structure data. *Monthly Notices of the Royal Astronomical Society* **406**, 60–85 (2010). URL https://doi.org/10.1111/j.1365-2966.2010.16610.x. (Cited on page 97.)

[207] Dunsby, P. K. S., Elizalde, E., Goswami, R., Odintsov, S. & Saez-Gomez, D. ΛCDM universe in $f(R)$ gravity. *Physical Review D* **82**, 023519 (2010). URL https://link.aps.org/doi/10.1103/PhysRevD.82.023519. Publisher: American Physical Society. (Cited on pages 99 and 108.)

[208] Mehrabi, A. & Rezaei, M. Cosmographic Parameters in Model-independent Approaches. *The Astrophysical Journal* **923**, 274 (2021). URL https://dx.doi.org/10.3847/1538-4357/ac2fff. Publisher: The American Astronomical Society. (Cited on pages 99 and 100.)

[209] Mukherjee, P. & Banerjee, N. Non-parametric reconstruction of the cosmological jerk parameter. *The European Physical Journal C* **81**, 36 (2021). URL https://doi.org/10.1140/epjc/s10052-021-08830-5. (Cited on page 99.)

[210] Zhai, Z.-X., Zhang, M.-J., Zhang, Z.-S., Liu, X.-M. & Zhang, T.-J. Reconstruction and constraining of the jerk parameter from OHD and SNe Ia observations. *Physics Letters B* **727**, 8–20 (2013). URL https://www.sciencedirect.com/science/article/pii/S0370269313008137. (Cited on pages 99 and 100.)

[211] Amirhashchi, H. & Amirhashchi, S. Recovering ΛCDM model from a cosmographic study. *General Relativity and Gravitation* **52**, 13 (2020). URL https://doi.org/10.1007/s10714-020-2664-5. (Cited on page 99.)

[212] Mukherjee, A. & Banerjee, N. Parametric reconstruction of the cosmological jerk from diverse observational data sets. *Physical Review D* **93**, 043002 (2016). URL https://link.aps.org/doi/10.1103/PhysRevD.93.043002. Publisher: American Physical Society. (Cited on page 99.)

[213] Sahni, V., Saini, T. D., Starobinsky, A. A. & Alam, U. Statefinder—A new geometrical diagnostic of dark energy. *Journal of Experimental and Theoretical Physics Letters* **77**, 201–206 (2003). URL https://doi.org/10.1134/1.1574831. (Cited on page 99.)

[214] Alam, U., Sahni, V., Deep Saini, T. & Starobinsky, A. A. Exploring the expanding Universe and dark energy using the statefinder diagnostic. *Monthly Notices of the Royal Astronomical Society* **344**, 1057–1074 (2003). URL https://doi.org/10.1046/j.1365-8711.2003.06871.x. (Cited on page 100.)

[215] Chakraborty, S., Gregoris, D. & Mishra, B. On the uniqueness of ΛCDM-like evolution for homogeneous and isotropic cosmology in General Relativity. *Physics Letters B* **842**, 137962 (2023). URL https://www.sciencedirect.com/science/article/pii/S0370269323002964. (Cited on page 100.)

[216] Tsujikawa, S., Gannouji, R., Moraes, B. & Polarski, D. Dispersion of growth of matter perturbations in $f(R)$ gravity. *Physical Review D* **80**, 084044 (2009). URL https://link.aps.org/doi/10.1103/PhysRevD.80.084044. (Cited on page 105.)

[217] Pogosian, L. & Silvestri, A. Pattern of growth in viable $f(R)$ cosmologies. *Physical Review D* **77**, 023503 (2008). URL https://link.aps.org/doi/10.1103/PhysRevD.77.023503. Publisher: American Physical Society. (Cited on page 105.)

[218] Huterer, D. *et al.* Growth of cosmic structure: Probing dark energy beyond expansion. *Astroparticle Physics* **63**, 23–41 (2015). URL https://www.sciencedirect.com/science/article/pii/S0927650514001005. (Cited on page 108.)

[219] Narikawa, T. & Yamamoto, K. Characterizing the linear growth rate of cosmological density perturbations in an $f(R)$ model. *Physical Review D* **81**, 043528 (2010). URL https://link.aps.org/doi/10.1103/PhysRevD.81.043528. Publisher: American Physical Society. (Cited on page 108.)

[220] Mirzatuny, N. & Pierpaoli, E. An accurate fitting function for scale-dependent growth rate in Hu-Sawicki $f(R)$ gravity. *Journal of Cosmology and Astroparticle Physics* **2019**, 066 (2019). URL https://dx.doi.org/10.1088/1475-7516/2019/09/066. (Cited on page 108.)

[221] Tsujikawa, S. Modified gravity models of dark energy (2010). URL http://arxiv.org/abs/1101.0191. ArXiv:1101.0191. (Cited on page 108.)

[222] Fonseca, J., Viljoen, J.-A. & Maartens, R. Constraints on the growth rate using the observed galaxy power spectrum. *Journal of Cosmology and Astroparticle Physics* **2019**, 028 (2019). URL https://dx.doi.org/10.1088/1475-7516/2019/12/028. (Cited on page 108.)

Appendix A

General propagation equations of the $1+3$ covariant formalism

Expansion propagation (generalised Raychaudhuri equation):

$$\dot{\Theta}+\frac{1}{3}\Theta^2+\sigma_{ab}\sigma^{ab}-2\omega_a\omega^a-\widetilde{\nabla}^a\dot{u}_a+\dot{u}_a\dot{u}^a+\frac{1}{2}\left(\tilde{\rho}^m+3\tilde{p}^m\right)=-\frac{1}{2}\left(\rho^R+3p^R\right)\,. \tag{A.1}$$

Vorticity propagation:

$$\dot{\omega}_{\langle a\rangle}+\frac{2}{3}\Theta\omega_a+\frac{1}{2}\operatorname{curl}\dot{u}_a-\sigma_{ab}\omega^b=0\,. \tag{A.2}$$

Shear propagation:

$$\dot{\sigma}_{\langle ab\rangle}+\frac{2}{3}\Theta\sigma_{ab}+E_{ab}-\widetilde{\nabla}_{\langle a}\dot{u}_{b\rangle}+\sigma_{c\langle a}\sigma_{b\rangle}{}^c+\omega_{\langle a}\omega_{b\rangle}-\dot{u}_{\langle a}\dot{u}_{b\rangle}=\frac{1}{2}\pi^R_{ab}. \tag{A.3}$$

Gravito-electric propagation:

$$\begin{aligned}&\dot{E}_{\langle ab\rangle}+\Theta E_{ab}-\operatorname{curl}H_{ab}+\frac{1}{2}\left(\tilde{\rho}^m+\tilde{p}^m\right)\sigma_{ab}-2\dot{u}^c\varepsilon_{cd(a}H_{b)}{}^d-3\sigma_{c\langle a}E_{b\rangle}{}^c+\omega^c\varepsilon_{cd(a}E_{b)}{}^d\\&\quad=-\frac{1}{2}\left(\rho^R+p^R\right)\sigma_{ab}-\frac{1}{2}\dot{\pi}^R_{\langle ab\rangle}-\frac{1}{2}\widetilde{\nabla}_{\langle a}q^R_{b\rangle}-\frac{1}{6}\Theta\pi^R_{ab}-\frac{1}{2}\sigma^c{}_{\langle a}\pi^R_{b\rangle c}-\frac{1}{2}\omega^c\varepsilon^d_{c(a}\pi^R_{b)d}\,.\end{aligned} \tag{A.4}$$

Gravito-magnetic propagation:

$$\begin{aligned}&\dot{H}_{\langle ab\rangle}+\Theta H_{ab}+\operatorname{curl}E_{ab}-3\sigma_{c\langle a}H_{b\rangle}{}^c+\omega^c\varepsilon_{cd(a}H_{b)}{}^d+2\dot{u}^c\varepsilon_{cd(a}E_{b)}{}^d\\&\quad=\frac{1}{2}\operatorname{curl}\pi^R_{ab}-\frac{3}{2}\omega_{\langle a}q^R_{b\rangle}+\frac{1}{2}\sigma^c{}_{(a}\varepsilon_{b)c}{}^dq^R_d\,.\end{aligned} \tag{A.5}$$

Vorticity constraint:

$$\widetilde{\nabla}^a\omega_a-\dot{u}^a\omega_a=0\,. \tag{A.6}$$

Shear constraint:

$$\widetilde{\nabla}^b\sigma_{ab}-\operatorname{curl}\omega_a-\frac{2}{3}\widetilde{\nabla}_a\Theta+2[\omega,\dot{u}]_a=-q^R_a\,. \tag{A.7}$$

Gravito-magnetic constraint:

$$\operatorname{curl}\sigma_{ab}+\widetilde{\nabla}_{\langle a}\omega_{b\rangle}-H_{ab}+2\dot{u}_{\langle a}\omega_{b\rangle}=0\,. \tag{A.8}$$

Gravito-electric divergence:

$$\begin{aligned}&\widetilde{\nabla}^bE_{ab}-\frac{1}{3}\widetilde{\nabla}_a\tilde{\rho}^m-[\sigma,H]_a+3H_{ab}\omega^b\\&\quad=\frac{1}{2}\sigma^b_aq^R_b-\frac{3}{2}\left[\omega,q^R\right]_a-\frac{1}{2}\widetilde{\nabla}^b\pi^R_{ab}+\frac{1}{3}\widetilde{\nabla}_a\rho^R-\frac{1}{3}\Theta q^R_a\,.\end{aligned} \tag{A.9}$$

Gravito-magnetic divergence:

$$\begin{aligned}&\widetilde{\nabla}^b H_{ab} - (\tilde{\rho}^m + \tilde{p}^m)\,\omega_a + [\sigma, E]_a - 3E_{ab}\omega^b \\ &\quad = -\frac{1}{2}\operatorname{curl} q_a^R + \left(\rho^R + p^R\right)\omega_a - \frac{1}{2}\left[\sigma, \pi^R\right]_a - \frac{1}{2}\pi_{ab}^R\omega^b\,. \end{aligned} \tag{A.10}$$

Standard Matter Conservation:

$$\dot{\rho}^m = -\Theta\left(\rho^m + p^m\right)\,, \tag{A.11}$$

$$\widetilde{\nabla}^a p^m = -\left(\rho^m + p^m\right)\dot{u}^a\,. \tag{A.12}$$

Angle brackets applied to a vector denote the projection of this vector on the tangent 3-spaces

$$V_{\langle a\rangle} = h_a{}^b V_b\,. \tag{A.13}$$

Appendix B

Covariant Identities

On a flat Friedmann background, the following covariant linearised identities hold:

$$\widetilde{\nabla}_a \dot{f} = \left(\widetilde{\nabla}_a f\right)^{\cdot} + \frac{1}{3}\Theta\widetilde{\nabla}_a f - \dot{f}\dot{u}_a, \tag{B.1}$$

$$\widetilde{\nabla}^2\left(\widetilde{\nabla}_a f\right) = \widetilde{\nabla}_a\left(\widetilde{\nabla}^2 f\right) - \frac{2K}{a^2}\widetilde{\nabla}_a f + 2\dot{f}\omega_a, \tag{B.2}$$

$$\widetilde{\nabla}^2 \dot{f} = \left(\widetilde{\nabla}^2 f\right)^{\cdot} + \frac{2}{3}\Theta\widetilde{\nabla}^2 f - \dot{f}\widetilde{\nabla}^a \dot{u}_a, \tag{B.3}$$

$$\left(\widetilde{\nabla}_a V_b\right)^{\cdot} = \widetilde{\nabla}_a \dot{V}_b - \frac{1}{3}\Theta\widetilde{\nabla}_a V_b, \tag{B.4}$$

$$\widetilde{\nabla}_{[a}\widetilde{\nabla}_{b]} V_c = -\frac{K}{a^2} V_{[a} h_{b]c}, \tag{B.5}$$

$$\widetilde{\nabla}^b \widetilde{\nabla}_{\langle a} V_{b\rangle} = \frac{1}{2}\widetilde{\nabla}^2 V_a + \frac{1}{6}\widetilde{\nabla}_a\left(\widetilde{\nabla}^b V_b\right) + \frac{K}{a^2} V_a, \tag{B.6}$$

$$\left(\widetilde{\nabla}_a W_{cd}\right)^{\cdot} = \widetilde{\nabla}_a \dot{W}_{cd} - \frac{1}{3}\Theta\widetilde{\nabla}_a W_{cd}\,, \tag{B.7}$$

where $V_a = V_{\langle a\rangle}$ and $W_{ab} = W_{\langle ab\rangle}$ are first-order quantities.

Appendix C

Coefficient Definitions

C.1 Equation (4.93b)

Coefficient A:

$$A = \frac{q\big[(j+18)x - 4j - 6(1+w)\Omega + 12\big] + 2x\big[-2j+s+10\big] + 2(j-2)\big[3(1+w)\Omega - 2\big] + q^2(x+4)}{x(j-q-2)}\,.$$

Coefficient B:

$$B = \frac{B_1 + B_2 + B_3 + B_4 + B_5}{x(j-q-2)^2}\,,$$

where

$$\begin{aligned}
B_1 &= x\Big[11j^2 - j(l+4s+34) + 2\big(l + (s+4)(s+12)\big)\Big] \\
B_2 &= q\Big[j(5j-22) + l + 36s + 206\Big] \\
B_3 &= -6q(w+1)\Omega\Big[3j(w-3) + s - 6w + 24\Big] \\
B_4 &= q^2\Big[x(5j+4s+133) + 4(-9j+s+32) + 3(w+1)\Omega\big(2j + 3(w-7)\big)\Big] \\
B_5 &= -4(j-2)(s+6) + 3(w+1)\Omega(j-2)\Big[3(j-2)w - j + 2s + 14\Big] \\
&\quad + q^3\Big[-4j - 6(w+1)\Omega + 22x + 44\Big] + q^4\Big[2(x+2)\Big]\,.
\end{aligned}$$

Coefficient C:

$$C = \frac{6(j-q-2)\big[4w(1+q+2x) - (1+w)(1-3w)\Omega\big]}{(1+w)x}\,.$$

Coefficient D:

$$D = \frac{6(1-w)(j-q-2)}{(1+w)}\,.$$

C.2 Equation (4.94)

Coefficient E:

$$E = \frac{\hat{k}^2}{h^2}(1+z)^2 + \frac{1}{x(j-q-2)^2}\Big(E_1 + E_2 + E_3 + E_4 + E_5\Big),$$

where

$$\begin{aligned}
E_1 &= x\Big[11j^2 - j(34+l+4s) + 2\big(l + (4+s)(12+s)\big)\Big],\\
E_2 &= q\Big[136 + 12s - 4j(14+s) + x\big(206 + j(-22+5j) + l + 36s\big)\\
&\quad - 6\big(24 + s + 3j(-3+w) - 6w\big)(1+w)\Omega\Big],\\
E_3 &= q^2\Big[4(32 - 9j + s) + (133 + 5j + 4s)x + 3(2j + 3(-7+w))(1+w)\Omega\Big],\\
E_4 &= q^3\Big[44 - 4j + 22x - 6(1+w)\Omega\Big],\\
E_5 &= q^4\Big[2 + x\Big] + (j-2)\Big[-4(6+s) + 3(1+w)\big(14 - j + 2s + 3(j-2)w\big)\Omega\Big].
\end{aligned}$$

Coefficient F:

$$F = 4w(q + 2x + 1) - (w+1)(3w+1)\Omega\,.$$