



Relativistic cosmology and large-scale structure

Christos G. Tsagas^{a,*}, Anthony Challinor^{b,c}, Roy Maartens^d

^a Section of Astrophysics, Astronomy and Mechanics, Department of Physics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece

^b Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

^c DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

^d Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth P01 2EG, UK

ARTICLE INFO

Article history:

Accepted 25 March 2008

Available online 28 March 2008

editor: M.P. Kamionkowski

PACS:

98.80.-k

98.80.Jk

98.65.Dx

Keywords:

Cosmology

Large-scale structure

ABSTRACT

General relativity marked the beginning of modern cosmology and it has since been at the centre of many of the key developments in this field. In the present review, we discuss the general-relativistic dynamics and perturbations of the standard cosmological model, the Friedmann–Lemaître universe, and how these can explain and predict the properties of the observable universe. Our aim is to provide an overview of the progress made in several major research areas, such as linear and non-linear cosmological perturbations, large-scale structure formation and the physics of the cosmic microwave background radiation, in view of current and upcoming observations. We do this by using a single formalism throughout the review, the 1 + 3 covariant approach to cosmology, which allows for a uniform and balanced presentation of technical information and physical insight.

© 2008 Elsevier B.V. All rights reserved.

Contents

1.	Relativistic cosmology.....	63
1.1.	Cosmology at the dawn of the 21st century.....	63
1.2.	The 1 + 3 covariant description.....	67
1.2.1.	Local spacetime splitting.....	67
1.2.2.	The gravitational field.....	67
1.2.3.	Matter fields.....	68
1.3.	Covariant relativistic cosmology.....	69
1.3.1.	Kinematics.....	69
1.3.2.	Electromagnetic fields.....	70
1.3.3.	Conservation laws.....	71
1.3.4.	Equilibrium thermodynamics.....	71
1.3.5.	Spatial curvature.....	72
1.3.6.	Weyl curvature.....	73
1.4.	The Friedmann–Lemaître universe.....	74
1.4.1.	The FLRW metric.....	74
1.4.2.	FLRW cosmologies.....	74
1.4.3.	Luminosity distance.....	75
1.4.4.	Scale-factor evolution in FLRW cosmologies.....	75
1.5.	The Bianchi universes.....	77
1.5.1.	Classification of Bianchi cosmologies.....	77

* Corresponding author.

E-mail address: tsagas@astro.auth.gr (C.G. Tsagas).

1.5.2.	Bianchi I cosmologies.....	78
1.5.3.	Bianchi VII_h cosmologies.....	79
1.5.4.	Bianchi IX cosmologies.....	80
1.5.5.	Isotropisation of Bianchi cosmologies.....	80
1.5.6.	Kantowski–Sachs cosmologies.....	80
2.	Inhomogeneous relativistic cosmologies.....	81
2.1.	The gauge problem in cosmology.....	81
2.1.1.	Gauge freedom.....	81
2.1.2.	Gauge dependence.....	81
2.2.	Covariant and gauge-invariant perturbations.....	82
2.2.1.	Criteria for gauge invariance.....	82
2.2.2.	Gauge-invariant inhomogeneities.....	82
2.3.	Inhomogeneous single-fluid cosmologies.....	82
2.3.1.	Imperfect fluids.....	82
2.3.2.	Perfect fluids.....	83
2.3.3.	Covariant conserved quantities.....	83
2.4.	Inhomogeneous multi-fluid cosmologies.....	84
2.4.1.	4-velocity fields.....	84
2.4.2.	Multi-component perfect fluids.....	84
2.4.3.	Conservation laws of the total fluid.....	85
2.4.4.	Conservation laws of the i -th fluid.....	85
2.4.5.	Nonlinear density perturbations in the i -th species.....	86
2.5.	Inhomogeneous magnetised cosmologies.....	86
2.5.1.	The ideal MHD approximation.....	86
2.5.2.	Conservation laws.....	87
2.5.3.	Nonlinear density perturbations of the magnetised fluid.....	87
2.6.	Inhomogeneous scalar-field cosmologies.....	88
2.6.1.	Minimally coupled scalar fields.....	88
2.6.2.	Scalar-field kinematics.....	88
2.6.3.	Scalar fields as perfect fluids.....	89
2.6.4.	Conservation laws.....	89
2.6.5.	Nonlinear scalar-field perturbations.....	89
3.	Linear cosmological perturbations.....	90
3.1.	Linearisation.....	90
3.1.1.	Interpreting the inhomogeneity variables.....	90
3.2.	Single-fluid perturbations.....	91
3.2.1.	Linear evolution equations.....	91
3.2.2.	Three types of inhomogeneity.....	92
3.2.3.	Density perturbations.....	92
3.2.4.	Isocurvature perturbations.....	93
3.2.5.	Density vortices.....	94
3.2.6.	Dissipative effects.....	94
3.3.	Multi-fluid perturbations.....	95
3.3.1.	Linearised evolution equations for the i -th species.....	95
3.3.2.	Entropy perturbations in a multi-fluid system.....	96
3.3.3.	Density perturbations in the i -th species.....	96
3.3.4.	A radiation and dust universe.....	97
3.3.5.	A CDM and baryon universe.....	97
3.4.	Magnetised perturbations.....	98
3.4.1.	The Alfvén speed.....	98
3.4.2.	Magnetised density perturbations.....	99
3.4.3.	Radiation era.....	99
3.4.4.	Dust era.....	100
3.5.	Scalar-field perturbations.....	101
3.5.1.	The effective fluid characteristics.....	101
3.5.2.	Density perturbations.....	101
3.5.3.	Standard slow-roll inflation.....	102
3.5.4.	Coasting universe.....	102
3.6.	Gravitational wave perturbations.....	102
3.6.1.	Isolating tensor modes.....	103
3.6.2.	Covariant description of the gravitational-wave energy density.....	103
3.6.3.	Evolution of gravitational waves.....	103
3.7.	Perturbed non-Friedmannian cosmologies.....	104
3.7.1.	The Bianchi I universe.....	105
3.7.2.	The Einstein static universe.....	105
3.7.3.	The Gödel universe.....	106

3.7.4.	The Lukash Bianchi VII_h universe	107
4.	Kinetic theory and the cosmic neutrino and microwave backgrounds.....	108
4.1.	Distribution functions and the Liouville equation	108
4.1.1.	Bulk properties	109
4.1.2.	Linearisation around FLRW cosmologies	110
4.2.	Cosmic neutrino background	114
4.2.1.	Scalar perturbations.....	114
4.2.2.	Tensor perturbations	115
4.3.	Cosmic microwave background.....	116
4.3.1.	CMB observables	117
4.3.2.	Transformation properties under change of frame	119
4.3.3.	Radiative transfer	120
4.3.4.	Scalar perturbations.....	122
4.3.5.	Vector perturbations.....	126
4.3.6.	Tensor perturbations	128
4.3.7.	Cosmic microwave background in other cosmological models	129
5.	Beyond the linear regime.....	130
5.1.	Nonlinear peculiar kinematics	130
5.1.1.	$1 + 3$ peculiar-velocity decomposition	130
5.1.2.	Nonlinear peculiar motions.....	131
5.1.3.	The case of dust	132
5.2.	The mildly nonlinear regime.....	132
5.2.1.	Second order evolution.....	132
5.2.2.	Rescaling the variables	133
5.3.	The relativistic Zeldovich approximation	133
5.3.1.	The Zeldovich ansatz	133
5.3.2.	The Zeldovich pancakes.....	134
5.4.	Averaging and backreaction.....	135
5.4.1.	Spatial averaging.....	135
5.4.2.	The averaged equations.....	135
5.5.	Backreaction and accelerated expansion	136
5.5.1.	Conditions for acceleration	136
5.5.2.	Attractive aspects and caveats	137
6.	Summary and outlook.....	137
	Acknowledgements	138
Appendix A.	138
A.1.	$1 + 3$ Covariant decomposition	138
A.2.	Transformations under a 4-velocity boost.....	139
A.3.	Covariant commutation laws.....	140
A.4.	Scalar, vector and tensor modes.....	141
Appendix B.	Notation.....	141
	References	142

1. Relativistic cosmology

Cosmology is the study of the dynamics and make-up of the Universe as a whole, or at least the maximally observable region of the Universe. Less than 100 years ago, the prevailing view (shared by Einstein) was that the Universe was static, and the existence of galaxies beyond our own remained unknown. A revolution was initiated via observations by Leavitt, Hubble and others which showed that the Universe was in fact expanding and contained many distant galaxies. Friedmann, Lemaitre and other theorists showed how the expansion could be explained by a spatially homogeneous and isotropic model obeying the field equations of General Relativity. The expansion pointed to an extremely hot origin of the Universe, the Big Bang, and Gamow and others showed how this should leave a thermal relic radiation, and also how nucleosynthesis of the light elements would take place in the hot early universe. However, it took many decades for observations to catch up and confirm this, and to lay the basis for further developments.

1.1. Cosmology at the dawn of the 21st century

Cosmology has come of age as an observationally based physical science in the last few decades, driven by the tremendous growth in data from increasingly high-precision experiments. Key milestones since 1990 include:

- the Cosmic Background Explorer (COBE), that detected the large-angle anisotropies in the Cosmic Microwave Background (CMB) temperature, and its successors, especially the Wilkinson Microwave Anisotropy Probe (WMAP), that measured the anisotropies at small angles and detected the acoustic peaks;

- the 2-degree Field (2dF) Galaxy Redshift Survey, that measured the matter power spectrum based on over 200k galaxies, and its successor, the Sloan Digital Sky Survey (SDSS);
- the Supernova Cosmology Project (SCP), that measured the magnitude–redshift data for more than 40 supernovae (SNe), and its successors, including the Supernova Legacy Survey (SNLS).

Collectively, these experiments and others, including measurements of weak lensing, underpin our current understanding of the evolutionary history and contents of the Universe. A broad range of new and upcoming experiments will aim to refine and extend this understanding.

The theoretical efforts to interpret the observational data and to make further predictions that can be tested against observations, have involved an important interplay between general relativity, astrophysics, particle physics and computation. The current model of large-scale structure formation in the Universe is based on the following:

- A spatially homogeneous and isotropic Friedmann–Lemaître–Robertson–Walker (FLRW) background spacetime,

$$ds^2 = -dt^2 + a^2(t) \left[dr^2 + f_K^2(r) (d\theta^2 + \sin^2 \theta d\phi^2) \right], \quad (1.1.1)$$

where the form of $f_K(r)$ depends on the model's spatial curvature—see Section 1.4.1, represents the average dynamics on large scales.

- The expansion rate, H , is governed by the Friedmann equation,

$$H^2 \equiv \left(\frac{\dot{a}}{a} \right)^2 = \frac{1}{3} (\rho_{(r)} + \rho_{(c)} + \rho_{(b)} + \rho_{(de)}) - \frac{K}{a^2}, \quad (1.1.2)$$

with $K = 0, \pm 1$. Thus, H is determined by the radiation, cold dark matter, baryonic matter and dark energy content of the Universe once K , the spatial curvature index, is chosen. The energy densities redshift with expansion according to the conservation law,

$$\dot{\rho}_{(i)} + 3H(1 + w_{(i)})\rho_{(i)} = 0, \quad (1.1.3)$$

where $w_{(i)} = p_{(i)}/\rho_{(i)}$ with $w_{(r)} = 1/3$, $w_{(c)} = 0 = w_{(b)}$, $w_{(de)} < -1/3$. The primordial radiation-dominated era is preceded by a brief burst of inflationary expansion, driven by a scalar field (or fields). Radiation decouples from baryonic matter soon after the total matter begins to dominate. At relatively recent times, matter begins to give way to a negative-pressure dark energy component which starts to accelerate the expansion again. The simplest model of dark energy has $\rho_{de} = \Lambda$, the cosmological constant, representing the vacuum energy density, with $w_\Lambda = -1$.

- Large-scale structure emerges as small over-densities begin to grow in the matter-dominated era via gravitational instability. In order to grow the galaxies fast enough, non-baryonic cold dark matter is needed in the standard model based on general relativity. The seeds of these over-densities are provided by the vacuum fluctuations of the inflaton field. The simplest inflation models predict a nearly scale-invariant and Gaussian spectrum of density perturbations, and a sub-dominant component of gravitational wave perturbations. The imprint of these primordial density perturbations is recorded in the CMB anisotropies, and the subsequent evolution is measured via the evolving galaxy distribution. The inflationary model provides not only the seeds for the emergence of observed large-scale structure, but also resolves the critical puzzle within non-inflationary models: that widely separated parts of the CMB sky were never in causal contact and, yet, have the same temperature.

The current “standard model” of cosmology is the inflationary Cold Dark Matter (CDM) model with cosmological constant, usually called Λ CDM, which is based on general relativity and particle physics (i.e. the Standard Model, with minimal extensions). The Λ CDM model provides an excellent fit to the wealth of high-precision observational data, on the basis of a remarkably small number of cosmological parameters (see, e.g., [1,2]). In particular, independent data sets from CMB anisotropies, galaxy surveys and supernova luminosities, lead to a consistent set of best-fit model parameters. This is illustrated in Figs. 1 and 2.

The Λ CDM model is remarkably successful, but we know that its theoretical foundation, general relativity, breaks down at high enough energies, $E \gtrsim M_{\text{fundamental}}$, where the fundamental scale at which new physics kicks in is usually taken to be the Planck scale, $M_{\text{fundamental}} = M_p \sim 10^{16}$ TeV, but could possibly be a lower scale, with particle collider constraints indicating that $M_{\text{fundamental}} \gtrsim 1$ TeV. Λ CDM can only provide limited insight into the very early universe. Indeed, the crucial role played by inflation belies the fact that inflation remains an effective theory without yet a basis in fundamental theory. A quantum gravity theory will be able to probe higher energies and earlier times, and should provide a consistent basis for inflation, or an alternative that replaces inflation within the standard cosmological model (for recent work, see e.g. Refs. [3–8]).

An even bigger theoretical problem than inflation is that of the late-time acceleration in the expansion of the universe [9–16]. In terms of the fundamental energy density parameters, $\Omega_{(i)} = \rho_{(i)}/3H_0^2$, we can rewrite the Friedmann equation using the conservation equations,

$$\left(\frac{H}{H_0} \right)^2 = (\Omega_{(c)} + \Omega_{(b)})(1+z)^3 + \Omega_{(r)}(1+z)^4 + \Omega_\Lambda + \Omega_K(1+z)^2, \quad (1.1.4)$$

where the redshift is $z = a^{-1} - 1$ with $a_0 = 1$ today. The data indicates that the present cosmic energy budget is given by

$$\Omega_\Lambda \approx 0.75, \quad \Omega_{(m)} \equiv \Omega_{(c)} + \Omega_{(b)} \approx 0.25, \quad |\Omega_K| \ll 1, \quad (1.1.5)$$

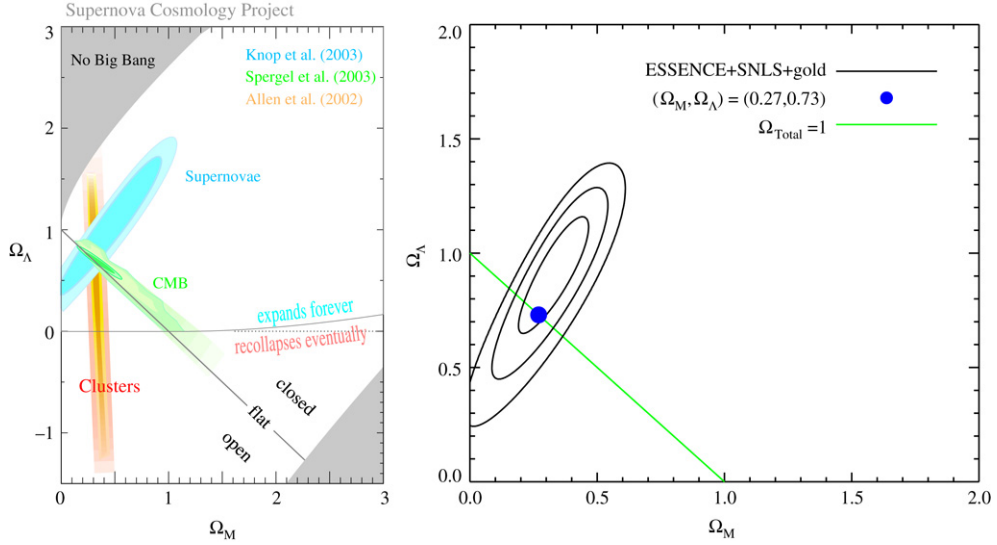


Fig. 1. Observational constraints in the $(\Omega_{(m)}, \Omega_{\Lambda})$ plane: joint constraints (left) (from [17]); recent compilation of supernova constraints (right) (from [18]).

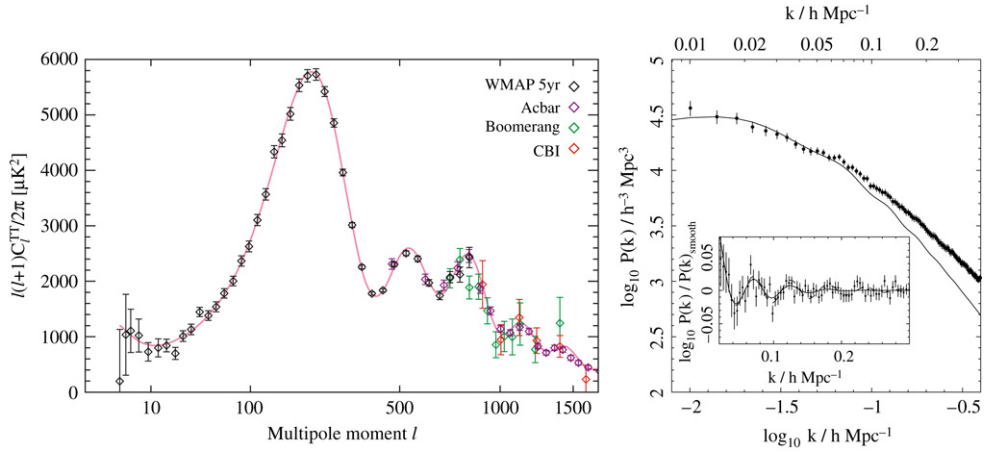


Fig. 2. Left: Power spectrum of CMB temperature anisotropies, showing data from WMAP5 [19], the 2003 flight of BOOMERANG [20], CBI [21] and the full ACBAR dataset [22]. The red line is the best-fit LCDM model to the data. Right: Matter power spectrum, showing data from the SDSS 2006 data release and the best-fit LCDM curve; the inset shows the imprint (in Fourier space) of the CMB acoustic peaks, known as the baryon acoustic oscillations (from [2]).

so that the Universe is currently accelerating, $\ddot{a}_0 > 0$, and (nearly) spatially flat.

Within the framework of general relativity, the acceleration typically originates from a dark energy field with negative pressure. In LCDM, this is the vacuum energy ($w_{\Lambda} = -1$), but dynamical dark energy fields have also been considered. For the simplest option of vacuum energy, the observed value of the cosmological constant is overwhelmingly smaller than the prediction of current particle physics. In particular,

$$\rho_{\Lambda, \text{obs}} = \Lambda \sim H_0^2 M_p^2 \sim (10^{-33} \text{ eV})^2 (10^{19} \text{ GeV})^2 \sim (10^{-3} \text{ eV})^4, \tag{1.1.6}$$

whereas

$$\rho_{\Lambda, \text{theory}} \sim M_{\text{fundamental}}^4 \gtrsim 1 \text{ TeV}^4 \gg \rho_{\Lambda, \text{obs}}. \tag{1.1.7}$$

In addition, the Λ value needs to be strongly fine-tuned to be of the same order of magnitude today as the current matter density, i.e.,

$$\rho_{\Lambda} \sim \rho_{(m)0} \Rightarrow \Omega_{\Lambda} \sim \Omega_{(m)}, \tag{1.1.8}$$

otherwise galaxies and then life could not emerge in the universe. The question is how this “coincidence” arises at late times, given that

$$\rho_{\Lambda} = \text{constant}, \quad \text{while } \rho_{(m)} \propto (1+z)^3. \tag{1.1.9}$$

No convincing or natural explanation has yet been proposed. String theory provides a tantalising possibility in the form of the “landscape” of vacua [23,24]. There appears to be a vast number of vacua admitted by string theory, with a broad range of energies above and below zero. The idea is that our observable region of the universe corresponds to a particular small positive vacuum energy, whereas other regions with greatly different vacuum energies will look entirely different. This multitude of regions forms in some sense a “multiverse”. This is an interesting idea, but it is highly speculative, and it is not clear how much of it will survive the further development of string theory and cosmology.

A different approach is based on the idea that there is no material dark energy field, but instead the Universe accelerates due to gravitational effects. Within general relativity, this has been proposed via nonlinear effects from structure formation [25–35]. As structure forms and the matter density perturbations become nonlinear, there are two questions that are posed: (1) what is the back-reaction effect of this nonlinear process on the background cosmology? (2) how do we perform a covariant and gauge-invariant averaging over the inhomogeneous universe to arrive at the correct FRW background? The simplistic answers to these questions are (1) the effect is negligible since it occurs on scales too small to be cosmologically significant; (2) in light of this, the background is independent of structure formation, i.e. it is the same as in the linear regime. A quantitative analysis is needed to fully resolve both issues. However, this is very complicated because it involves the nonlinear features of general relativity in an essential way.

There have been claims that these simplistic answers are wrong, and that, on the contrary, the effects are large enough to accelerate the universe. Note the possibility that averaging effects could be significant, even if they do not lead to acceleration. This would indeed be a dramatic and satisfying resolution of the coincidence problem, without the need for any dark energy field. Of course, this would not solve the problem of the vacuum energy, but would only re-define the problem as: why does the vacuum not gravitate? However, the claims for acceleration via nonlinear effects have been widely disputed, and it is fair to say that there is as yet no convincing demonstration that this is possible.

A more drastic form of a gravitational explanation for late-time acceleration, is that general relativity breaks down on the largest scales, and a modified gravity theory takes over on these scales. Schematically, this means modifying the geometric side of the field equations,

$$G_{ab} + G_{ab}^{(\text{dark})} = 8\pi G T_{ab}, \quad (1.1.10)$$

rather than the matter side,

$$G_{ab} = 8\pi G (T_{ab} + T_{ab}^{(\text{dark})}), \quad (1.1.11)$$

as in the standard general relativity approach. Modified gravity represents an intriguing possibility for resolving the theoretical crisis posed by late-time acceleration. However, it turns out to be extremely difficult to modify general relativity at low energies in cosmology, without violating the solar system constraints, or without introducing ghosts and other instabilities into the theory, or without altering the expansion rate in the matter-dominated era. Up to now, there is no convincing alternative to the general relativistic dark-energy models. It is indicative of the stir the supernovae observations have caused, that even the Copernican principle itself has been questioned [36].

In addition to the theoretical problems of inflation and dark energy, there is also the problem of “missing mass”, i.e. the fact that we cannot account for the observed matter power spectrum, given the observed CMB power spectrum, if we invoke only baryonic matter and use general relativistic dynamics. The general relativistic solution to this problem is non-baryonic CDM, dominating over baryonic matter at roughly 5 to 1. This solution simultaneously accounts for the rotation curves of spiral galaxies. Extensions of the Standard Model of particle physics predict various candidate particles for the cold dark matter, and a range of experiments is underway or planned to constrain, detect or rule out some of these.

A more radical approach to the missing mass problem is to reject non-baryonic CDM, and to propose, instead, a modification to gravity at low accelerations, similar in spirit to the modified gravity approach to the dark energy problem. Modified Newtonian dynamics (MOND) can account for the galactic rotation curves, and there are covariant relativistic modifications of general relativity that can reproduce MOND in the Newtonian limit [37,38]. These modified theories typically require both scalar and vector degrees of freedom in the gravitational field, in addition to the tensor. They lack a simple and natural motivation – much like most of the modified theories that are alternatives to dark energy. Future developments may lead to a low-energy modification of general relativity that does not require dark matter or dark energy, that preserves the successes of general relativity from solar system to cosmological scales, and that has some motivation in fundamental theory and a level of internal simplicity and naturalness.

In this review, we will not discuss further the key theoretical challenges posed by inflation, by the “missing mass” problem, and by the late-time acceleration of the Universe. Instead, we will adopt the standard view, based on a simple phenomenological model of inflation, on non-baryonic cold dark matter, and on the cosmological constant model of dark energy – with general relativity applying on all scales from the inflationary energy scale downwards. Our aim is to study the dynamics of structure formation within this framework and by means of a single formalism, the 1 + 3 covariant approach to cosmology. We begin with a comprehensive presentation of the covariant formalism, followed by a discussion of the standard cosmological model, the Friedmann–Lemaître universe. In Section 2, we analyse the nonlinear behaviour of a general cosmological spacetime, containing matter in the form of a single fluid, a mixture of interacting fluids, a minimally coupled scalar fields and in the presence of a large-scale magnetic field. We linearise the nonlinear formulae in Section 3 and then use them to discuss the key features and the evolution of perturbed Friedmannian models in various environments

and during different epochs. Section 4 provides an overview of the covariant kinetic theory, before applying it to the study of the cosmic microwave and neutrino backgrounds. We conclude this review with a brief summary of the currents trends in cosmological research and a look to the future in Section 6. Finally, in the Appendices, we provide the reader with all the technical information that is necessary for the detailed study of this manuscript.

1.2. The 1 + 3 covariant description

The covariant approach to general relativity and cosmology dates back to the work of Heckmann, Schücking, and Raychaudhuri in the 1950s [39,40] and it has since been employed in numerous applications by many authors (see [41–44] for details). The formalism uses the kinematic quantities, the energy–momentum tensor of the fluid(s) and the gravito-electromagnetic parts of the Weyl tensor, instead of the metric, which in itself does not provide a covariant description. The key equations are the Ricci and Bianchi identities, applied to the fluid 4-velocity vector, while Einstein's equations are incorporated via algebraic relations between the Ricci and the energy–momentum tensor.

1.2.1. Local spacetime splitting

Consider a general spacetime with a Lorentzian metric g_{ab} of signature $(-, +, +, +)$ and introduce a family of observers with worldlines tangent to the timelike 4-velocity vector¹

$$u^a = \frac{dx^a}{d\tau}, \quad (1.2.1)$$

where τ is the observers' proper time, so that $u_a u^a = -1$. This fundamental velocity field introduces a local 1 + 3 'threading' of the spacetime into time and space. The vector u_a determines the time direction, while the tensor $h_{ab} = g_{ab} + u_a u_b$ projects orthogonal to the 4-velocity into the observers' instantaneous rest space at each event. In the absence of vorticity the 4-velocity is hypersurface-orthogonal and h_{ab} is the metric of the 3-dimensional spatial sections orthogonal to u_a .

The vector field u_a and its tensor counterpart h_{ab} allow for a unique decomposition of every spacetime quantity into its irreducible timelike and spacelike parts. These fields are also used to define the covariant time and spatial derivatives of any tensor field $S_{ab\dots}{}^{cd\dots}$ according to

$$\dot{S}_{ab\dots}{}^{cd\dots} = u^e \nabla_e S_{ab\dots}{}^{cd\dots} \quad \text{and} \quad D_e S_{ab\dots}{}^{cd\dots} = h_e^s h_a^f h_b^p h_c^q h_r^d \dots \nabla_s S_{fp\dots}{}^{qr\dots}, \quad (1.2.2)$$

respectively.

The effective volume element in the observer's instantaneous rest space is given by contracting the spacetime volume element (η_{abcd}) along the time direction,

$$\varepsilon_{abc} = \eta_{abcd} u^d. \quad (1.2.3)$$

The totally antisymmetric pseudotensor η_{abcd} has $\eta^{0123} = [-\det(g_{ab})]^{-1/2}$, it is covariantly constant and satisfies the identities $\eta_{abcd} \eta^{efpq} = -4! \delta_{[a}^e \delta_b^f \delta_c^p \delta_{d]}^q$. It follows that $\varepsilon_{abc} u^a = 0$,

$$\eta_{abcd} = 2u_{[a} \varepsilon_{b]cd} - 2\varepsilon_{ab[c} u_{d]} \quad \text{and} \quad \varepsilon_{abc} \varepsilon^{def} = 3! h_{[a}^d h_b^e h_{c]}^f. \quad (1.2.4)$$

Note that $D_c h_{ab} = 0 = D_d \varepsilon_{abc}$, while $\dot{h}_{ab} = 2u_{(a} A_{b)}$ and $\dot{\varepsilon}_{abc} = 3u_{[a} \varepsilon_{b]c]d} A^d$ (with $A_a = \dot{u}_a$ – see Section 1.3.1 below).

1.2.2. The gravitational field

In the general relativistic geometrical interpretation of gravity, matter determines the spacetime curvature, while the latter dictates the motion of the matter. This interaction is realised via the Einstein field equations,

$$G_{ab} \equiv R_{ab} - \frac{1}{2} R g_{ab} = T_{ab} - \Lambda g_{ab}, \quad (1.2.5)$$

where G_{ab} is the Einstein tensor, $R_{ab} = R_{acb}{}^c$ is the spacetime Ricci tensor (with trace R), T_{ab} is the total energy–momentum tensor of the matter fields and Λ is the cosmological constant. The twice contracted Bianchi identities guarantee that $\nabla^b T_{ab} = 0$ and total energy–momentum conservation.

The Ricci tensor describes the local gravitational field at each event due to matter there. The non-local, long-range gravitational field, mediated via gravitational waves and tidal forces, is encoded in the Weyl conformal curvature tensor C_{abcd} . The splitting of the gravitational field into its local and non-local parts is given by the decomposition of the Riemann tensor,

$$R_{abcd} = C_{abcd} + \frac{1}{2} (g_{ac} R_{bd} + g_{bd} R_{ac} - g_{bc} R_{ad} - g_{ad} R_{bc}) - \frac{1}{6} R (g_{ac} g_{bd} - g_{ad} g_{bc}), \quad (1.2.6)$$

¹ Latin indices vary between 0 and 3 and refer to arbitrary coordinate or tetrad frames. Greek indices run from 1 to 3. We use geometrised units with $c = 1 = 8\pi G$, which means that all geometrical variables have physical dimensions that are integer powers of length.

where the Weyl tensor shares all the symmetries of the Riemann tensor and is also trace-free, $C^c{}_{acb} = 0$. Relative to the fundamental observers, the conformal curvature tensor decomposes further into its irreducible parts (e.g. see [45,46])

$$E_{ab} = C_{acbd}u^c u^d \quad \text{and} \quad H_{ab} = \frac{1}{2} \varepsilon_a{}^{cd} C_{cdbe} u^e. \quad (1.2.7)$$

Then,

$$C_{abcd} = (\mathcal{G}_{abpq}\mathcal{G}_{cdsr} - \eta_{abqp}\eta_{cdsr}) u^q u^s E^{pr} - (\eta_{abqp}\mathcal{G}_{cdsr} + \mathcal{G}_{abqp}\eta_{cdsr}) u^q u^s H^{pr}, \quad (1.2.8)$$

where $\mathcal{G}_{abcd} = \mathcal{G}_{ac}\mathcal{G}_{bd} - \mathcal{G}_{ad}\mathcal{G}_{bc}$. Alternatively,

$$C_{ab}{}^{cd} = 4 \left(u_{[a} u^{[c} + h_{[a}{}^{[c} \right) E_{b]}{}^{d]} + 2\varepsilon_{abe} u^{[c} H^{d]e} + 2u_{[a} H_{b]}{}^e \varepsilon^{cde}. \quad (1.2.9)$$

The spatial, symmetric and trace-free tensors E_{ab} and H_{ab} are known as the electric and magnetic Weyl components. The electric part generalises the tidal tensor of the Newtonian gravitational potential, but H_{ab} has no Newtonian counterpart. Note that both tensors must be present for a nonzero super-energy flux vector ($P_a = \varepsilon_{abc} E^{bd} H^c{}_d$), which is essential for the propagation of gravitational waves.

The Weyl tensor represents the part of the curvature that is not determined locally by matter. However, its dynamics are not arbitrary because the Riemann tensor satisfies the Bianchi identities, whose contraction gives [45]

$$\nabla^d C_{abcd} = \nabla_{[b} R_{a]c} + \frac{1}{6} \mathcal{G}_{c[b} \nabla_{a]} R, \quad (1.2.10)$$

using decomposition (1.2.6). In a sense the once contracted Bianchi identities act as the field equations for the Weyl tensor, determining the part of the spacetime curvature that depends on the matter distribution at other points. Eq. (1.2.10) splits into a set of two propagation and two constraint equations, which govern the dynamics of the electric and magnetic Weyl components (see Section 1.3.6).

1.2.3. Matter fields

With respect to the fundamental observers, the energy–momentum tensor of a general (imperfect) fluid decomposes into its irreducible parts as²

$$T_{ab} = \rho u_a u_b + p h_{ab} + 2q_{(a} u_{b)} + \pi_{ab}, \quad (1.2.11)$$

where $\rho = T_{ab} u^a u^b$ is the matter energy density, $p = T_{ab} h^{ab}/3$ is the effective isotropic pressure of the fluid, namely the sum between the equilibrium pressure and the associated bulk viscosity, $q_a = -h_a{}^b T_{bc} u^c$ is the total energy–flux vector, and $\pi_{ab} = h_{(a}{}^c h_{b)}{}^d T_{cd}$ is the symmetric and trace-free anisotropic stress tensor.³

The 4-velocity u_a is generally arbitrary and a velocity boost of the form $u_a \rightarrow \tilde{u}_a$ induces changes in the dynamical quantities, given explicitly in Appendix A.2. When the fluid is perfect, however, there is a unique hydrodynamic 4-velocity, relative to which q_a, π_{ab} are identically zero and the effective pressure reduces to the equilibrium one. As a result,

$$T_{ab} = \rho u_a u_b + p h_{ab}. \quad (1.2.13)$$

If we additionally assume that $p = 0$, we have the simplest case of pressure-free matter, namely ‘dust’, which includes baryonic matter (after decoupling) and cold dark matter. Otherwise, we need to determine p as a function of ρ and potentially of other thermodynamic variables. In general, the equation of state takes the form $p = p(\rho, s)$, where s is the specific entropy. Finally, for a barotropic medium we have $p = p(\rho)$ (see Section 1.3.4 for further discussion).

Expression (1.2.11) describes any type of matter, including electromagnetic fields, scalar fields, etc. (see Sections 1.3.2 and 2.6.1). Since $R = 4\Lambda - T$, with $T = T_a{}^a$, Einstein’s equations are recast into

$$R_{ab} = T_{ab} - \frac{1}{2} T g_{ab} + \Lambda g_{ab}. \quad (1.2.14)$$

The successive contraction of the above, assuming that T_{ab} is given by Eq. (1.2.11), leads to a set of algebraic relations that will prove useful later

$$R_{ab} u^a u^b = \frac{1}{2} (\rho + 3p) - \Lambda, \quad (1.2.15)$$

$$h_a{}^b R_{bc} u^c = -q_a, \quad (1.2.16)$$

$$h_a{}^c h_b{}^d R_{cd} = \frac{1}{2} (\rho - p) h_{ab} + \Lambda h_{ab} + \pi_{ab}. \quad (1.2.17)$$

² For a multi-component medium, or when allowing for peculiar velocities, one needs to account for the differing 4-velocities of the matter components and the fundamental observers (see Section 2.4).

³ Angled brackets denote the symmetric and trace-free part of spatially projected second-rank tensors and the projected part of vectors according to

$$S_{(ab)} = h_{(a}{}^c h_{b)}{}^d S_{cd} = h_{(a}{}^c h_{b)}{}^d S_{cd} - \frac{1}{3} h^{cd} S_{cd} h_{ab} \quad \text{and} \quad V_{(a)} = h_a{}^b V_b, \quad (1.2.12)$$

respectively (with $S_{(ab)} h^{ab} = 0$). The reader is referred to the Appendix (see Appendix A.1 there) for more details on covariant decomposition.

1.3. Covariant relativistic cosmology

There are various physical choices in cosmology for the fundamental 4-velocity field that defines the 1 + 3 splitting of spacetime. Some possibilities include the frame in which the dipole of the CMB anisotropy vanishes and the local rest-frame of the matter (these are generally assumed to coincide when averaged on sufficiently large scales). In specific situations, it may be appropriate to choose the frame that simplifies the physics (for example, for a perfect-fluid cosmology it makes sense to adopt the rest-frame of the fluid), and we shall make several choices for u_a throughout this review. Once u_a is specified, its integral curves define the worldlines of the fundamental observers introduced in Section 1.2.1.

1.3.1. Kinematics

The observers' motion is characterised by the irreducible kinematical quantities of the u_a -congruence, which emerge from the covariant decomposition of the 4-velocity gradient

$$\nabla_b u_a = \sigma_{ab} + \omega_{ab} + \frac{1}{3} \Theta h_{ab} - A_a u_b, \quad (1.3.1)$$

where $\sigma_{ab} = D_{(b} u_{a)}$, $\omega_{ab} = D_{[b} u_{a]}$, $\Theta = \nabla^a u_a = D^a u_a$ and $A_a = \dot{u}_a = u^b \nabla_b u_a$ are, respectively, the shear and the vorticity tensors, the volume expansion (or contraction) scalar, and the 4-acceleration vector. The latter represents non-gravitational forces and vanishes when matter moves under gravity alone. By construction we have $\sigma_{ab} u^a = 0 = \omega_{ab} u^a = A_a u^a$. Also, on using the orthogonally projected alternating tensor ε_{abc} (with $\dot{\varepsilon}_{abc} = 3u_{[a} \varepsilon_{bc]d} A^d$), one defines⁴ the vorticity vector $\omega_a = \varepsilon_{abc} \omega^{bc} / 2$ (with $\omega_{ab} = \varepsilon_{abc} \omega^c$). We note that the tensor $v_{ab} = D_b u_a = \sigma_{ab} + \omega_{ab} + (\Theta/3)h_{ab}$ describes the relative motion of neighbouring observers. In particular, $v_a = v_{ab} \chi^b$ monitors the relative velocity between the observers' worldlines, with χ_a representing the relative position vector between the same two flow lines (e.g. see [41,42] for details). The volume scalar determines the average separation between neighbouring observers and is also used to introduce a representative length scale (a) by means of the definition $\dot{a}/a = \Theta/3$. The effect of the vorticity is to change the orientation of a given fluid element without modifying its volume or shape. The shear, on the other hand, changes the shape but leaves the volume unaffected.

The non-linear covariant kinematics are determined by a set of propagation and constraint equations, which are purely geometrical in origin and essentially independent of the Einstein equations. Both sets emerge after applying the Ricci identities

$$2\nabla_{[a} \nabla_{b]} u_c = R_{abcd} u^d, \quad (1.3.2)$$

to the fundamental 4-velocity vector defined in (1.2.1). Substituting in from (1.3.1), using decompositions (1.2.6) and (1.2.8) and the auxiliary relations (1.2.15)–(1.2.17), the timelike and spacelike parts of the resulting expression lead to a set of three propagation and three constraint equations. The former contains Raychaudhuri's formula

$$\dot{\Theta} = -\frac{1}{3} \Theta^2 - \frac{1}{2} (\rho + 3p) - 2(\sigma^2 - \omega^2) + D^a A_a + A_a A^a + \Lambda, \quad (1.3.3)$$

for the time evolution of Θ ; the shear propagation equation

$$\dot{\sigma}_{(ab)} = -\frac{2}{3} \Theta \sigma_{ab} - \sigma_{c(a} \sigma^c_{b)} - \omega_{(a} \omega_{b)} + D_{(a} A_{b)} + A_{(a} A_{b)} - E_{ab} + \frac{1}{2} \pi_{ab}, \quad (1.3.4)$$

which describes kinematical anisotropies; and the evolution equation of the vorticity

$$\dot{\omega}_{(a)} = -\frac{2}{3} \Theta \omega_a - \frac{1}{2} \text{curl} A_a + \sigma_{ab} \omega^b. \quad (1.3.5)$$

Note that $\sigma^2 = \sigma_{ab} \sigma^{ab} / 2$ and $\omega^2 = \omega_{ab} \omega^{ab} / 2 = \omega_a \omega^a$ are, respectively, the scalar square magnitudes of the shear and the vorticity, while E_{ab} is the electric component of the Weyl tensor (see Section 1.2.2). Also, $\text{curl} v_a = \varepsilon_{abc} D^b v^c$ for any orthogonally projected vector v_a , which means that $D^b \omega_{ab} = \text{curl} \omega_a$.

The spacelike component of (1.3.2) leads to a set of three complementary constraints. These are the shear or $(0, \alpha)$ constraint

$$D^b \sigma_{ab} = \frac{2}{3} D_a \Theta + \text{curl} \omega_a + 2\varepsilon_{abc} A^b \omega^c - q_a, \quad (1.3.6)$$

the vorticity-divergence identity

$$D^a \omega_a = A_a \omega^a, \quad (1.3.7)$$

⁴ The sign conventions are such that $\vec{\omega} = -\vec{\nabla} \times \vec{v} / 2$ in the Newtonian limit and agree with those in the majority of the related articles. Note that in [44] the vorticity tensor (ω_{ab}) and the orthogonally projected volume element (ε_{abc}) have opposite signs, relative to the ones defined here. The reader should have this in mind when comparing the equations of the two papers.

and the magnetic Weyl equation

$$H_{ab} = \text{curl } \sigma_{ab} + D_{(a}\omega_{b)} + 2A_{(a}\omega_{b)}. \quad (1.3.8)$$

Raychaudhuri's formulae (see [40] and also [47,48] for recent reviews) is the key to the study of gravitational collapse, as it describes the evolution of the average separation between two neighbouring observers. For this reason Eq. (1.3.3) has been at the core of all the singularity theorems (see [45,49] and references therein). Negative terms in the right-hand side of (1.3.3) lead to contraction and positive resist the collapse, which means that conventional (non-phantom) matter is always attractive unless $p < -\rho/3$.

1.3.2. Electromagnetic fields

The Maxwell field is determined by the antisymmetric electromagnetic (Faraday) tensor F_{ab} , which relative to a fundamental observer decomposes into an electric and a magnetic component as [42,50]

$$F_{ab} = 2u_{[a}E_{b]} + \varepsilon_{abc}B^c. \quad (1.3.9)$$

In the above, $E_a = F_{ab}u^b$ and $B_a = \varepsilon_{abc}F^{bc}/2$ are respectively the electric and magnetic fields experienced by the observer (with $E_a u^a = 0 = B_a u^a$). The Faraday tensor also determines the energy–momentum tensor of the Maxwell field according to

$$T_{ab}^{(em)} = -F_{ac}F^c{}_b - \frac{1}{4}F_{cd}F^{cd}g_{ab}, \quad (1.3.10)$$

in Heaviside-Lorentz units. The above expression combines with (1.3.9) to give the irreducible decomposition for $T_{ab}^{(em)}$, relative to the u_a -frame

$$T_{ab}^{(em)} = \frac{1}{2}(E^2 + B^2)u_a u_b + \frac{1}{6}(E^2 + B^2)h_{ab} + 2\mathcal{P}_{(a}u_{b)} + \Pi_{ab}. \quad (1.3.11)$$

Here $E^2 = E_a E^a$ and $B^2 = B_a B^a$ are the square magnitudes of the two fields, $\mathcal{P}_a = \varepsilon_{abc}E^b B^c$ is the electromagnetic Poynting vector and $\Pi_{ab} = -E_{(a}E_{b)} - B_{(a}B_{b)}$. Expression (1.3.11) allows for a fluid description of the electromagnetic field and manifests its generically anisotropic nature. In particular, the Maxwell field corresponds to an imperfect fluid with energy density $(E^2 + B^2)/2$, isotropic pressure $(E^2 + B^2)/6$, anisotropic stresses given by Π_{ab} and an energy–flux vector represented by \mathcal{P}_a . Eq. (1.3.11) also ensures that $T_a^{(em)a} = 0$, in agreement with the trace-free nature of the radiation stress–energy tensor.

We follow the evolution of the electromagnetic field by means of Maxwell's equations. In their standard tensor form these read

$$\nabla_{[c}F_{ab]} = 0 \quad \text{and} \quad \nabla^b F_{ab} = J_a, \quad (1.3.12)$$

where (1.3.12)(a) reflects the existence of a 4-potential and J_a is the 4-current that sources the electromagnetic field. With respect to the u_a -congruence, the 4-current splits into its irreducible parts according to

$$J_a = \mu u_a + \mathcal{J}_a, \quad (1.3.13)$$

with $\mu = -J_a u^a$ representing the charge density and $\mathcal{J}_a = h_a{}^b J_b$ the orthogonally projected current (i.e. $\mathcal{J}_a u^a = 0$).

Relative to a fundamental observer, each one of Maxwell's equations decomposes into a timelike and a spacelike component. The timelike parts of (1.3.12)(a), (b) lead to a set of two propagation equations

$$\dot{E}_{(a)} = -\frac{2}{3}\Theta E_a + (\sigma_{ab} + \varepsilon_{abc}\omega^c)E^b + \varepsilon_{abc}A^b B^c + \text{curl } B_a - \mathcal{J}_a, \quad (1.3.14)$$

$$\dot{B}_{(a)} = -\frac{2}{3}\Theta B_a + (\sigma_{ab} + \varepsilon_{abc}\omega^c)B^b - \varepsilon_{abc}A^b E^c - \text{curl } E_a, \quad (1.3.15)$$

while their spacelike components provide the constraints

$$D^a E_a + 2\omega^a B_a = \mu \quad \text{and} \quad D^a B_a - 2\omega^a E_a = 0. \quad (1.3.16)$$

Expressions (1.3.14)–(1.3.16) are 1 + 3 covariant versions of Ampère's law, Faraday's law, Coulomb's law and Gauss' law respectively. Therefore, in addition to the usual 'curl' and 'divergence' terms, the covariant form of (1.3.14) and (1.3.15) contains terms generated by the relative motion of the neighbouring observers. Also, according to (1.3.16), the magnetic vector is not solenoidal unless $\omega^a E_a = 0$.

1.3.3. Conservation laws

The twice contracted Bianchi identities guarantee the conservation of the total energy momentum tensor, namely that $\nabla^b T_{ab} = 0$. This constraint splits into a timelike and a spacelike part, which respectively lead to the energy and the momentum conservation laws. When dealing with a general imperfect fluid, the former is

$$\dot{\rho} = -\Theta(\rho + p) - D^a q_a - 2A^a q_a - \sigma^{ab} \pi_{ab}, \quad (1.3.17)$$

while the latter satisfies the expression

$$(\rho + p)A_a = -D_a p - \dot{q}_{(a)} - \frac{4}{3} \Theta q_a - (\sigma_{ab} + \omega_{ab})q^b - D^b \pi_{ab} - \pi_{ab} A^b. \quad (1.3.18)$$

When the fluid is perfect, the energy–momentum tensor is given by (1.2.13) and the above reduce to

$$\dot{\rho} = -\Theta(\rho + p) \quad \text{and} \quad (\rho + p)A_a = -D_a p, \quad (1.3.19)$$

respectively. It follows, from (1.3.19)(b), that the sum $\rho + p$ describes the relativistic total inertial mass of the medium. Then, when the inertial mass is zero (i.e. for $\rho + p \rightarrow 0$), consistency demands that the pressure gradients must also vanish. For a barotropic fluid the latter immediately implies zero density gradients as well. Also, if the ‘phantom divide’ is crossed [51, 52], the inertial mass becomes negative and the acceleration antiparallel to the force that caused it.⁵

The energy momentum tensor of the electromagnetic field obeys the constraint $\nabla^b T_{ab}^{(em)} = -F_{ab} j^b$, with the Faraday tensor given by (1.3.9) and the quantity in the right-hand side representing the Lorentz 4-force. Thus, for charged matter the conservation of the total energy–momentum tensor $T_{ab} = T_{ab}^{(m)} + T_{ab}^{(em)}$ leads to the formulae

$$\dot{\rho} = -\Theta(\rho + p) - D^a q_a - 2A^a q_a - \sigma^{ab} \pi_{ab} + E_a \mathcal{J}^a \quad (1.3.20)$$

for the energy density, and

$$(\rho + p)A_a = -D_a p - \dot{q}_{(a)} - \frac{4}{3} \Theta q_a - (\sigma_{ab} + \omega_{ab})q^b - D^b \pi_{ab} - \pi_{ab} A^b + \mu E_a + \varepsilon_{abc} \mathcal{J}^b B^c, \quad (1.3.21)$$

for the momentum density. We note the electromagnetic terms in the right-hand side of the above, the effect of which depends on the electrical properties of the medium (see Section 2.5.1). The last two terms in the right-hand side of (1.3.21), in particular, represent the more familiar form of the Lorentz force.

The antisymmetry of the Faraday tensor (see Eq. (1.3.9)) and the second of Maxwell’s formulae (see Eq. (1.3.12)(b)) imply $\nabla^a J_a = 0$ and the conservation of the 4-current. Then, on using decomposition (1.3.13), we arrive at the covariant charge-density conservation law

$$\dot{\mu} = -\Theta \mu - D^a \mathcal{J}_a - A^a \mathcal{J}_a. \quad (1.3.22)$$

Hence, in the absence of spatial currents, the evolution of the charge density depends entirely on the volume expansion (or contraction) of the fluid element.

1.3.4. Equilibrium thermodynamics

In relativistic thermodynamics the physical state of a medium is monitored by means of the energy momentum tensor (T_{ab}), the particle flux vector (N_a) and the entropy flux vector (S_a) (e.g. see [61]). For isolated systems, the former of these three ‘primary variables’ satisfies the conservation law $\nabla^b T_{ab} = 0$. The entropy flux, on the other hand, obeys the second law of thermodynamics and, provided the particle number does not change, N_a satisfies the particle number conservation law. Written in covariant terms, these read

$$\nabla^a S_a \geq 0 \quad \text{and} \quad \nabla^a N_a = 0, \quad (1.3.23)$$

respectively. For a system in equilibrium (or for a perfect fluid) there is no entropy production, which implies that

$$\nabla^a S_a = 0. \quad (1.3.24)$$

Moreover, all three of the aforementioned primary variables can be expressed in terms of a unique timelike 4-velocity field, according to

$$T_{ab} = \rho u_a u_b + p h_{ab}, \quad S_a = S u_a \quad \text{and} \quad N_a = n u_a, \quad (1.3.25)$$

where $S = -S_a u^a$ and $n = -N_a u^a$ are the entropy and particle densities respectively. Note that the last two of the above combine to give

$$S_a = s N_a, \quad (1.3.26)$$

with $s = S/n$ representing the specific entropy (the entropy per particle) of the system.

⁵ Phantom cosmologies violate the dominant energy condition (i.e. have $\rho + p < 0$) and generally lead to future ‘big rip’ singularities (see [51,52] and also [53–58]). On the other hand, finite-time (sudden) future singularities can arise even when the standard energy conditions are observed [59,60].

Applying the conservation law $\nabla^b T_{ab} = 0$ to the energy–momentum tensor (1.3.25)(a), leads to the familiar energy and momentum density conservation laws of a perfect fluid (given in Section 1.3.3 by expressions (1.3.17) and (1.3.18) respectively). On the other hand, substituting (1.3.25)(c) into the left-hand side of (1.3.23)(b) provides the conservation equation of the particle number

$$\dot{n} = -\Theta n. \quad (1.3.27)$$

Similarly, inserting (1.3.26) into entropy conservation law (1.3.24) and then using Eqs. (1.3.23)(b) and (1.3.25)(c) we arrive at

$$\dot{s} = 0. \quad (1.3.28)$$

This ensures that the specific entropy of the system does not change along the fluid motion, which is another way of saying that the flow is adiabatic. When the spatial gradients of the specific entropy also vanish, we have $\nabla_a s = D_a s - \dot{s} u_a = 0$ and the medium is said to be isentropic. Note that an isentropic fluid has a barotropic equation of state and vice versa [62].

An additional thermodynamic scalar is the temperature (T) of the system, which satisfies the Gibbs equation

$$T ds = d\left(\frac{\rho}{n}\right) + p d\left(\frac{1}{n}\right), \quad (1.3.29)$$

with $df = dx^a \nabla_a f$ (e.g. see [63,62]). Of the five thermodynamic scalars (ρ , p , n , s and T), two are needed as independent variables. Selecting the energy density and the specific entropy as our independent quantities, the equation of state of a (perfect) fluid acquires the form $p = p(\rho, s)$. Then,

$$\dot{p} = \left(\frac{\partial p}{\partial \rho}\right)_s \dot{\rho} + \left(\frac{\partial p}{\partial s}\right)_\rho \dot{s}, \quad (1.3.30)$$

which (for $\dot{s} = 0$) gives

$$c_s^2 \equiv \left(\frac{\partial p}{\partial \rho}\right)_s = \frac{\dot{p}}{\dot{\rho}}, \quad (1.3.31)$$

namely the square of the adiabatic sound speed. In addition to (1.3.30), the above given equation of state also leads to

$$D_a p = \left(\frac{\partial p}{\partial \rho}\right)_s D_a \rho + \left(\frac{\partial p}{\partial s}\right)_\rho D_a s, \quad (1.3.32)$$

thus connecting the spatial gradients of the pressure to perturbations in the energy density and the specific entropy of the system. Consequently, when applied to adiabatic and then to isentropic systems, the relations (1.3.31) and (1.3.32) combine to give

$$D_a p = \left(\frac{\dot{p}}{\dot{\rho}}\right) D_a \rho + \left(\frac{\partial p}{\partial s}\right)_\rho D_a s \quad \text{and} \quad D_a p = \left(\frac{\dot{p}}{\dot{\rho}}\right) D_a \rho, \quad (1.3.33)$$

respectively. At this point we note that, according to the standard thermodynamic nomenclature, we distinguish between adiabatic and isentropic perturbations. The former are characterised by $\dot{s} = 0$, while the latter demand that the specific entropy is a spacetime invariant. Although the two concepts are distinct, it is not uncommon for cosmologists to say adiabatic and imply isentropic [64].

1.3.5. Spatial curvature

When the fluid flow is irrotational, the rest-space tangent planes of the fundamental observers mesh together to form spacelike hypersurfaces orthogonal to their worldlines. These are normal to the u_a -congruence and define the hypersurfaces of simultaneity for all the comoving observers. In the presence of vorticity, however, Frobenius' theorem forbids the existence of such integrable hypersurfaces (e.g. see [49,65]). Then the observers' rest-spaces no longer mesh together smoothly. The projected Riemann tensor is defined by

$$\mathcal{R}_{abcd} = h_a^q h_b^s h_c^f h_d^p R_{qsfp} - v_{ac} v_{bd} + v_{ad} v_{bc}, \quad (1.3.34)$$

where $v_{ab} = D_b u_a$ is the relative flow tensor between two neighbouring observers (see Section 1.3.1). On using Eqs. (1.2.5)–(1.2.15) and decompositions (1.2.6), (1.2.8), we find [66]

$$\begin{aligned} \mathcal{R}_{abcd} = & -\varepsilon_{abq} \varepsilon_{cds} E^{qs} + \frac{1}{3} \left(\rho - \frac{1}{3} \Theta^2 + \Lambda \right) (h_{ac} h_{bd} - h_{ad} h_{bc}) + \frac{1}{2} (h_{ac} \pi_{bd} + \pi_{ac} h_{bd} - h_{ad} \pi_{bc} - \pi_{ad} h_{bc}) \\ & - \frac{1}{3} \Theta [h_{ac} (\sigma_{bd} + \omega_{bd}) + (\sigma_{ac} + \omega_{ac}) h_{bd} - h_{ad} (\sigma_{bc} + \omega_{bc}) - (\sigma_{ad} + \omega_{ad}) h_{bc}] \\ & - (\sigma_{ac} + \omega_{ac}) (\sigma_{bd} + \omega_{bd}) + (\sigma_{ad} + \omega_{ad}) (\sigma_{bc} + \omega_{bc}). \end{aligned} \quad (1.3.35)$$

This provides an irreducible decomposition of the projected Riemann tensor. If $\omega_a = 0$, then \mathcal{R}_{abcd} is the 3-Riemann tensor of the hypersurfaces of simultaneity orthogonal to u_a . In analogy to 4-dimensions, the projected Ricci tensor and Ricci scalar are respectively defined by

$$\mathcal{R}_{ab} = h^{cd} \mathcal{R}_{acbd} = \mathcal{R}^c{}_{acb} \quad \text{and} \quad \mathcal{R} = h^{ab} \mathcal{R}_{ab}. \quad (1.3.36)$$

The algebraic symmetries of \mathcal{R}_{abcd} are given by

$$\mathcal{R}_{abcd} = \mathcal{R}_{[ab][cd]} \quad (1.3.37)$$

and

$$\mathcal{R}_{abcd} - \mathcal{R}_{cdab} = -\frac{2}{3} \Theta (h_{ac}\omega_{bd} + \omega_{ac}h_{bd} - h_{ad}\omega_{bc} - \omega_{ad}h_{bc}) - 2(\sigma_{ac}\omega_{bd} + \omega_{ac}\sigma_{bd} - \sigma_{ad}\omega_{bc} - \omega_{ad}\sigma_{bc}). \quad (1.3.38)$$

It follows from the above that $\mathcal{R}_{abcd} = \mathcal{R}_{cdab}$ in the absence of vorticity. In that case, the spatial Riemann tensor possesses all the symmetries of its 4-dimensional counterpart.

Contracting (1.3.35) on the first and third indices we arrive at what is usually referred to as the Gauss–Codacci formula

$$\mathcal{R}_{ab} = E_{ab} + \frac{2}{3} \left(\rho - \frac{1}{3} \Theta^2 + \sigma^2 - \omega^2 + \Lambda \right) h_{ab} + \frac{1}{2} \pi_{ab} - \frac{1}{3} \Theta (\sigma_{ab} + \omega_{ab}) + \sigma_{c(a} \sigma^c{}_{b)} - \omega_{c(a} \omega^c{}_{b)} + 2\sigma_{c[a} \omega^c{}_{b]}, \quad (1.3.39)$$

while a further contraction leads to the generalised Friedmann equation

$$\mathcal{R} = h^{ab} \mathcal{R}_{ab} = 2 \left(\rho - \frac{1}{3} \Theta^2 + \sigma^2 - \omega^2 + \Lambda \right). \quad (1.3.40)$$

Finally, one may combine Eqs. (1.3.39) and (1.3.40) to obtain

$$\mathcal{R}_{ab} = \frac{1}{3} \mathcal{R} h_{ab} + E_{ab} + \frac{1}{2} \pi_{ab} - \frac{1}{3} \Theta (\sigma_{ab} + \omega_{ab}) + \sigma_{c(a} \sigma^c{}_{b)} - \omega_{c(a} \omega^c{}_{b)} + 2\sigma_{c[a} \omega^c{}_{b]}, \quad (1.3.41)$$

where all terms on the right, with the exception of the first, are trace-free.

It should be noted that the matter variables used in this section represent the total fluid. For example, when dealing with a mixture of pressure-free dust and isotropic radiation $\rho = \rho^{(d)} + \rho^{(\gamma)}$ and $\pi_{ab} = 0$, in the presence of an electromagnetic field (see Section 1.3.2) we have $\rho = \rho^{(m)} + \rho^{(em)} = \rho^{(m)} + (E^2 + B^2)/2$, $\pi_{ab} = \pi_{ab}^{(m)} + I_{ab}$, etc.

1.3.6. Weyl curvature

The 1 + 3 splitting of the once contracted Bianchi identities (see Eq. (1.2.10) in Section 1.2.2) leads to a set of two propagation and two constraint equations that monitor the evolution of the long range gravitational field, namely tidal forces and gravity waves. In particular, on using the decomposition (1.2.8), the timelike component of (1.2.10) leads to [44]

$$\begin{aligned} \dot{E}_{(ab)} = & -\Theta E_{ab} - \frac{1}{2} (\rho + p) \sigma_{ab} + \text{curl} H_{ab} - \frac{1}{2} \dot{\pi}_{ab} - \frac{1}{6} \Theta \pi_{ab} - \frac{1}{2} D_{(a} q_{b)} - A_{(a} q_{b)} \\ & + 3\sigma_{(a}{}^c \left(E_{b)c} - \frac{1}{6} \pi_{b)c} \right) + \varepsilon_{cd(a} \left[2A^c H_{b)}{}^d - \omega^c \left(E_{b)}{}^d + \frac{1}{2} \pi_{b)}{}^d \right) \right] \end{aligned} \quad (1.3.42)$$

and

$$\dot{H}_{(ab)} = -\Theta H_{ab} - \text{curl} E_{ab} + \frac{1}{2} \text{curl} \pi_{ab} + 3\sigma_{(a}{}^c H_{b)c} - \frac{3}{2} \omega_{(a} q_{b)} - \varepsilon_{cd(a} \left(2A^c E_{b)}{}^d - \frac{1}{2} \sigma^c{}_{b)} q^d + \omega^c H_{b)}{}^d \right). \quad (1.3.43)$$

Taking the time derivatives of the above, one arrives at a pair of wavelike equations for the electric and the magnetic parts of the Weyl tensor, showing how curvature distortions propagate in the form of gravitational waves like ripples in the spacetime fabric. These waves are also subjected to a set of constraints, which emerge from the spacelike component of the decomposed Eq. (1.2.10) and are given by

$$D^b E_{ab} = \frac{1}{3} D_a \rho - \frac{1}{2} D^b \pi_{ab} - \frac{1}{3} \Theta q_a + \frac{1}{2} \sigma_{ab} q^b - 3H_{ab} \omega^b + \varepsilon_{abc} \left(\sigma^b{}_d H^{cd} - \frac{3}{2} \omega^b q^c \right) \quad (1.3.44)$$

and

$$D^b H_{ab} = (\rho + p) \omega_a - \frac{1}{2} \text{curl} q_a + 3E_{ab} \omega^b - \frac{1}{2} \pi_{ab} \omega^b - \varepsilon_{abc} \sigma^b{}_d \left(E^{cd} + \frac{1}{2} \pi^{cd} \right), \quad (1.3.45)$$

respectively [44]. The above expressions are remarkably similar to Maxwell's formulae, which explains the names of E_{ab} and H_{ab} . In fact, the Maxwell-like form of the free gravitational field underlines the rich correspondence between electromagnetism and general relativity and it has been the subject of theoretical debate for decades (see [67–70] for a representative list).

1.4. The Friedmann–Lemaître universe

In the previous sections we have considered general inhomogeneous and anisotropic cosmological spacetimes with imperfect total energy–momentum tensor. However, the current observational evidence (principally the CMB) and our theoretical prejudice (the Copernican principle), strongly support a universe that is homogeneous and isotropic on cosmological scales, namely a Friedmann–Lemaître–Robertson–Walker (FLRW) universe.

1.4.1. The FLRW metric

The geometry of the simplest non-static, non vacuum solution of the Einstein field equations is described by the Robertson–Walker line element. In suitable (comoving) coordinates the latter takes the form

$$ds^2 = -dt^2 + a^2(t) \left[dr^2 + f_K^2(r)(d\theta^2 + \sin^2\theta d\phi^2) \right], \quad (1.4.1)$$

where a is the scale factor and f_K depends on the geometry of the 3-D hypersurfaces. The scale factor defines a characteristic length scale and leads to the familiar Hubble parameter $H = \dot{a}/a$, which determines the rate of the (isotropic) expansion. The isotropy of the 3-space means that the latter has curvature equal to $\mathcal{R} = 6K/a^2$, with the curvature index K normalised to ± 1 when it is not zero. Then,

$$f_K(r) = \begin{cases} \sin r & \text{for } K = +1, \\ r & \text{for } K = 0, \\ \sinh r & \text{for } K = -1. \end{cases} \quad (1.4.2)$$

When $K = +1$ the 3-space is closed with spherical geometry and finite total volume. Alternatively, we have flat Euclidean 3-D hypersurfaces for $K = 0$ and open, hyperbolic, ones when $K = -1$. In either of these two cases the 3-space is unbounded unless nontrivial topologies are employed.

1.4.2. FLRW cosmologies

The high symmetry of the Friedmann models means that all kinematical and dynamical variables are functions of time only and any quantity that represents anisotropy or inhomogeneity vanishes identically. Thus, in covariant terms an FLRW model has $\Theta = 3H(t) \neq 0$, $\sigma_{ab} = 0 = \omega_a = A_a$, $E_{ab} = 0 = H_{ab}$, where $H = \dot{a}/a$ is the familiar Hubble parameter. The isotropy of the Friedmann models also constrains their matter content, which can only have the perfect-fluid form (with $\rho = \rho(t)$ and $p = p(t)$). In addition, due to the spatial homogeneity, all orthogonally projected gradients (e.g. $D_a\rho$, D_ap , etc) are by definition zero. These mean that the only nontrivial equations are the FLRW version of Raychaudhuri's formula, the equation of continuity and the Friedmann equation. These follow from expressions (1.3.3), (1.3.17) and (1.3.40) and are given by

$$\dot{H} = -H^2 - \frac{1}{6}(\rho + 3p) + \frac{1}{3}\Lambda, \quad \dot{\rho} = -3H(\rho + p) \quad (1.4.3)$$

and

$$H^2 = \frac{1}{3}\rho - \frac{K}{a^2} + \frac{1}{3}\Lambda, \quad (1.4.4)$$

respectively. Note that the isotropy of the FLRW models and Eq. (1.3.35) imply that the associated 3-Riemann tensor is given by $\mathcal{R}_{abcd} = (K/a^2)(h_{ac}h_{bd} - h_{ad}h_{bc})$. Introducing the density parameters $\Omega_\rho = \rho/3H^2$, $\Omega_\Lambda = \Lambda/3H^2$ and $\Omega_K = -K/(aH)^2$, the Friedmann equation takes the form

$$1 = \Omega_\rho + \Omega_K + \Omega_\Lambda. \quad (1.4.5)$$

Thus, in the absence of a cosmological constant, the 3-space is flat (i.e. $K = 0$) when the matter density takes the critical value $\rho = \rho_c = 3H^2$ and the Friedmann equation reduces to $K/a^2 = H^2(\Omega_\rho - 1)$. In that case $\Omega_\rho = 1$ ensures Euclidean 3-D hypersurfaces, while $\Omega_\rho > 1$ leads to spherical and $\Omega_\rho < 1$ to hyperbolic spatial geometry. One may also combine Eqs. (1.4.3)(a) and (1.4.4) to obtain an alternative form for the Raychadhuri equation with an explicit 3-curvature dependence, namely

$$\dot{H} = -\frac{1}{2}(\rho + p) + \frac{K}{a^2}. \quad (1.4.6)$$

After a little algebra, expression (1.4.3)(a) takes the alternative form

$$qH^2 = \frac{1}{6}(\rho + 3p) - \frac{1}{3}\Lambda, \quad (1.4.7)$$

where $q = -\ddot{a}/\dot{a}^2 = -[1 + (\dot{H}/H^2)]$ is the dimensionless deceleration parameter. When the latter is negative the universe accelerates, which means that in exact FLRW models with vanishing Λ we need to violate the strong energy condition (i.e. set $\rho + 3p < 0$) to achieve accelerated expansion.

The expansion rate also defines a representative length scale, the Hubble radius,

$$\lambda_H = H^{-1}. \quad (1.4.8)$$

In most FLRW models the scale-factor evolution (see Section 1.4.4) ensures that the Hubble length effectively coincides with the particle horizon ($d_H \propto t$). In that case, the Hubble radius determines the regions of causal contact.

The scale factor of an FLRW spacetime with non-Euclidean spatial geometry also defines the curvature scale ($\lambda_K = a$) of the model. This is the threshold at which any departures from Euclidean flatness in the geometry of the spatial hypersurfaces start becoming important (e.g. see [71]). Scales smaller than the curvature length are termed subcurvature, while those exceeding λ_K are referred to as supercurvature. The former are essentially immune to the effects of spatial geometry, which become prominent only on supercurvature lengths. However, the dynamics of fluctuations is sensitive to the spatial geometry on all scales, through its effect on the expansion rate (via the Friedmann equations). The relation between the curvature scale and the Hubble radius is determined by Eq. (1.4.4). In the absence of a cosmological constant, the latter takes the form

$$\left(\frac{\lambda_K}{\lambda_H}\right)^2 = -\frac{K}{1 - \Omega_\rho}. \quad (1.4.9)$$

Therefore, for $K = -1$ we find that $\lambda_K > \lambda_H$ always, with $\lambda_K \rightarrow \infty$ as $\Omega_\rho \rightarrow 1$ and $\lambda_K \rightarrow \lambda_H$ for $\Omega_\rho \rightarrow 0$. In practice, this means that supercurvature scales in spatially open FLRW cosmologies are never causally connected. When dealing with closed models, on the other hand, expression (1.4.9) shows that $\lambda_K > \lambda_H$ when $\Omega_\rho < 2$ and $\lambda_K \leq \lambda_H$ if $\Omega_\rho \geq 2$. We finally note that the importance of spatial geometry within a comoving region is always the same, since the curvature scale simply redshifts with the expansion.

1.4.3. Luminosity distance

The luminosity distance of an object at redshift z is $D_L = a_0(1+z)r_0$, where a_0 and r_0 are the current values of the scale factor and of the object's radial distance (e.g. see [72]). The latter is determined by integrating the line element $dt = (a/\sqrt{1-Kr^2})dr$ of a null geodesic. Assuming a spatially flat FLRW model, the result reads

$$r_0 = a_0^{-1} \int_0^z H^{-1} dx, \quad (1.4.10)$$

and it is easily integrated through the various epochs of the expansion. In view of the recent supernovae observations, however, it helps to express the above in terms of kinematical quantities, and particularly in terms of the deceleration parameter. Following Section 1.4.2 and recalling that $dz = -(1+z)Hdt$, we have [73]

$$\int_H^{H_0} H^{-1} dH = \ln\left(\frac{H_0}{H}\right) = -\int_0^z (1+q)d[\ln(1+x)], \quad (1.4.11)$$

which substituted into Eq. (1.4.10) leads to

$$a_0 r_0 = H_0^{-1} \int_0^z e^{-\int_0^x (1+q)d[\ln(1+y)]} dx. \quad (1.4.12)$$

Consequently, expressed in terms of the kinematical parameters of a spatially flat Friedmann model, the luminosity distance of an object at redshift z is given by⁶

$$D_L = (1+z)H_0^{-1} \int_0^z e^{-\int_0^x (1+q)d[\ln(1+y)]} dx. \quad (1.4.13)$$

When compared with the measured luminosity distance from remote type Ia supernovae, the above expression indicated that our universe has recently entered a phase of accelerating expansion [73,75].

1.4.4. Scale-factor evolution in FLRW cosmologies

The $K = 0$ case: To close the (1.4.3) system one needs to introduce an equation of state for the matter component. Before doing so, we will first briefly refer to the de Sitter universe. This is an exponentially expanding $K = 0$ model, containing no matter and having a positive cosmological constant. Applied to the de Sitter space, expression (1.4.4) reduces to $H = \sqrt{\Lambda/3} = \text{constant}$. The latter integrates immediately, giving $a \propto e^{\sqrt{\Lambda/3}t}$ and thus guaranteeing the exponential nature of the expansion.

⁶ For further discussion, extending to FLRW models with $K \neq 0$, and for expressions of the luminosity distance in terms of higher order derivatives of the scale factor, the reader is referred to [74].

In what follows we will consider barotropic perfect fluids, mainly in the form of non-relativistic dust or isotropic radiation (with $p = 0$ and $p = \rho/3$ respectively). When $w = p/\rho$ is the constant barotropic index of the cosmic medium, the continuity equation (see (1.4.3)(b)) gives

$$\rho = \rho_0 \left(\frac{a_0}{a} \right)^{3(1+w)}. \quad (1.4.14)$$

Substituting this result into the Friedmann equation, assuming Euclidean spatial sections (i.e. $K = 0$), $w \neq -1$ and setting the cosmological constant to zero, we arrive at the following expression for the scale factor

$$a = a_0 \left(\frac{t}{t_0} \right)^{2/3(1+w)}, \quad (1.4.15)$$

having normalised our solution so that $a(t = 0) = 0$. When dealing with non-relativistic matter with $w = 0$ (e.g. baryonic ‘dust’ or non-baryonic cold dark matter), we have what is known as the Einstein–de Sitter universe with $a \propto t^{2/3}$. Alternatively, we obtain $a \propto t^{1/2}$ in the case of relativistic species (e.g. isotropic radiation) and $a \propto t^{1/3}$ for a stiff medium with $w = 1$. An additional special case is that of matter with zero gravitational mass, which corresponds to $w = -1/3$ and leads to ‘coasting’ expansion with $a \propto t$. Solution (1.4.15) does not apply to a medium with $w = -1$ (and therefore with zero inertial mass – see Eq. (1.3.19)(b)). In that case, (1.4.14) guarantees that $\rho = \rho_0 = \text{constant}$, which when substituted into (1.4.4) leads to $H = H_0 = \text{constant}$ and subsequently to exponential expansion (inflation) with $a \propto e^{H_0(t-t_0)}$. Note that during a phase of exponential (de Sitter-type) expansion, the Hubble radius remains constant, while the particle horizon increases in the usual manner (see Section 1.4.2).

The $K = +1$ case: The equation of continuity does not depend on the curvature of the 3-space, which means that expression (1.4.14) monitors the evolution of the matter density irrespective of the model’s spatial curvature. When the FLRW spacetime has non-Euclidean spatial geometry it helps to parametrise the scale-factor evolution in terms of the conformal time (η – defined by $\dot{\eta} = 1/a$). Then, for $K = +1$, $\Lambda = 0$ and $w \neq -1/3$ relations (1.4.3) and (1.4.4) combine to give

$$a = a_0 \left\{ \frac{\sin[(1+3w)\eta/2 + \mathcal{C}]}{\sin[(1+3w)\eta_0/2 + \mathcal{C}]} \right\}^{2/(1+3w)}, \quad (1.4.16)$$

where $(1+3w)\eta/2 + \mathcal{C} \in (0, \pi)$. Normalising so that $a(\eta \rightarrow 0) \rightarrow 0$, the $\eta = \pi/(1+3w)$ threshold corresponds to the moment of maximum expansion when $a = a_{\text{max}} = a_0 \{\sin[(1+3w)\eta_0/2]\}^{-2/(1+3w)}$. For non-relativistic matter $w = 0$ and the above solution reduces to $a \propto \sin^2(\eta/2)$ [72], while we obtain $a \propto \sin \eta$ if radiation dominates. When $w = -1/3$ one can no longer use solution (1.4.16). Instead, Eq. (1.4.3)(a) leads immediately to the familiar coasting-expansion phase with $a \propto t$. Expressions (1.4.3)(b), (1.4.4) also provide the relation between scale factor and proper time in the $w = -1$ case. Just like in spatially flat models, Eq. (1.4.3)(b) ensures that $\rho = \rho_0 = \text{constant}$ and then (1.4.4) leads to $a(1 + \sqrt{3/\rho_0}H) \propto e^{\sqrt{(\rho_0/3)}t}$.

The Einstein universe corresponds to a static $K = +1$ model with positive cosmological constant. In such an environment the density of the matter component is also constant, while Eqs. (1.4.3)(a), (1.4.4) and (1.4.6) reduce to the constraints

$$\rho + 3p = 2\Lambda, \quad \frac{1}{a^2} = \frac{1}{3}(\rho + \Lambda) \quad \text{and} \quad \frac{1}{a^2} = \frac{1}{2}(\rho + p), \quad (1.4.17)$$

respectively. The Einstein universe has long been known to be unstable under homogeneous perturbations, though its stability to inhomogeneous distortions is less straightforward (see Section 3.7.2).

The $K = -1$ case: Applied to FLRW cosmologies with hyperbolic spatial geometry, zero cosmological constant and $w \neq -1/3$, the analysis described above leads to the following (ever expanding) evolution law for the scale factor

$$a = a_0 \left\{ \frac{\sinh[(1+3w)\eta/2 + \mathcal{C}]}{\sinh[(1+3w)\eta_0/2 + \mathcal{C}]} \right\}^{2/(1+3w)}, \quad (1.4.18)$$

where now $(1+3w)\eta/2 + \mathcal{C} > 0$. Not surprisingly, the above can be also obtained from (1.4.16), after the trigonometric functions are replaced with their hyperbolic counterparts. Assuming pressure-free ‘dust’ and normalising as before, we find $a \propto \sinh^2(\eta/2)$ [72]. On the other hand, solution (1.4.18) implies $a \propto \sinh \eta$ for an open FLRW universe dominated by relativistic species. We finally note that, similarly to the $K = +1$ case, the system (1.4.3) and (1.4.4) ensures that $a \propto t$ when $w = -1/3$ and $a(1 + \sqrt{3/\rho_0}H) \propto e^{\sqrt{(\rho_0/3)}t}$ for $w = -1$.

A special model with open spatial geometry is the vacuum Milne universe. Similarly to the $w = -1/3$ case, the absence of matter means that (1.4.3)(a) integrates to give a coasting scale factor of the general form $a \propto t$. Here, however, the Friedmann equation – see expression (1.4.4) – guarantees that $a = t$.

We finally point out that, in the absence of a cosmological constant and after introducing the transformation $y = a^{(1+3w)/2}$, with $w \neq -1/3$, Raychaudhuri’s equation (see expression (1.4.3)(a) in Section 1.4.2) reduces to a simple harmonic-oscillator of the form

$$y'' = -K \left(\frac{1+3w}{2} \right)^2 y, \quad (1.4.19)$$

Table 1

The non-tilted Bianchi spacetimes classified into two group classes and ten group types (see [82])

Group class	Group type	n_1	n_2	n_3	FLRW as special case
$A(a = 0)$	I	0	0	0	$K = 0$
	II	+	0	0	–
	VI_0	0	+	–	–
	VII_0	0	+	+	$K = 0$
	VIII	–	+	+	–
	IX	+	+	+	$K = +1$
$B(a \neq 0)$	V	0	0	0	$K = -1$
	IV	0	0	+	–
	VI_h	0	+	–	–
	VII_h	0	+	+	$K = -1$

with primes indicating conformal time derivatives [76]. This expression is particularly useful when addressing spatially closed or open FLRW models, with $\Lambda = 0$ and $w \neq -1/3$. For instance, it is straightforward to verify that for $K = \pm 1$ the above given equation leads immediately to solutions (1.4.16) and (1.4.18) respectively.

1.5. The Bianchi universes

Despite the success of the Friedmann–Lemaître models, the structure that we observe today, means that our universe is neither homogeneous nor isotropic, at least on certain scales and to a certain extent. To follow the (late time) evolution of the universe on these scales one needs models with more degrees of freedom than the FLRW ones. The spatially homogeneous and anisotropic Bianchi models have long been used to understand the observed level of isotropy in our universe and also to probe the nature of the initial singularity. Here, we will briefly consider members of the Bianchi family that contain the FLRW models as special cases, referring the reader to review articles and monographs for further details [77–85].

1.5.1. Classification of Bianchi cosmologies

Time is essentially the only dynamical coordinate in the Bianchi spacetimes, the spatial homogeneity of which has ‘removed’ all the inhomogeneous degrees of freedom and has reduced Einstein’s equations to a set of ordinary differential equations. Despite this, the Bianchi family provides a rich set of models where one can study the fully nonlinear theory. We generally distinguish between two different kinds of Bianchi cosmologies. The non-tilted (or orthogonal) models, with the flow-lines of the fluid normal to the hypersurfaces of homogeneity and the tilted models where this is no longer true. In the latter case the ‘peculiar’ velocity of the matter enters the equations as an additional dynamical variable (see Section 5.1).

The literature contains three basic ways of classifying the orthogonal Bianchi models, all based on the commutation laws of the associated tetrad basis vectors.⁷ Thus, one may use the tetrad $\{\mathbf{e}_a, a = 0, 1, 2, 3\}$, so that \mathbf{e}_0 is the normal vector to the hypersurfaces of homogeneity and

$$[\mathbf{e}_a, \mathbf{e}_b] = \gamma^c_{ab} \mathbf{e}_c, \tag{1.5.1}$$

with the commutation functions $\gamma^a_{bc} = \gamma^a_{bc}(t)$ treated as dynamical variables themselves [77]. The spatial commutators $\gamma^\alpha_{\beta\gamma}$, with $\alpha, \beta, \gamma = 1, 2, 3$, are then decomposed into the time-dependent pair $n_{\alpha\beta}$ and a_α that satisfy the condition

$$n_{\alpha\beta} a^\beta = 0. \tag{1.5.2}$$

Choosing the tetrad so that $n_{\alpha\beta}$ is a diagonalisable matrix (i.e. $n_{\alpha\beta} = \text{diag}(n_1, n_2, n_3)$) and $a_\alpha = (a, 0, 0)$, the above reduces to $n_1 a = 0$. Consequently, one can immediately define two major classes of Bianchi spacetimes. Those with $a = 0$ are known as class A models and those with $a \neq 0$ are termed class B. Further classification is achieved by the signs of the eigenvalues of $n_{\alpha\beta}$. When dealing with the class B models one may also introduce the scalar h , which satisfies the constraint

$$a^2 = h n_2 n_3. \tag{1.5.3}$$

This means that the h -parameter is well defined only in class B models with $n_2 n_3 \neq 0$. The general Bianchi classification is given in Table 1, showing that $h < 0$ in type VI_h and $h > 0$ in VII_h . We finally note that some Bianchi groups allow for subspaces of higher symmetry, like isotropic or locally rotationally symmetric models. As a result, the FLRW universes may sometimes appear as special cases in certain Bianchi cosmologies.

⁷ This means classifying the Lie algebras of the Killing vector fields and therefore the associated group of the G_3 -isometries (see [77,82] for details).

1.5.2. Bianchi I cosmologies

The simplest anisotropically expanding cosmologies, which are also the simplest generalisation of the spatially flat FLRW universe, are the non-tilted Bianchi I models. These are class A spacetimes (see Table 1), with Euclidean 3-geometry and line elements of the form

$$ds^2 = -dt^2 + X^2(t)dx^2 + Y^2(t)dy^2 + Z^2(t)dz^2, \quad (1.5.4)$$

in comoving coordinates.⁸ The above allows for different expansion rates along the three spatial directions, with the average scale factor and the mean Hubble parameter given by $a = \sqrt[3]{XYZ}$ and $H = \dot{\theta}/3 = \dot{a}/a$ respectively. The spatial homogeneity of the Bianchi I spacetimes ensures that all invariants depend on time only. The flow lines of the fundamental observers are irrotational geodesics and all spatial gradients vanish identically. Therefore, in covariant terms, all Bianchi I cosmologies are characterised by

$$\omega_a = 0 = A_a = H_{ab} = \mathcal{R}_{ab}, \quad (1.5.5)$$

which means that the only nonzero quantities are the volume scalar, the shear tensor and the electric part of the Weyl field. Also, because of their generic anisotropy, the Bianchi I spacetimes can support imperfect fluids with non-vanishing anisotropic pressure (i.e. $\pi_{ab} \neq 0$ though $q_a = 0$). For example, the type-I models are natural hosts of large-scale magnetic fields [86].

As with the FLRW case, the 1 + 3 covariant formulae monitoring the evolution of the Bianchi I cosmologies are obtained from the general expressions given in Section 1.3. Thus, applied to a type-I environment and in the absence of a cosmological constant, Eqs. (1.3.3), (1.3.17) and (1.3.40) reduce to

$$\dot{H} = -H^2 - \frac{1}{6}(\rho + 3p) - \frac{2}{3}\sigma^2, \quad \dot{\rho} = -3H(\rho + p) - \sigma_{ab}\pi^{ab} \quad (1.5.6)$$

and

$$H^2 = \frac{1}{3}(\rho + \sigma^2), \quad (1.5.7)$$

where the latter can be seen as the Bianchi I analogue of the Friedmann equation. Also note that, on using the density parameter Ω_ρ defined in Section 1.4.2, relation (1.5.7) takes the form

$$1 = \Omega_\rho + \Sigma, \quad (1.5.8)$$

with $\Sigma = \sigma^2/3H^2$ providing a measure of the model's shear anisotropy. Similarly, in a Bianchi I spacetime, expressions (1.3.4), (1.3.41) and (1.3.42) reduce to

$$\dot{\sigma}_{ab} = -2H\sigma_{ab} - \sigma_{c(a}\sigma^c_{b)} - E_{ab} + \frac{1}{2}\pi_{ab}, \quad E_{ab} = H\sigma_{ab} - \sigma_{c(a}\sigma^c_{b)} - \frac{1}{2}\pi_{ab} \quad (1.5.9)$$

and

$$\dot{E}_{ab} = -3HE_{ab} - \frac{1}{2}(\rho + p)\sigma_{ab} - \frac{1}{2}(\dot{\pi}_{ab} + H\pi_{ab}) + 3\sigma_{(a}^c(E_{b)c} - \frac{1}{6}\pi_{b)c}), \quad (1.5.10)$$

respectively. We also note that Eqs. (1.5.9)(b) recasts (1.5.9)(a) into

$$\dot{\sigma}_{ab} = -3H\sigma_{ab} + \pi_{ab}. \quad (1.5.11)$$

The latter ensures that, in the absence of anisotropic pressures, the shear depletes as a^{-3} , where a is the average (over the three spatial directions) scale factor.

Once an equation of state for the matter is introduced, the set (1.5.6)–(1.5.11) governs the dynamics of a Bianchi I spacetime fully. In the special case of matter in the perfect-fluid form with $p = w\rho$, expression (1.5.6)(b) integrates to $\rho \propto a^{-3(1+w)}$. Following (1.5.7), this means that the shear will dominate the early expansion, no matter how small the anisotropy may be today (unless the matter component has $w = 1$ —stiff fluid). These shear dominated early stages correspond to the vacuum Kasner regime, in which case equation (1.5.7) leads to $a \propto t^{1/3}$. The line element of the Kasner solution has the form

$$ds^2 = -dt^2 + t^{2p_1}dx^2 + t^{2p_2}dy^2 + t^{2p_3}dz^2, \quad (1.5.12)$$

where $p_1 + p_2 + p_3 = 1 = p_1^2 + p_2^2 + p_3^2$ (e.g. see [87–89]). Together, these conditions guarantee that either exactly one of the three exponents is negative or two of them are zero. In the former case the spacetime expands in two directions and contracts along the third, with a cigar-like initial singularity. In the latter case we have motion (expansion) in one direction

⁸ For a discussion on the classification of the Bianchi models in two major classes (A and B) on the basis of their structure constants the reader is referred to [77,82].

only, which corresponds to a pancake-type singularity.⁹ Note that the total volume always increases with time, since the average scale factor grows as (recall that $a \propto t^{1/3}$). Given that kinematics dictate the Kasner phase, it is not surprising that the exponents of (1.5.12) are determined by the model's shear anisotropy according to [82]

$$p_1 = \frac{1}{3}(1 - 2\Sigma_+) \quad \text{and} \quad p_{2,3} = \frac{1}{3}(1 + \Sigma_+ \pm \sqrt{3}\Sigma_-), \quad (1.5.13)$$

with $\Sigma_{\pm} = \sigma_{\pm}/H$. Note that $\sigma_+ = (\sigma_2 + \sigma_3)/2$ and $\sigma_- = (\sigma_2 - \sigma_3)/2\sqrt{3}$, where σ_2 and σ_3 are the two independent components of the shear tensor (recall that $\sigma_{\alpha\beta} = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ and $\sigma_1 + \sigma_2 + \sigma_3 = 0$ —see also [82] for more technical details and further references).

1.5.3. Bianchi VII_h cosmologies

The most general Bianchi spacetimes that contain the open FLRW universe as a special case are the type-VII_h cosmologies (see Table 1). The late-time attractors of this class B family is, for a broad range of initial data and matter properties, the Lukash plane-wave solution. These vacuum spacetimes are self-similar equilibrium points and have a line element of the form¹⁰

$$ds^2 = -dt^2 + t^2 dx^2 + t^{2r} e^{2rx} [(A dy + B dz)^2 + (C dy + A dz)^2]. \quad (1.5.14)$$

Note that r is a constant parameter in the range $0 < r < 1$, $A = \cos \nu$, $B = f^{-1} \sin \nu$, $C = -f \sin \nu$ and $\nu = k(x + \ln t)$. Also, f and k are constants related to r by

$$\frac{k^2(1-f^2)^2}{f^2} = 4r(1-r) \quad \text{and} \quad r^2 = hk^2, \quad (1.5.15)$$

where h is the associated group parameter. Note that r determines the amount and the nature of the model's anisotropy. When $r \rightarrow 1$, in particular, the anisotropy vanishes and (1.5.14) reduces to the metric of the empty Milne universe. At the $r \rightarrow 0$ limit, on the other hand, the anisotropy is maximised [90,91].

Due to the absence of matter and given its irrotational nature, the Lukash spacetime is covariantly characterised by the irreducible sets

$$\rho = 0 = p = q_a = \pi_{ab} \quad \text{and} \quad A_a = 0 = \omega_a, \quad (1.5.16)$$

which imply that the only nonzero quantities are Θ , σ_{ab} , E_{ab} and H_{ab} . Note that the Weyl components have equal magnitudes and are orthogonal to each other (i.e. $E^2 = H^2$ and $E_{ab}H^{ab} = 0$, respectively), in line with the Petrov-type N nature of the solution. The absence of matter means that the Lukash universe is Ricci flat, although the curvature of the 3-space (i.e. the 3-Ricci tensor) is nonzero. Setting the cosmological constant to zero, the kinematics of the model is governed by the propagation formulae [91]

$$\dot{H} = -H^2 - \frac{2}{3}\sigma^2, \quad \dot{\sigma}_{ab} = -3H\sigma_{ab} - \mathcal{R}_{(ab)}, \quad (1.5.17)$$

which are supplemented by the constraints

$$H^2 = \frac{1}{3}\left(\sigma^2 - \frac{1}{2}\mathcal{R}\right) \quad \text{and} \quad H_{ab} = \text{curl } \sigma_{ab}. \quad (1.5.18)$$

The evolution of the Weyl field, on the other hand, is monitored by the set

$$\dot{E}_{ab} = -3HE_{ab} + \text{curl } H_{ab} + 3\sigma_{c(a}E^c{}_{b)}, \quad \dot{H}_{ab} = -3HH_{ab} - \text{curl } H_{ab} + 3\sigma_{c(a}H^c{}_{b)}. \quad (1.5.19)$$

The average kinematic anisotropy is measured by means of the shear parameter defined by the dimensionless, expansion normalised parameter $\Sigma = \sigma^2/3H^2 = (1-r)/(1+2r)$ [91]. This means that Σ remains constant during the model's evolution and lies between zero and unity. Also, minimum anisotropy corresponds to $\Sigma \rightarrow 0$ and maximum to $\Sigma \rightarrow 1$, ensuring that the shear of the Lukash model is always bounded. On using the Σ -parameter, Eqs. (1.5.17)(a) and (1.5.18)(a) are recast into

$$\dot{H} = -3H^2(1+2\Sigma) \quad \text{and} \quad \mathcal{R} = 6H^2(\Sigma - 1), \quad (1.5.20)$$

respectively. The former leads to the power-law evolution $a \propto t^{1/(1+2\Sigma)}$ of the average scale factor and the latter ensures that the spatial sections are open. Thus, when the shear is at its minimum, we approach $a \propto t$ and the Milne universe. At the $\Sigma \rightarrow 1$ limit, on the other hand, we find the familiar Kasner solution (i.e. $a \propto t^{1/3}$ —see Section 1.5.2). Note that for maximum shear anisotropy the 3-curvature vanishes, whereas the Weyl field tends to zero at both limits [90,91]. Finally, the deceleration parameter of the Lukash universe is $q = 2\Sigma$, with $0 < q < 2$.

⁹ With the exception of the $(p_1, p_2, p_3) = (0, 0, 1)$ triplet, which corresponds to a flat spacetime, the initial singularity cannot be eliminated by any coordinate transformation [88].

¹⁰ Self-similar Bianchi solutions, vacuum or with a non-tilted perfect fluid, have been studied in [90].

1.5.4. Bianchi IX cosmologies

The type-IX model is a class A spacetime and the only Bianchi cosmology that contains the closed FLRW universe as a special case (see Table 1). The model is well known for its oscillatory behaviour with chaotic characteristics and with the matter becoming dynamically negligible as it approaches the initial singularity [92]. The spacetime metric has the form

$$ds^2 = -dt^2 + X^2(t)(\cos \psi d\theta + \sin \psi \sin \theta d\phi)^2 + Y^2(t)(\sin \psi d\theta - \cos \psi \sin \theta d\phi)^2 + Z^2(t)(d\psi + \cos \theta d\phi)^2, \quad (1.5.21)$$

parametrised by the Euler angles with $0 \leq \psi \leq 4\pi$, $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. Allowing H to assume all real values and therefore permitting a contracting epoch, the generalised Friedmann equation of the model reads

$$H^2 = \frac{1}{3} \left(\rho + \sigma^2 - \frac{1}{2} \mathcal{R} \right). \quad (1.5.22)$$

A key feature of the type-IX cosmologies is that their spatial curvature changes sign during the model's evolution. In particular, the associated 3-Ricci scalar is given by [82]

$$\mathcal{R} = -\frac{1}{2} \left[n_1^2 + n_2^2 + n_3^2 - 2(n_1 n_2 + n_2 n_3 + n_3 n_1) \right], \quad (1.5.23)$$

where $n_{1,2,3} > 0$ according to Table 1. Alternatively, one may express \mathcal{R} in terms of the individual scale factors as [93]

$$\mathcal{R} = \frac{2(X^2 Y^2 + X^2 Z^2 + Y^2 Z^2) - (X^4 + Y^4 + Z^4)}{2(XYZ)^2}. \quad (1.5.24)$$

The right-hand side in both of the above expressions can take either sign. In fact, \mathcal{R} is predominantly negative and becomes positive only when the model approaches isotropy (i.e. for $n_1 = n_2 = n_3$ or $X = Y = Z$). Since an expansion maximum occurs when the curvature is positive (see Eq. (1.5.22)), one may argue that the model cannot recollapse while still anisotropic [94].

The past attractor of the Bianchi-IX spacetimes is the so-called *Mixmaster* oscillatory singularity. A qualitative analysis shows that the model approaches the initial singularity through a sequence oscillatory eras. Each era consists of alternating Kasner phases, with a metric given by (1.5.12) and the negative exponent shifting between two of the directions [87,88]. This means that distances along the associated two axes oscillate, while those in the third decrease. As the model passes through the different eras, the 'decreasing axis' bounces from one direction to the next, with the process asymptotically acquiring a random character [95]. The chaotic behaviour of the type-IX model was suggested as a way of achieving sufficient 'mixing' between the three spatial directions that could remove the horizon problem [92]. As it eventually turned out, however, this mechanism does not work.

1.5.5. Isotropisation of Bianchi cosmologies

Bianchi cosmologies have been traditionally studied qualitatively, primarily by means of dynamical system methods [82, 83]. These techniques have revealed an interesting property of many Bianchi models, namely their "intermediate isotropisation". The latter occurs because a number of Bianchi-type spacetimes have phase planes where the FLRW solutions are acting as saddle points. This means that these models can isotropise and therefore look very much like a Friedmann universe, over an extended period of their evolution, despite the fact that they start off and end up quite unlike the FLRW spacetimes.

Another issue is whether the Bianchi models show any tendency to isotropise, either at early or at late times. Following [96], the set of Bianchi models, with conventional matter, that isotropise asymptotically to the future is of zero measure. It has been shown, however, that Bianchi cosmologies tend towards isotropy at late times when a cosmological constant is present. Taken at face value, this implies that inflation should smooth the anisotropy of these models out. Nevertheless, the existing results apply primarily to non-tilted Bianchi types and also seem to depend on the amount of the initial anisotropy.

1.5.6. Kantowski–Sachs cosmologies

These are spatially homogeneous spacetimes that do not belong to the Bianchi family [97,98]. The Kantowski–Sachs class of models have local rotational symmetry, with metrics which in comoving coordinates read [44,84]

$$ds^2 = -dt^2 + A^2(t)dr^2 + B^2(t) \left[d\theta^2 + f^2(\theta) d\phi^2 \right], \quad (1.5.25)$$

where $f(\theta) = \sin \theta$. Note that in general the function $f(\theta)$ obeys an expression of the form (1.4.2) – see Section 1.4.1 – in which case one refers to Kantowski–Sachs-like metrics [99]. However, only the $K = +1$ model falls outside the Bianchi family. Those with zero and negative spatial curvature reduce to axisymmetric type-I and type-III cosmologies respectively [81]. Note that, in the Bianchi classification of Section 1.5.1, the missing type-III spacetime corresponds to the VI₋₁ model [82].

2. Inhomogeneous relativistic cosmologies

The simplest inhomogeneous cosmologies are spherically symmetric, like the Lemaitre–Tolman–Bondi (LTB) model. The latter has closed spatial sections and matter in the form of irrotational dust. The LTB universe possesses a centre of symmetry, in fact it can allow for up to two such centres, but it is not isotropic about an arbitrary observer; an inevitable consequence of the model's spatial inhomogeneity [81]. There also exist inhomogeneous solutions of the EFE without symmetries, with the Szekeres quasi-spherical model probably being the most celebrated (see [81] for an extended discussion and references). An additional class of cosmologies without (global) symmetry are the so called Swiss–Cheese models, obtained by cutting and pasting segments of spherically symmetric spacetimes (see [44]). The universe we live in is also believed to be free of symmetries.

2.1. The gauge problem in cosmology

It has long been known that the study of cosmological perturbations is plagued by what is known as the gauge problem, reflecting the fact that in perturbation theory we deal with two spacetime manifolds [100–104]. The first is the physical spacetime, \mathcal{W} , that corresponds to the real universe and the second, denoted here by $\overline{\mathcal{W}}$, is a fictitious idealised mathematical model.

2.1.1. Gauge freedom

In most cosmological studies the idealised, background, spacetime is represented by the homogeneous and isotropic FLRW models. To proceed, one needs to establish a one-to-one correspondence, namely a gauge $\phi : \overline{\mathcal{W}} \rightarrow \mathcal{W}$, between the two spacetimes. Such a point-identification map is generally arbitrary, although particular ones may be more suitable for specific cases. When a coordinate system is introduced in $\overline{\mathcal{W}}$, the gauge carries it to \mathcal{W} and vice versa. As a result, a smooth spacetime is defined into the real universe. Any change in $\phi : \overline{\mathcal{W}} \rightarrow \mathcal{W}$, keeping the background coordinates fixed, is known as a gauge transformation. This introduces a coordinate transformation in the physical spacetime but also changes the event in \mathcal{W} which is associated with a given event of $\overline{\mathcal{W}}$. Gauge transformations are therefore different from coordinate transformations which merely relabel events. The gauge problem stems from our inherent freedom to make gauge transformations. Although, the gauge freedom is usually expressed as a freedom of coordinate choice in \mathcal{W} , it should be understood that it generally changes the point-identification between the two spacetimes.

In the study of cosmological perturbations we consider the realistic universe and define perturbations by specifying the map $\phi : \overline{\mathcal{W}} \rightarrow \mathcal{W}$ between the \mathcal{W} and its fictitious counterpart. However, although we can always perturb away from a given background spacetime, recovering the smooth metric from a given perturbed one is not a uniquely defined process. This is a problem because it is always possible to choose an alternative background and therefore arrive at different perturbation values (see also [104]). Selecting an unperturbed spacetime from a given lumpy one corresponds to a gauge choice. Determining the best gauge is known as the fitting problem in cosmology and there is no unique answer to it [105].

2.1.2. Gauge dependence

By definition, the perturbation of any quantity is the difference between its value at some event in the real spacetime and its value at the corresponding, through the gauge, event in the background. Then, even scalar quantities that have nonzero and position-dependent background values, will lead to gauge-dependent perturbations. Following [106,107], we consider an one parameter family of 4-manifolds $\mathcal{W}_\epsilon = \mathcal{W}(\epsilon)$ embedded in a 5-manifold \mathcal{M} . Each one of these 4-manifolds represents a realistic spacetime, perturbed relative to the background manifold $\overline{\mathcal{W}}$. We define a point-identification map between $\overline{\mathcal{W}}$ and \mathcal{W}_ϵ , by introducing in \mathcal{M} a vector field X_A (with $A = 0, \dots, 4$), which is everywhere transverse to the embeddings \mathcal{W}_ϵ . Points lying along the same integral curves of X_A , which are parametrised by ϵ for convenience, will be regarded as the 'same'. Thus, selecting a specific vector field X_A corresponds to a choice of gauge. If Q_ϵ is some geometrical quantity defined on \mathcal{W}_ϵ , for small ϵ we have

$$\overline{h}_\epsilon(Q_\epsilon) = \overline{Q} + \epsilon \overline{\mathcal{L}}_X Q_\epsilon + \mathcal{O}(\epsilon^2), \quad (2.1.1)$$

where an overbar refers to quantities evaluated in $\overline{\mathcal{W}}$, \mathcal{L}_X is the Lie derivative along X_A and \overline{h}_ϵ is the pullback of \mathcal{W}_ϵ to $\overline{\mathcal{W}}$. The quantity $\delta Q = \overline{h}_\epsilon(Q_\epsilon) - \overline{Q} = \epsilon \overline{\mathcal{L}}_X Q_\epsilon$ is what we usually call linear perturbation of \overline{Q} and it clearly depends on our gauge choice [106,107].

According to (2.1.1), even quantities that behave like scalars under coordinate changes will not remain invariant under gauge transformations. To see this from a different, less technical, point of view we will follow [104]. Consider the familiar density perturbation $\delta\rho = \rho - \overline{\rho}$, where ρ is the matter density. In the right-hand side of the above we have the difference in the value of the matter density between two corresponding points in the background and the real spacetime. A gauge transformation will generally change this correspondence and therefore the perturbation value. This means that the value of $\delta\rho$ is entirely gauge-dependent and therefore arbitrary. For instance, one can select the gauge so that the surfaces of constant background density are the surfaces of constant real density, thus setting $\delta\rho$ to zero [104].¹¹

¹¹ With this gauge choice the fluid flow lines are not orthogonal to the surfaces of constant density and comoving observers will still measure a nonzero density variation.

2.2. Covariant and gauge-invariant perturbations

One way of addressing the gauge problem is by completely fixing the point-identification map between the background and the real spacetimes. However, determining the best gauge for a given physical problem is not a trivial task and it might lead to spurious, gauge-dependent, results. Alternatively, we may only partially fix the gauge (leaving some residual gauge freedom and always keeping track of its consequences) or employ gauge-invariant variables [103,104].

2.2.1. Criteria for gauge invariance

Gauge-independent quantities must remain invariant under gauge transformations between the idealised and the realistic spacetimes. According to the Stewart and Walker lemma, the simplest cases are scalars that are constant in the background universe or tensors that vanish there [106]. In both cases the mapped quantity is also constant and gauge changes are irrelevant because they all define the same perturbation. The only other possibility are tensors that can be written as linear combinations of products of the Kronecker deltas with constant coefficients. The same general criteria also apply to second order perturbations, but this time the Stewart and Walker requirements must be satisfied by the first-order variables [108].

Most cosmological applications deal with FLRW models. One would therefore like to know which quantities satisfy this criterion on Friedmannian backgrounds. Since the only invariantly defined constant is the cosmological constant and because constant products of the Kronecker deltas do not occur naturally, the only remaining option is to look for quantities that vanish in FLRW environments. Given the symmetries of the Friedmann models, any variable that describes spatial inhomogeneity or anisotropy must vanish there and therefore its linear perturbation should remain invariant under gauge transformations.

2.2.2. Gauge-invariant inhomogeneities

Covariantly, spatial inhomogeneities in the distribution of any physical quantity are described by the orthogonally projected gradient of the quantity in question. For the purposes of structure formation the key variable is the comoving fractional gradient in the energy density of the matter given by [104]

$$\Delta_a = \frac{a}{\rho} D_a \rho. \quad (2.2.1)$$

The above, which monitors density variations as measured by a pair of neighbouring fundamental observers (see [104] and also Section 3.1.1 here), is identically zero in spacetimes with homogeneous spatial sections. Indeed, by definition we have

$$D_a \rho = h_a^b \nabla_b \rho = h_a^0 \nabla_0 \rho + h_a^\alpha \nabla_\alpha \rho = 0, \quad (2.2.2)$$

since $h_a^0 = 0$ in a comoving frame and $\nabla_\alpha \rho = 0$ because $\rho = \rho(t)$. The vanishing of $D_a \rho$ in spatially-homogeneous models, like the Friedmann universes, guarantees that Δ_a satisfies the Stewart and Walker lemma [106]. Consequently, Δ_a describes density inhomogeneities within perturbed almost-FLRW spacetimes in a gauge independent way.

The density gradient can be supplemented by a number of auxiliary variables that describe spatial inhomogeneities in other physical quantities. Here, for the sake of economy, we will only introduce a variable for the volume-expansion gradients and for those in the magnetic energy density. Following [104,109], these are defined as

$$\mathcal{Z}_a = a D_a \Theta \quad \text{and} \quad \mathcal{B}_a = \frac{a}{B^2} D_a B^2, \quad (2.2.3)$$

respectively. Both vanish in spatially homogeneous spacetimes and therefore they also comply with the Stewart and Walker criterion for gauge invariance. For the rest of this section, we will analyse the behaviour of these quantities in different cosmological environments and then, in Section 3, we will use them to study the evolution of perturbed almost-FLRW universes.

2.3. Inhomogeneous single-fluid cosmologies

It is broadly accepted that the present large-scale structure of the universe is the result of Jeans-type instabilities, where small inhomogeneities in the initial density distribution of the cosmic medium grow gravitationally to form the galaxies and the voids seen in the universe today. Here, we will present the main equations governing the nonlinear evolution of density inhomogeneities, within single-fluid cosmologies, in a covariant and gauge-invariant manner.

2.3.1. Imperfect fluids

Consider a general spacetime filled with a single imperfect fluid. Spatial inhomogeneities in the matter density, as measured by a pair of neighbouring observers, are monitored by the orthogonally projected dimensionless comoving

gradient $\Delta_a = (a/\rho)D_a\rho$. Taking the covariant derivative of the above and using the energy and momentum conservation laws, respectively given by (1.3.17) and (1.3.18), we obtain

$$\begin{aligned} \dot{\Delta}_{(a)} = & \frac{p}{\rho} \Theta \Delta_a - \left(1 + \frac{p}{\rho}\right) \mathcal{Z}_a + \frac{a\Theta}{\rho} \left(\dot{q}_{(a)} + \frac{4}{3} \Theta q_a\right) - \frac{a}{\rho} D_a D^b q_b + \frac{a\Theta}{\rho} D^b \pi_{ab} - (\sigma^b{}_a + \omega^b{}_a) \Delta_b \\ & - \frac{a}{\rho} D_a (2A^b q_b + \sigma^{bc} \pi_{bc}) + \frac{a\Theta}{\rho} (\sigma_{ab} + \omega_{ab}) q^b + \frac{a\Theta}{\rho} \pi_{ab} A^b + \frac{1}{\rho} (D^b q_b + 2A^b q_b + \sigma^{bc} \pi_{bc}) (\Delta_a - aA_a). \end{aligned} \quad (2.3.1)$$

In the right-hand side of the above we notice a number of agents which act as sources of density perturbations. Hence, even if Δ_a is initially zero, it will not generally remain so. One of the key sources of density inhomogeneities is \mathcal{Z}_a , the volume expansion gradient, the nonlinear evolution of which is obtained by taking its time derivative and then using the associated expression of the Raychaudhuri equation. The result is

$$\begin{aligned} \dot{\mathcal{Z}}_{(a)} = & -\frac{2}{3} \Theta \mathcal{Z}_a - \frac{1}{2} \rho \dot{\Delta}_a - \frac{3}{2} a D_a p - a \left[\frac{1}{3} \Theta^2 + \frac{1}{2} (\rho + 3p) - \Lambda \right] A_a + a D_a D^b A_b \\ & - (\sigma^b{}_a + \omega^b{}_a) \mathcal{Z}_b - 2a D_a (\sigma^2 - \omega^2) + 2a A^b D_a A_b - a \left[2 (\sigma^2 - \omega^2) - D^b A_b - A^b A_b \right] A_a. \end{aligned} \quad (2.3.2)$$

2.3.2. Perfect fluids

When dealing with a perfect fluid, a choice of frame can be made in which the fluid appears isotropic, i.e. there is no energy–flux or anisotropic pressure and the associated variables vanish identically (i.e. $q_a = 0 = \pi_{ab}$). This choice considerably simplifies Eq. (2.3.1), meaning that perfect-fluid density inhomogeneities evolve as

$$\dot{\Delta}_{(a)} = \frac{p}{\rho} \Theta \Delta_a - \left(1 + \frac{p}{\rho}\right) \mathcal{Z}_a - (\sigma^b{}_a + \omega^b{}_a) \Delta_b. \quad (2.3.3)$$

There is no change in the propagation equation of the expansion gradients, which maintains the algebraic form of (2.3.2). The only difference the perfect fluid makes, is that Δ_a is now monitored by (2.3.3) and the 4-acceleration is given by (1.3.19)(b) instead of (1.3.18). If the medium is also barotropic (i.e. for $p = p(\rho)$), the pressure gradients are directly related to those in the density by $D_a p = c_s^2 D_a \rho$, where $c_s^2 = \dot{p}/\dot{\rho}$ is the adiabatic sound speed (see Section 1.3.4).

2.3.3. Covariant conserved quantities

In the metric-based perturbative formalism, the curvature perturbation on uniform density hypersurfaces is conserved for the adiabatic growing mode on super-Hubble scales, and the geometric interpretation of this can be understood via the perturbation of the expansion e-folds, $N = \ln a$, using the so-called “separate universe” picture [110]. A covariant version of this result is based on defining an appropriate spatial-gradient quantity, and leads to a simple geometric nonlinear conserved quantity for a perfect fluid [111].

Along each worldline of the perfect fluid, we define the generalised, covariant local e-fold function

$$\alpha = \frac{1}{3} \int \Theta d\tau, \quad (2.3.4)$$

where τ is proper time. Applying the commutation law (A.31) to the above scalar, as well as the density of the fluid, gives¹²

$$\frac{1}{3} D_a \Theta = \mathcal{L}_u D_a \alpha - \dot{\alpha} A_a, \quad (2.3.5)$$

where \mathcal{L}_u is the Lie derivative along u_a . Then, the projected gradient of the energy–density conservation law (see Eq. (1.3.19)(a)) leads to

$$\mathcal{L}_u (D_a \rho) + 3(\rho + p) \mathcal{L}_u (D_a \alpha) + \Theta D_a (\rho + p) = 0. \quad (2.3.6)$$

Defining the auxiliary projected vector $\zeta_a = D_a \alpha - (\dot{\alpha}/\rho) D_a \rho$, we can simplify expression (2.3.6) into

$$\mathcal{L}_u \zeta_a = -\frac{\Theta}{3(\rho + p)} \Gamma_a, \quad (2.3.7)$$

with $\Gamma_a = D_a p - (\dot{p}/\dot{\rho}) D_a \rho$. For isentropic perturbations, Γ_a vanishes identically (see expression (1.3.33) in Section 1.3.4) and the above guarantees that

$$\mathcal{L}_u \zeta_a = 0. \quad (2.3.8)$$

In other words, ζ_a is a conserved quantity in the isentropic/barotropic case on all scales and at all perturbative orders.

¹² Using the Lie derivative along u_a , simplifies the identity (A.31) to $D_a \dot{f} = \mathcal{L}_u (D_a f) - \dot{f} A_a$.

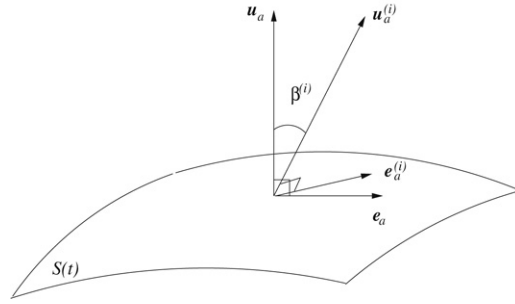


Fig. 3. In a multi-component system, the 4-velocity $u_a^{(i)}$ of the i -th fluid makes a hyperbolic angle $\beta^{(i)}$ with the fundamental 4-velocity field u_a , normal to the hypersurfaces of homogeneity $S(t)$. The unit vectors e_a and $e_a^{(i)}$ are orthogonal to u_a and $u_a^{(i)}$ respectively. Following definition (2.4.2), the peculiar velocity of the i -th species is $v_a^{(i)} = v^{(i)}e_a$, with $v_{(i)}^2 = v_a^{(i)}v_a^{(i)}$.

2.4. Inhomogeneous multi-fluid cosmologies

During its evolution the universe goes through epochs where the matter is better described by a mixture of several fluids, rather than a single component. This brings about the need for studies of inhomogeneous multi-component systems (see [112,113] for non-covariant treatments). When studying the effects of inhomogeneities on the CMB, for example, one usually considers a mixture of radiation, baryonic matter and neutrinos (e.g. see [114] and Section 4 here). Studies of nonlinear gravitational collapse also require the use of a multi-fluid description, in order to incorporate the effects of peculiar velocities (see Section 5.1 for more details).

2.4.1. 4-velocity fields

Consider spacetime filled with a mixture of different fluids. Suppose that u_a is the 4-velocity of the fundamental observers and $u_a^{(i)}$ that of the i -th fluid component (i.e. $u_a^{(i)}u_a^{(i)} = -1$). The tensors projecting orthogonal to u_a and $u_a^{(i)}$ are

$$h_{ab} = g_{ab} + u_a u_b \quad \text{and} \quad h_{ab}^{(i)} = g_{ab} + u_a^{(i)} u_b^{(i)}, \quad (2.4.1)$$

respectively. The relation between u_a and $u_a^{(i)}$ is determined by the Lorentz boost

$$u_a^{(i)} = \gamma^{(i)} (u_a + v_a^{(i)}), \quad (2.4.2)$$

where $v_a^{(i)}u_a = 0$. Here, $\gamma^{(i)} = (1 - v_{(i)}^2)^{-1/2}$ is the Lorentz-boost factor and $v_a^{(i)}$ is the peculiar velocity of the i -th component relative to u_a . For non-relativistic peculiar motions $\gamma^{(i)} \simeq 1$.

The boost relation can also be recast in terms of the hyperbolic, tilt, angle $\beta_{(i)}$ between the two 4-velocity vectors (see Fig. 3). Noting that $\cosh \beta^{(i)} = -u_a^{(i)}u_a = \gamma^{(i)}$ and $\sinh \beta^{(i)}e_a = \gamma^{(i)}v_a^{(i)} = h_a{}^b u_b^{(i)}$, where $v_a^{(i)} = v^{(i)}e_a$, expression (2.4.2) reads [115]

$$u_a^{(i)} = \cosh \beta^{(i)}u_a + \sinh \beta^{(i)}e_a. \quad (2.4.3)$$

In addition, $v_{(i)} = \tanh \beta_{(i)}$, which means that when the tilt angle is small (i.e. for $\beta_{(i)} \ll 1$) we have $v_{(i)} \simeq \beta_{(i)}$ and non-relativistic peculiar velocities.

2.4.2. Multi-component perfect fluids

The non-equilibrium state of a material medium is described by its energy momentum tensor T_{ab} , its particle flux vector N_a and by the entropy flux vector S_a . The former two are conserved ($\nabla_b T^{ab} = 0 = \nabla_a N^a$) and the last obeys the second law of thermodynamics ($\nabla_a S^a \geq 0$ —see [61] and also Section 1.3.4 here). When the strong energy condition holds, the energy-momentum tensor of a fluid has a unique timelike eigenvector u_a^E , normalised so that $u_a^E u_a^E = -1$. One may also define a unitary timelike vector parallel to N_a by $u_a^N = N_a / \sqrt{-N_a N^a}$. Provided that the fluid is perfect (or in equilibrium) all three vectors u_a^E , u_a^N and S_a are parallel and define a unique hydrodynamic 4-velocity vector (the rest-frame of the fluid flow). This is the only frame the energy-momentum tensor of the matter assumes the perfect-fluid form (see expression (1.2.13)).

When dealing with an imperfect fluid there is no a uniquely defined hydrodynamic 4-velocity. Then, the energy momentum tensor and the particle flux vector decompose as

$$T_{ab} = \rho u_a u_b + p h_{ab} + 2q_{(a} u_{b)} + \pi_{ab} \quad \text{and} \quad N_a = n u_a + \mathcal{N}_a, \quad (2.4.4)$$

respectively. Here $\mathcal{N}_a = h_a{}^b N_b$ is the particle drift and recall that $n = -N_a u^a$ is the particle number density. Also, the total energy-flux vector in Eq. (2.4.4)(a) is given by

$$q_a = Q_a + \frac{1}{n} (\rho + p) \mathcal{N}_a, \quad (2.4.5)$$

with Q_a representing the associated heat flux. In the case of an imperfect medium two frames are of special status. The first, which has $u_a \equiv u_a^E$ and zero total energy flux, is known as the energy – or Landau – frame [88]. There, following (2.4.5), the comoving observers see a nonzero particle drift equal to $\mathcal{N}_a = -[n/(\rho + p)]Q_a$. The alternative option is the Eckart – or particle – frame with $u_a \equiv u_a^N$ [116]. Here, the observers see no particle drift and therefore $q_a = Q_a$.

Consider a mixture of perfect fluids, where the i -th component has energy density $\rho^{(i)}$, isotropic pressure $p^{(i)}$ and moves along the timelike 4-velocity field $u_a^{(i)}$. Relative to this frame, the energy–momentum tensor and the particle flux of the individual species respectively read

$$T_{ab}^{(i)} = \rho^{(i)} u_a^{(i)} u_b^{(i)} + p^{(i)} h_{ab}^{(i)} \quad \text{and} \quad N_a^{(i)} = n^{(i)} u_a^{(i)}, \quad (2.4.6)$$

with $h_{ab}^{(i)}$ given by (2.4.1)(b) and $n^{(i)}$ representing the number density of each component in their own rest frame. With respect to the u_a -frame, however, the above become

$$T_{ab}^{(i)} = \hat{\rho}^{(i)} u_a u_b + \hat{p}^{(i)} h_{ab} + 2u_{(a} \hat{q}_{b)}^{(i)} + \hat{\pi}_{ab}^{(i)} \quad \text{and} \quad N_a^{(i)} = \hat{n}^{(i)} u_a + \hat{\mathcal{N}}_a^{(i)}, \quad (2.4.7)$$

respectively. The former corresponds to the energy–momentum tensor of an imperfect fluid with

$$\hat{\rho}^{(i)} = \gamma_{(i)}^2 (\rho^{(i)} + p^{(i)}) - p^{(i)}, \quad \hat{p}^{(i)} = p^{(i)} + \frac{1}{3} (\gamma_{(i)}^2 - 1) (\rho^{(i)} + p^{(i)}), \quad (2.4.8)$$

$$\hat{q}_a^{(i)} = \gamma_{(i)}^2 (\rho^{(i)} + p^{(i)}) v_a^{(i)}, \quad \text{and} \quad \hat{\pi}_{ab}^{(i)} = \gamma_{(i)}^2 (\rho^{(i)} + p^{(i)}) \left(v_a^{(i)} v_b^{(i)} - \frac{1}{3} v_{(i)}^2 h_{ab} \right). \quad (2.4.9)$$

Similarly, expression (2.4.7)(b) is that of an imperfect fluid with particle number density and particle drift given by

$$\hat{n}^{(i)} = \gamma^{(i)} n^{(i)} \quad \text{and} \quad \hat{\mathcal{N}}_a^{(i)} = \hat{n}^{(i)} v_a^{(i)}, \quad (2.4.10)$$

respectively. When dealing with non-relativistic peculiar velocities, we may ignore terms quadratic in $v_{(i)}$ and therefore set $\gamma_{(i)}$ to unity. Then, expressions (2.4.8) and (2.4.9) reduce to $\hat{\rho}^{(i)} = \rho^{(i)}$, $\hat{p}^{(i)} = p^{(i)}$, $\hat{q}_a^{(i)} = (\rho^{(i)} + p^{(i)}) v_a^{(i)}$ and $\hat{\pi}_{ab}^{(i)} = 0$ [117]. Also, (2.4.10) simplify to $\hat{n}^{(i)} = n^{(i)}$ and $\hat{\mathcal{N}}_a^{(i)} = n^{(i)} v_a^{(i)}$. From now on we will always assume that $v_{(i)} \ll 1$, unless stated otherwise.

2.4.3. Conservation laws of the total fluid

Measured relative to the u_a -frame the total fluid has an effective energy density $\rho = \Sigma_i \rho^{(i)}$, pressure $p = \Sigma_i p^{(i)}$, energy flux $q_a = \Sigma_i \hat{q}_a^{(i)}$ and

$$T_{ab} = \rho u_a u_b + p h_{ab} + 2u_{(a} q_{b)}. \quad (2.4.11)$$

The latter is conserved which means that $\nabla^b T_{ab} = 0$. Assuming an effective equation of state of the form $p = p(\rho, s)$, where s is the associated specific entropy, we have

$$D_a p = \frac{c_s^2 \rho}{a} \Delta_a + \frac{p}{a} \mathcal{E}_a, \quad (2.4.12)$$

where $c_s^2 = (\partial p / \partial \rho)_s$ and $\mathcal{E}_a = (a/p)(\partial p / \partial s)_\rho D_a s$ are the square of the effective sound speed and the effective entropy perturbation respectively. Therefore, the momentum–density conservation law of the total fluid reads

$$(\rho + p) A_a = -\frac{c_s^2 \rho}{a} \Delta_a - \frac{p}{a} \mathcal{E}_a - \dot{q}_{(a)} - \frac{4}{3} \Theta q_a - (\sigma_a^b + \omega_a^b) q_b, \quad (2.4.13)$$

while the effective total energy satisfies (1.3.17). We note that the expansion dynamics is determined by the total fluid and spatial inhomogeneities in the volume expansion are monitored via Eq. (2.3.2) with $D_a p$ and A_a given by (2.4.12) and (2.4.13) respectively.

2.4.4. Conservation laws of the i -th fluid

Assuming a mixture of interacting and non-comoving perfect fluids, the energy–momentum of the i -th species reads (see Section 2.4.2)

$$T_{ab}^{(i)} = \rho^{(i)} u_a u_b + p^{(i)} h_{ab} + 2u_{(a} q_{b)}^{(i)}, \quad (2.4.14)$$

with $q_a^{(i)} = (\rho^{(i)} + p^{(i)}) v_a^{(i)}$ and $v_a^{(i)}$ representing the peculiar velocity of the component. The above satisfies the conservation law

$$\nabla^b T_{ab}^{(i)} = \mathcal{I}_a^{(i)}, \quad (2.4.15)$$

where the interaction term has $\sum_i \mathcal{I}_a^{(i)} = 0$ due to the conservation of $T_{ab} = \sum_i T_{ab}^{(i)}$. Setting $\mathcal{I}^{(i)} = -\mathcal{I}_a^{(i)} u^a$ and $\mathcal{I}_a^{(i)} = h_a^b \mathcal{I}_b^{(i)}$, the timelike and spacelike parts of (2.4.15) give

$$\dot{\rho}^{(i)} = -\Theta (\rho^{(i)} + p^{(i)}) - D^a q_a^{(i)} - 2A^a q_a^{(i)} + \mathcal{I}^{(i)} \quad (2.4.16)$$

and

$$\left(\rho^{(i)} + p^{(i)}\right) A_a = -\frac{c_s^{2(i)} \rho^{(i)}}{a} \Delta_a^{(i)} - \frac{p^{(i)}}{a} \mathcal{E}_a^{(i)} - \dot{q}_{(a)}^{(i)} - \frac{4}{3} \Theta q_a^{(i)} - \left(\sigma_a^b + \omega_a^b\right) q_b^{(i)} + \mathcal{J}_a^{(i)}, \quad (2.4.17)$$

where $c_s^{2(i)} = (\partial p^{(i)} / \partial \rho^{(i)})_{s^{(i)}}$ and $\mathcal{E}_a^{(i)} = (a/p^{(i)}) (\partial p^{(i)} / \partial s^{(i)})_{\rho^{(i)}} D_a s^{(i)}$ are respectively the square of the sound speed and the entropy perturbation of the species in question. Also, following definition (2.2.1), $\Delta_a^{(i)} = (a/\rho^{(i)}) D_a \rho^{(i)}$ describes inhomogeneities in the density distribution of the i -th fluid component, relative to the u_a -frame. Finally we note that $\sum_i \mathcal{J}^{(i)} = 0 = \sum_i \mathcal{J}_a^{(i)}$.

2.4.5. Nonlinear density perturbations in the i -th species

Taking the time derivative of $\Delta_a^{(i)}$ and using the conservation laws (2.4.16) and (2.4.17), we arrive at the following nonlinear expression

$$\begin{aligned} \dot{\Delta}_{(a)}^{(i)} &= \frac{p^{(i)}}{\rho^{(i)}} \Theta \Delta_a^{(i)} - \left(1 + \frac{p^{(i)}}{\rho^{(i)}}\right) \mathcal{Z}_a + \frac{a \Theta}{\rho^{(i)}} \left(\dot{q}_{(a)}^{(i)} + \frac{4}{3} \Theta q_a^{(i)}\right) - \frac{a}{\rho^{(i)}} D_a \left(D^b q_b^{(i)} - \mathcal{J}^{(i)}\right) - \left(\sigma_a^b + \omega_a^b\right) \Delta_b^{(i)} - \frac{2a}{\rho^{(i)}} D_a \left(A^b q_b^{(i)}\right) \\ &+ \frac{a \Theta}{\rho^{(i)}} \left(\sigma_a^b + \omega_a^b\right) q_b^{(i)} + \frac{1}{\rho^{(i)}} \left(D^b q_b^{(i)} + 2A^b q_b^{(i)} - \mathcal{J}^{(i)}\right) \left(\Delta_a^{(i)} - a A_a\right) - \frac{a \Theta}{\rho^{(i)}} \mathcal{J}_a^{(i)}. \end{aligned} \quad (2.4.18)$$

When the interaction term is specified, the above describes the propagation of spatial inhomogeneities in the density distribution of the i -th species. Recall that $q_a^{(i)} = (\rho^{(i)} + p^{(i)}) v_a^{(i)}$ and that the nonlinear evolution of \mathcal{Z}_a is governed by Eq. (2.3.2).

2.5. Inhomogeneous magnetised cosmologies

From the Earth and the nearby stars, to galaxies, galaxy clusters and remote high-redshift protogalaxies, magnetic fields have been repeatedly observed [118–120]. Although the origin of cosmic magnetism is still a mystery, it appears that we live in a magnetised universe [121].

2.5.1. The ideal MHD approximation

With the exception of any period of inflation and early reheating, the universe has been a good conductor throughout its lifetime. As a result, B -fields of cosmological origin have remained frozen into the expanding cosmic fluid during most of their evolution. This allows us to study the magnetic effects on structure formation within the limits of the ideal magnetohydrodynamics (MHD) approximation. The latter is described by means of Ohm's law, which has the covariant form [122,123]

$$J_a = \mu u_a + \zeta E_a, \quad (2.5.1)$$

with ζ representing the scalar conductivity of the medium. Note that the quantities on the right-hand side are measured in the rest-frame of the plasma. Expression (2.5.1) splits the 4-current into a timelike part, due to the charge density, and a spacelike part,

$$\mathcal{J}_a = \zeta E_a, \quad (2.5.2)$$

from conduction. This form of Ohm's law covariantly describes the resistive magnetohydrodynamic (MHD) approximation in the single-fluid approach.¹³ Note the absence of the induced electric field from the above, reflecting the fact that the covariant form of Maxwell's formulae (see (1.3.14)–(1.3.16)) already incorporates the effects of relative motion. According to (2.5.2), zero electrical conductivity implies that the spatial currents vanish, even when the electric field is non-zero. On the other hand, non-zero spatial currents are compatible with a vanishing electric field as long as the conductivity of the medium is infinite (i.e. for $\zeta \rightarrow \infty$). Thus, at the ideal MHD limit, the electric field vanishes in the frame of the fluid and the energy–momentum tensor of the residual magnetic field simplifies to [109]

$$T_{ab}^{(B)} = \frac{1}{2} B^2 u_a u_b + \frac{1}{6} B^2 h_{ab} + \Pi_{ab}, \quad (2.5.3)$$

with $\Pi_{ab} = -B_{(a} B_{b)}$. This means that the B -field corresponds to an imperfect fluid with energy density $\rho_B = B^2/2$, isotropic pressure $p_B = B^2/6$ and anisotropic stresses represented by Π_{ab} . Similarly, Maxwell's equations (see Section 1.3.2) reduce to a single propagation formula,

$$\dot{B}_{(a)} = \left(\sigma_{ab} + \varepsilon_{abc} \omega^c - \frac{2}{3} \Theta h_{ab}\right) B^b, \quad (2.5.4)$$

¹³ Ohm's law relates the induced 3-current with the electric field and it is generally given in the form of a propagation equation of the electric current. For the fully nonlinear 1 + 3 covariant version of the generalised Ohm's law in the case of a hot multi-component plasma the reader is referred to [124].

and to the following three constraints

$$\text{curl } B_a + \varepsilon_{abc} A^b B^c = \mathcal{J}_a, \quad (2.5.5)$$

$$2\omega^a B_a = \mu, \quad D^a B_a = 0. \quad (2.5.6)$$

The right-hand side of (2.5.4) is due to the relative motion of the neighbouring observers and guarantees that the magnetic field lines always connect the same matter particles, namely that the field remains frozen-in with the highly conducting fluid. Expression (2.5.5) provides a direct relation between the spatial currents, which are responsible for keeping the field lines frozen-in with the matter, and the magnetic field itself (e.g. see [125]).

Contracting the magnetic induction Eq. (2.5.4) along B_a leads to the evolution law for the energy density of the field, which takes the nonlinear form

$$\dot{(B^2)} = -\frac{4}{3} \Theta B^2 - 2\sigma_{ab} \Pi^{ab}. \quad (2.5.7)$$

This shows that in a highly conducting cosmic medium we have $B^2 \propto a^{-4}$ always unless there is substantial anisotropy, in which case the B -field behaves as a dissipative radiative fluid. Note that in a spatially homogeneous, radiation-dominated universe with weak overall anisotropy, the shear term in the right-hand side of (2.5.7) means that the ratio $B^2/\rho^{(\gamma)}$ is no longer constant but displays a slow ‘quasi static’ logarithmic decay (see [126] and also [127,128]).

2.5.2. Conservation laws

Covariant studies of cosmic electromagnetic fields date back to the work of Elhers and Ellis (see also [129] for an analogous approach), while the Newtonian version of the relativistic approach was recently given in [130]. Following [42], the energy momentum tensor corresponding to a magnetised single perfect fluid of infinite conductivity is

$$T_{ab} = \left(\rho + \frac{1}{2} B^2\right) u_a u_b + \left(p + \frac{1}{6} B^2\right) h_{ab} + \Pi_{ab}, \quad (2.5.8)$$

and the medium behaves as an imperfect fluid with effective density $\rho + B^2/2$, isotropic pressure $p + B^2/6$ and solely magnetic anisotropic stresses represented by Π_{ab} . Applied to the above, and using the MHD form of Maxwell’s equations, the conservation law $\nabla^b T_{ab} = 0$ splits into two expressions that respectively describe energy and momentum conservation¹⁴

$$\dot{\rho} = -(\rho + p)\Theta, \quad (2.5.9)$$

$$\left(\rho + p + \frac{2}{3} B^2\right) A_a = -D_a p - \varepsilon_{abc} B^b \text{curl } B^c - \Pi_{ab} A^b. \quad (2.5.10)$$

Note the absence of magnetic terms in Eq. (2.5.9), since only the electric field contributes to (1.3.20). The magnetic energy is separately conserved, as guaranteed by the magnetic induction equation (2.5.4). The second last term in (2.5.10), which is often referred to as the magnetic Lorentz force, is always normal to the B -field lines and decomposes as

$$\varepsilon_{abc} B^b \text{curl } B^c = \frac{1}{2} D_a B^2 - B^b D_b B_a. \quad (2.5.11)$$

The last term in the above is the result of the magnetic tension. In so far as this tension stress is not balanced by the pressure gradients, the field lines are out of equilibrium and there is a non-zero Lorentz force acting on the magnetised fluid.

2.5.3. Nonlinear density perturbations of the magnetised fluid

In the presence of magnetic fields, the nonlinear evolution of spatial inhomogeneities in the density distribution of a single, highly conducting perfect fluid is described by the expression

$$\dot{\Delta}_{(a)} = \frac{p}{\rho} \Theta \Delta_a - \left(1 + \frac{p}{\rho}\right) \mathcal{Z}_a + \frac{a\Theta}{\rho} \varepsilon_{abc} B^b \text{curl } B^c + \frac{2a\Theta B^2}{3\rho} A_a - (\sigma_{ba} + \omega_{ba}) \Delta^b + \frac{a\Theta}{\rho} \Pi_{ab} A^b. \quad (2.5.12)$$

This is obtained after taking the projected time-derivative of definition (2.2.1) and then using relations (2.5.9) and (2.5.10). Similarly, starting from (2.2.3)(a) and using (2.5.9) we arrive at the nonlinear evolution equation for the expansion gradients

$$\begin{aligned} \dot{\mathcal{Z}}_{(a)} = & -\frac{2}{3} \Theta \mathcal{Z}_a - \frac{1}{2} \rho \Delta_a - \frac{1}{2} B^2 \mathcal{B}_a + \frac{3}{2} a \varepsilon_{abc} B^b \text{curl } B^c + a D_a D^b A_b + 2a A^b D_a A_b \\ & + \left[\frac{1}{2} \mathcal{R} - 3(\sigma^2 - \omega^2) + D^b A_b + A_b A^b \right] A_a - (\sigma_{ba} + \omega_{ba}) \mathcal{Z}^b + \frac{3}{2} a \Pi_{ab} A^b - 2a D_a (\sigma^2 - \omega^2), \end{aligned} \quad (2.5.13)$$

¹⁴ One can also obtain the conservation laws (2.5.9) and (2.5.10) by taking the MHD limit of expressions (1.3.20) and (1.3.21) of Section 1.3.3 and then substituting in the 3-current from constraint (2.5.5).

where \mathcal{R} is the Ricci scalar of the magnetised 3-D space orthogonal to the fluid motion. Finally, the non-linear propagation formula monitoring spatial inhomogeneities in the magnetic energy density (see definition (2.2.3)(b)) is

$$\begin{aligned} \dot{\mathcal{B}}_{(a)} &= \frac{4\rho}{3(\rho+p)} \dot{\Delta}_{(a)} - \frac{4p\Theta}{3(\rho+p)} \Delta_a - \frac{4a\Theta}{3(\rho+p)} \varepsilon_{abc} B^b \text{curl} B^c - \frac{4}{3} a\Theta \left[1 + \frac{2B^2}{3(\rho+p)} \right] A_a \\ &\quad - (\sigma_{ba} + \omega_{ba}) \mathcal{B}^b + \frac{4\rho}{3(\rho+p)} (\sigma_{ba} + \omega_{ba}) \Delta^b - \frac{4a\Theta}{3(\rho+p)} \Pi_{ab} A^b - \frac{2a}{B^2} \Pi^{bc} D_a \sigma_{bc} \\ &\quad - \frac{2a}{B^2} \sigma^{bc} D_a \Pi_{bc} + \frac{2}{B^2} \sigma_{bc} \Pi^{bc} \mathcal{B}_a - \frac{2a}{B^2} \sigma_{bc} \Pi^{bc} A_a. \end{aligned} \quad (2.5.14)$$

The above results from the time-derivative of (2.2.3)(b) by means of (2.5.7) and (2.5.13) and cannot be used when the magnetised medium has $p = -\rho$.

2.6. Inhomogeneous scalar-field cosmologies

Scalar-field dominated universes have come into prominence primarily through the inflationary scenarios. Covariantly, scalar-field cosmologies have been discussed in a series of papers by Madsen and Ellis [131–133] and more recently by Vernizzi and Langlois [111,134].

2.6.1. Minimally coupled scalar fields

Consider a general, pseudo-Riemannian spacetime filled with a single scalar field (φ), which is minimally coupled to gravity. The associated Lagrangian density is

$$\mathcal{L}_\varphi = -\sqrt{-g} \left[\frac{1}{2} \nabla_a \varphi \nabla^a \varphi + V(\varphi) \right], \quad (2.6.1)$$

where g is the determinant of the spacetime metric and $V(\varphi)$ is the effective potential that describes the self-interaction of the scalar field. The stress–energy tensor associated with (2.6.1) has the form

$$T_{ab}^{(\varphi)} = \nabla_a \varphi \nabla_b \varphi - \left[\frac{1}{2} \nabla_c \varphi \nabla^c \varphi + V(\varphi) \right] g_{ab}. \quad (2.6.2)$$

Applying the twice contracted Bianchi identities, namely the conservation law $\nabla^b T_{ab} = 0$, to the above and assuming that $\nabla_a \varphi \neq 0$ we arrive at the familiar Klein–Gordon equation

$$\nabla^a \nabla_a \varphi - V'(\varphi) = 0, \quad (2.6.3)$$

where in this case the prime indicates differentiation with respect to φ . We note that when $\nabla_a \varphi = 0$ expression (2.6.2) reduces to $T_{ab}^{(\varphi)} = -V(\varphi) g_{ab}$. The latter ensures that $\nabla_a V(\varphi) = 0$, since $\nabla^b T_{ab} = 0$, and φ behaves as an effective cosmological constant rather than a dynamical scalar field.

2.6.2. Scalar-field kinematics

In order to achieve an 1 + 3 covariant fluid-description of scalar fields, one first needs to assign a 4-velocity vector to the φ -field itself. Suppose that the 4-vector $\nabla_a \varphi$ is timelike, namely that $\nabla_a \varphi \nabla^a \varphi < 0$ over an open spacetime region. In this case $\nabla_a \varphi$ defines the normals to the spacelike hypersurfaces $\varphi(x^a) = \text{constant}$ and we define our 4-velocity field by [112,132]

$$u_a = -\frac{1}{\dot{\varphi}} \nabla_a \varphi, \quad (2.6.4)$$

with $\dot{\varphi} = u^a \nabla_a \varphi \neq 0$. This means that $\dot{\varphi}^2 = -\nabla_a \varphi \nabla^a \varphi > 0$ and $u_a u^a = -1$ as required. Also, the flow vector u_a defines our time direction and introduces a unique threading of the spacetime into time and space. The metric of the 3-space orthogonal to u_a is represented by the projection tensor

$$h_{ab} = g_{ab} + \frac{1}{\dot{\varphi}^2} \nabla_a \varphi \nabla_b \varphi, \quad (2.6.5)$$

which satisfies the standard requirements $h_{ab} = h_{(ab)}$, $h_{ab} u^b = 0$, $h_a^a = 3$ and $h_{ab} h^b{}_c = h_{ac}$. In addition, h_{ab} defines the covariant derivative operator $D_a = h_a^b \nabla_b$ orthogonal to u_a and therefore guarantees that

$$D_a \varphi = 0, \quad (2.6.6)$$

always. This result is a key feature of covariant scalar-field cosmologies and, as we will see in the following sections, it ‘dominates’ the fluid description of the φ -field and essentially dictates all aspects of its evolution.

The irreducible variables describing the kinematics of $u_a = -\nabla_a\varphi/\dot{\varphi}$ in a covariant manner are obtained in the usual way by means of the decomposition (1.3.1). More specifically, starting from definition (2.6.4), using the Klein–Gordon equation and the fact $D_a\varphi = 0$, we arrive at [135]

$$\Theta = -\frac{1}{\dot{\varphi}} [\ddot{\varphi} + V'(\varphi)], \quad \omega_{ab} = 0, \quad \sigma_{ab} = -\frac{1}{\dot{\varphi}} h_a{}^c h_b{}^d \nabla_c \nabla_d \varphi + \frac{1}{3\dot{\varphi}} [\ddot{\varphi} + V'(\varphi)] h_{ab} \quad (2.6.7)$$

and

$$A_a = -\frac{1}{\dot{\varphi}} D_a \dot{\varphi}. \quad (2.6.8)$$

Hence, the 4-velocity field (2.6.4) is irrotational and $\dot{\varphi}$ acts as an acceleration potential for the fluid flow. These expressions provide the irreducible kinematical variables associated with a minimally coupled scalar field. The reader is referred to [132] for a more general description of scalar-field kinematics, of which relations (2.6.7) and (2.6.8) emerge as a special case.

2.6.3. Scalar fields as perfect fluids

The introduction of the timelike velocity field (2.6.4) also allows for a dynamically convenient fluid-description of scalar fields. In particular, by means of (2.6.4) the energy–momentum tensor (2.6.2) has perfect-fluid form¹⁵

$$T_{ab}^{(\varphi)} = \rho^{(\varphi)} u_a u_b + p^{(\varphi)} h_{ab}, \quad (2.6.9)$$

with

$$\rho^{(\varphi)} = \frac{1}{2} \dot{\varphi}^2 + V(\varphi) \quad \text{and} \quad p^{(\varphi)} = \frac{1}{2} \dot{\varphi}^2 - V(\varphi). \quad (2.6.10)$$

Clearly, if we demand a positive definite energy density for the field, then $\dot{\varphi}^2 + 2V(\varphi) > 0$.

The two simplest cases correspond to a free scalar field with purely kinetic energy (i.e. $V(\varphi) = 0$), which has $p^{(\varphi)} = \rho^{(\varphi)}$ and behaves like a stiff-matter component. If the field's energy is purely potential, on the other hand, $p^{(\varphi)} = -\rho^{(\varphi)}$. Such an (effective) equation of state is approximately achieved during the slow-rolling regime of the standard inflationary scenarios. Then, the potential energy dominates over the kinetic, which means that $\dot{\varphi}^2 \ll V(\varphi)$.

In general, either

$$p^{(\varphi)} = \rho^{(\varphi)} - 2V(\varphi), \quad (2.6.11)$$

or $p^{(\varphi)} = \dot{\varphi}^2 - \rho^{(\varphi)}$. Therefore, minimally coupled scalar fields do not generally behave like barotropic media. Instead, the φ -field may be regarded as bulk viscous fluid with the potential playing the role of bulk viscosity [131].

2.6.4. Conservation laws

When applied to the energy–momentum tensor of a minimally coupled scalar field, the twice-contracted Bianchi identities lead to the familiar Klein–Gordon equation (see expression (2.6.3) in Section 2.6.1). Once the timelike 4-velocity vector (2.6.4) is introduced, we may use condition $D_a\varphi = 0$ to write $\nabla_a\varphi = -\dot{\varphi}u_a$. Then, the Klein–Gordon equation acquires its 1 + 3 covariant form

$$\ddot{\varphi} + \Theta\dot{\varphi} + V'(\varphi) = 0. \quad (2.6.12)$$

The above can also be seen as the energy conservation law associated with the perfect fluid description of a minimally coupled scalar field. Indeed, after substituting expressions (2.6.10) into the standard energy conservation equation we immediately recover (2.6.12). On the other hand, the conservation of the momentum is given by expression (2.6.8). This can be verified by inserting relations (2.6.10) into the familiar momentum conservation law of a single perfect fluid (see Eq. (1.3.19)(b) in Section 1.3.3).

2.6.5. Nonlinear scalar-field perturbations

Taking the orthogonally projected gradients of Eq. (2.6.11), which can be seen as the equation of state of a minimally coupled scalar field, we obtain

$$D_a p^{(\varphi)} = D_a \rho^{(\varphi)} = \dot{\varphi} D_a \dot{\varphi}. \quad (2.6.13)$$

This relation is guaranteed by our 4-velocity choice, which ensures that $D_a\varphi = 0$ always (see Eq. (2.6.6)) and therefore that $D_a V(\varphi) = 0$ as well. Introducing the variable $\Delta_a^{(\varphi)} = (a/\rho^{(\varphi)}) D_a \rho^{(\varphi)}$ that describes inhomogeneities in the (effective) energy density of the φ -field, expression (2.6.13) leads to

$$a D_a p^{(\varphi)} = \rho^{(\varphi)} \Delta_a^{(\varphi)}, \quad (2.6.14)$$

¹⁵ A non-minimally coupled scalar field corresponds to an imperfect medium [132].

or, equivalently, $\Delta_a^{(\varphi)} = (a\dot{\varphi}/\rho^{(\varphi)})D_a\dot{\varphi}$. From (2.6.13) follows that, despite its non-barotropic equation of state, in hydrodynamic terms the scalar field behaves as an effective stiff fluid. The non-barotropic nature of the φ -field emerges in the associated Γ_a -parameter (see Section 2.3.3), which does not vanish. Indeed, for a minimally coupled scalar field, we find that

$$\Gamma_a = D_a p^{(\varphi)} - \left(\frac{\dot{p}^{(\varphi)}}{\dot{\rho}^{(\varphi)}} \right) D_a \rho^{(\varphi)} = -\frac{2V'(\varphi)}{\Theta\dot{\varphi}} D_a \rho^{(\varphi)} \neq 0, \quad (2.6.15)$$

unless of course $V'(\varphi) = 0$. The aforementioned ‘duality’ in the nature of scalar fields represents a major departure from conventional perfect-fluid behaviour and is reflected in the statement that scalar-field perturbations are ‘non-adiabatic’.

Setting $\Delta_a^{(\varphi)} = (a/\rho^{(\varphi)})D_a\rho^{(\varphi)}$, $\Lambda = 0$ and using expression (2.6.14), we may adapt the system (2.3.1) and (2.3.2) to a minimally coupled scalar field. To be precise, we obtain the propagation equations

$$\dot{\Delta}_{(a)}^{(\varphi)} = \frac{p^{(\varphi)}}{\rho^{(\varphi)}} \Theta \Delta_a^{(\varphi)} - \left(1 + \frac{p^{(\varphi)}}{\rho^{(\varphi)}} \right) \mathcal{Z}_a - (\sigma^b{}_a + \omega^b{}_a) \Delta_b^{(\varphi)}, \quad (2.6.16)$$

and

$$\begin{aligned} \dot{\mathcal{Z}}_{(a)} = & -\frac{2}{3} \Theta \mathcal{Z}_a - 2\rho^{(\varphi)} \Delta_a^{(\varphi)} - a \left[\frac{1}{3} \Theta^2 + \frac{1}{2} (\rho^{(\varphi)} + 3p^{(\varphi)}) \right] A_a + a D_a D^b A_b - (\sigma^b{}_a + \omega^b{}_a) \mathcal{Z}_b \\ & - 2a D_a (\sigma^2 - \omega^2) + 2a A^b D_a A_b - a \left[2 (\sigma^2 - \omega^2) - D^b A_b - A^b A_b \right] A_a, \end{aligned} \quad (2.6.17)$$

respectively. Finally, combining (2.6.8) and (2.6.14) we arrive at the following expression for the 4-acceleration of the u_a -congruence defined in Section 2.6.2 (see Eq. (2.6.4) there)

$$a(\rho^{(\varphi)} + p^{(\varphi)})A_a = -\rho^{(\varphi)} \Delta_a^{(\varphi)}. \quad (2.6.18)$$

The above, together with Eqs. (2.6.16) and (2.6.17), monitors the evolution of nonlinear density perturbations in scalar-field dominated cosmologies.

3. Linear cosmological perturbations

3.1. Linearisation

Before linearising the full equations of the previous section, one needs to select the exact solution that will provide the unperturbed background about which the nonlinear formulae will be linearised. Terms with nonzero unperturbed value are assigned zero perturbative order, while those that vanish in the background are treated as first order perturbations [104, 136]. This guarantees that linear variables satisfy the criterion for gauge-invariance [106]. By definition, all the first-order quantities are assumed weak relative to the background ones and of perturbative order $\mathcal{O}(\epsilon)$ in a smallness parameter ϵ (e.g. see [82]). When linearising, products of $\mathcal{O}(\epsilon)$ variables are neglected.

Cosmological applications use the homogeneous and isotropic FLRW spacetime as the unperturbed zero-order model. When dealing with a Friedmann–Lemaître background, the only zero-order quantities are the matter energy density ρ , isotropic pressure p and the volume expansion $\Theta = 3H$. If the geometry of spatial sections is not Euclidean, these are supplemented by the 3-Ricci scalar \mathcal{R} .

3.1.1. Interpreting the inhomogeneity variables

To linear order, the inhomogeneity variables of Section 2.2.2 describe measurable spatial differences in the distribution of the physical quantities in question. To show this, consider a connecting vector between two worldlines of u_a (so that its Lie derivative along u_a vanishes). The projection of this vector perpendicular to u_a defines a spatial vector χ_a , which connects events on neighbouring worldlines that are simultaneous (as seen by u_a – see Section 1.3.1). Then, $\chi_a u^a = 0$ and

$$\dot{\chi}_{(a)} = \frac{1}{3} \Theta \chi_a + (\sigma_{ab} + \omega_{ab}) \chi^b, \quad (3.1.1)$$

which means that in a FLRW universe $\chi_a = a\chi_a^0$ (with $\chi_a^0 = \text{constant}$) [44]. If we Taylor expand the density about some worldline, then the density in the local rest space of u_a at displacement χ_a has

$$\tilde{\rho} - \rho = \chi^a D_a \rho, \quad (3.1.2)$$

to leading order. Using the definition (2.2.1) and the fact that $\chi_a \propto a$ in the FLRW background, the above translates into

$$\delta\rho \propto \chi_0^a \Delta_a, \quad (3.1.3)$$

where $\delta\rho = (\tilde{\rho} - \rho)/\rho$. In other words, the comoving fractional gradient Δ_a describes the measurable local density variation between two neighbouring fundamental observers. Moreover, Δ_a closely corresponds to the familiar energy–density contrast of the non-covariant studies.

3.2. Single-fluid perturbations

3.2.1. Linear evolution equations

Consider an FLRW universe filled with a single barotropic perfect fluid (i.e. assume that $p = p(\rho)$). On this background linear inhomogeneities in the energy density of the medium evolve according to

$$\dot{\Delta}_a = 3wH\Delta_a - (1+w)\mathcal{Z}_a. \quad (3.2.1)$$

The above is obtained from (2.3.3) after dropping its nonlinear terms and assuming a barotropic equation of state for the cosmic fluid. Also the parameter $w = p/\rho$ is the barotropic index of the fluid with

$$\dot{w} = -3H(1+w)(c_s^2 - w). \quad (3.2.2)$$

This means that $w = \text{constant}$ when $c_s^2 = w$ provided $H \neq 0$ and $w \neq -1$ (as happens during the radiation and dust eras, for example). Within the same almost-FLRW environment expression (2.3.2) linearises to

$$\dot{\mathcal{Z}}_a = -2H\mathcal{Z}_a - \frac{1}{2}\rho\Delta_a - \frac{3}{2}aD_a p - a \left[3H^2 + \frac{1}{2}\rho(1+3w) - \Lambda \right] A_a + aD_a D^b A_b, \quad (3.2.3)$$

where $\rho(1+w)A_a = -D_a p$ to linear order (see (1.3.19)(b)). Given the barotropic nature of the medium, we may define $c_s^2 = dp/d\rho$ as the square of the adiabatic sound speed. Then, the linearised conservation law for the momentum density reads

$$a(1+w)A_a = -c_s^2 \Delta_a. \quad (3.2.4)$$

Solving this for A_a , substituting the result into the right-hand side of (3.2.3) and keeping terms of up to linear order, we obtain

$$\dot{\mathcal{Z}}_a = -2H\mathcal{Z}_a - \frac{1}{2}\rho\Delta_a - \frac{c_s^2}{1+w} \left(D^2 \Delta_a + \frac{K}{a^2} \Delta_a \right) - 6ac_s^2 H \text{curl } \omega_a, \quad (3.2.5)$$

where $K = 0, \pm 1$ is the 3-curvature index of the FLRW background and $D^2 = D^a D_a$ is the orthogonally projected Laplacian operator. Note that in deriving the above we have also used the linear commutation laws between the orthogonally projected gradients of scalars and spacelike vectors (see Appendix A.3), as well as the zero-order Friedmann equation $\rho - 3H^2 + \Lambda = 3K/a^2$ (see Section 1.4.2).

The linear expansion of this model is determined by the associated expression of Raychaudhuri's formula (1.3.3), written in the form

$$qH^2 = \frac{1}{6}\rho(1+3w) + \frac{c_s^2}{3a^2(1+w)} \Delta - \frac{1}{3} \Lambda, \quad (3.2.6)$$

on using (3.2.4). Note that q is the deceleration parameter of the perturbed spacetime (see Eq. (1.4.7) in Section 1.4.2) and $\Delta = aD^a \Delta_a$ describes scalar density perturbations (see Section 3.2.2). When positive, the latter represents overdensities and adds to the gravitational pull of the matter. In the opposite case, Δ corresponds to an underdense region and tends to accelerate the expansion.

Substituting (3.2.4) into the linearised counterpart of expression (1.3.4), we find that, to first order, the shear anisotropy evolves as

$$\dot{\sigma}_{ab} = -2H\sigma_{ab} - E_{ab} - \frac{c_s^2}{a^2(1+w)} \Delta_{(ab)}, \quad (3.2.7)$$

where $\Delta_{(ab)}$ represents anisotropies in the distribution of the density gradients (see Section 3.2.2). Finally, the rotational behaviour of a perturbed FLRW universe, containing a single barotropic fluid, is monitored by the propagation equation of the vorticity vector. Starting from (1.3.5) and using (3.2.4), together with the linearised commutation law (A.28), we arrive at

$$\dot{\omega}_a = -2 \left(1 - \frac{3}{2}c_s^2 \right) H\omega_a. \quad (3.2.8)$$

Thus, kinematic vortices decay with the expansion unless the barotropic medium has an equation of state 'stiffer' than $w = 2/3$ (e.g. see [137]).

The above propagation equations are supplemented by a set of three linear constraints (see (1.3.6)–(1.3.8) for the nonlinear expressions), namely by

$$D^b \sigma_{ab} = \frac{2}{3} D_a \Theta + \text{curl } \omega_a, \quad D^a \omega_a = 0 \quad (3.2.9)$$

and

$$H_{ab} = \text{curl } \sigma_{ab} + D_{(a} \omega_{b)}. \quad (3.2.10)$$

When the right-hand side of (3.2.9)(a) vanishes, the shear is also transverse and describes pure-tensor perturbations, namely gravitational waves (see Section 3.6 below). Also, following (3.2.9)(b) and (3.2.10), ω_a is a linear solenoidal vector and in the absence of rotation the magnetic Weyl component is fully determined by the shear.

Additional constraints between the kinematical and the dynamical quantities are obtained by the linearised Gauss–Codacci formula (see Section 1.3.5), which takes the form

$$\mathcal{R}_{ab} = \frac{1}{3} \mathcal{R}h_{ab} - H(\sigma_{ab} + \omega_{ab}) + E_{ab}, \quad (3.2.11)$$

where

$$\mathcal{R} = 2 \left(\rho - \frac{1}{3} \Theta^2 + \Lambda \right), \quad (3.2.12)$$

may be seen as the linear counterpart of Friedmann's equation.

3.2.2. Three types of inhomogeneity

The comoving, orthogonally projected gradient Δ_a contains collective information about three types of inhomogeneities: density perturbations (a scalar mode); vortices (a vector mode) and shape distortions. Noting that the latter is not a pure tensor mode, but a combination of scalar and vector perturbations, we refer the reader to Appendix A.4 for details on the covariant definition of pure scalar, vector and tensor modes. All the information is encoded in the dimensionless projected gradient $\Delta_{ab} = aD_b\Delta_a$ and it is decoded by splitting Δ_{ab} into its irreducible components as

$$\Delta_{ab} = \Delta_{(ab)} + \Delta_{[ab]} + \frac{1}{3} \Delta h_{ab}, \quad (3.2.13)$$

where $\Delta_{(ab)} = aD_{(b}\Delta_{a)}$, $\Delta_{[ab]} = aD_{[b}\Delta_{a]}$ and $\Delta = aD^a\Delta_a$ [138]. The first of these variables describes variations in the anisotropy pattern of the gradient field (e.g. pancakes or cigar-like structures). The second term in the right-hand side of (3.2.13) is related to magnitude preserving changes of Δ_a (i.e. vortex-like distortions), while Δ monitors scalar variations in the spatial distribution of the matter (i.e. overdensities or underdensities).

3.2.3. Density perturbations

Linearising the orthogonally projected gradients of (3.2.1) and (3.2.3) and then taking the trace of the resulting expressions we find that

$$\dot{\Delta} = 3wH\Delta - (1+w)\mathcal{Z} \quad (3.2.14)$$

and

$$\dot{\mathcal{Z}} = -2HZ - \left[\frac{1}{2} \rho + \frac{3Kc_s^2}{a^2(1+w)} \right] \Delta - \frac{c_s^2}{1+w} D^2 \Delta, \quad (3.2.15)$$

respectively. This system governs the linear gravitational clumping of matter in a perturbed almost-FLRW universe filled with a single perfect fluid. Note that in deriving Eq. (3.2.15) we have used the linear result $D^a \text{curl } \omega_a = 0$. Taking the time derivative of (3.2.14), recalling expression (3.2.2) and using the background relations (1.4.4), (1.4.6), we arrive at [136]

$$\ddot{\Delta} = -2 \left(1 - 3w + \frac{3}{2} c_s^2 \right) H \dot{\Delta} + \left[\left(\frac{1}{2} + 4w - 3c_s^2 - \frac{3}{2} w^2 \right) \rho + (5w - 3c_s^2) \Lambda - \frac{12(w - c_s^2)K}{a^2} \right] \Delta + c_s^2 D^2 \Delta. \quad (3.2.16)$$

This wave-like equation, with extra terms due to gravity and the universal expansion, monitors the linear evolution of matter aggregations in a almost-FLRW universe filled with a single barotropic perfect fluid.

In deriving expression (3.2.16) we have assumed that $\dot{w} \neq 0$. Although the equation of state of the cosmic medium changes during the lifetime of the universe, it remains essentially constant throughout several cosmological epochs of interest (e.g. $w = 1/3$ during the radiation era and $w = 0$ after equality). Setting $w = \text{constant}$ means $w = c_s^2$ (see (3.2.2)) and simplifies Eq. (3.2.16) to

$$\ddot{\Delta} = -2 \left(1 - \frac{3}{2} w \right) H \dot{\Delta} + \left[\frac{1}{2} \rho(1-w)(1+3w) + 2\Lambda w \right] \Delta + c_s^2 D^2 \Delta, \quad (3.2.17)$$

which is now independent of the background 3-curvature. On introducing the standard scalar harmonics $\mathcal{Q}^{(k)}$, we may expand the perturbed variable in a series of harmonics modes as $\Delta = \Sigma_k \Delta^{(k)} \mathcal{Q}^{(k)}$, where $D_a \Delta^{(k)} = 0$ and k is the comoving wavenumber of the associated mode (see Appendix A.4 for details). Then, expression (3.2.17) takes the form

$$\ddot{\Delta}^{(k)} = -2 \left(1 - \frac{3}{2} w \right) H \dot{\Delta}^{(k)} + \left\{ \frac{1}{2} [\rho(1-w)(1+3w) + 4\Lambda w] - \frac{k^2 c_s^2}{a^2} \right\} \Delta^{(k)}. \quad (3.2.18)$$

The last term in the right-hand side of the above demonstrates the competing effects of gravitational attraction and pressure support, with collapse occurring when the quantity within the angled brackets is positive. Noting that $\lambda_k = a/k$ is the physical wavelength of the perturbed scalar mode, we conclude that gravitational contraction will take place only on scales larger than the associated Jeans length, $\lambda_k > \lambda_J$, where

$$\lambda_J \simeq \frac{c_s}{\sqrt{\rho(1-w)(1+3w)+4\Lambda w}}. \quad (3.2.19)$$

The above means that when the cosmic medium has a ‘stiff’ equation of state (i.e. for $w = 1$), the linear gravitational aggregation of the perturbed matter component is only supported by a positive cosmological constant (see also (3.2.18)). The Jeans scale also determines an associated Jeans mass. In the absence of a cosmological constant, the latter is given by

$$M_J \propto \rho_b \lambda_J^3 \simeq 10^{16} \left(\frac{\Omega_b}{\Omega} \right) (\Omega h^2)^{-1/2} M_\odot, \quad (3.2.20)$$

which is approximately the mass of a supercluster of galaxies. Note that Ω , Ω_b are respectively the total and the baryonic density parameters, while $H = 100 \text{ h km s}^{-1} \text{ Mpc}^{-1}$ (e.g. see [139]).

Setting $K = 0 = \Lambda$ we will seek analytical solutions to Eq. (3.2.18). During the radiation era (i.e. when $w = 1/3 = c_s^2$), $H = 1/(2t)$, $\rho = 3/(4t^2)$ and in a comoving frame we have

$$\frac{d^2 \Delta^{(k)}}{dt^2} + \frac{1}{2t} \frac{d\Delta^{(k)}}{dt} - \frac{1}{2t^2} \left[1 - \frac{1}{6} \left(\frac{k}{aH} \right)^2 \right] \Delta^{(k)} = 0, \quad (3.2.21)$$

where $k/aH = \lambda_H/\lambda_k$ and $\lambda_H = 1/H = 2t$. On super-Hubble scales, where $k/aH \ll 1$ and the pressure support is negligible, the solution is

$$\Delta = \Delta_1 t + \Delta_2 t^{-1/2}, \quad (3.2.22)$$

with $\Delta_{1,2} = \text{constant}$. Therefore, during the radiation era, large-scale matter aggregations grow as $\Delta \propto a^2$. Well inside the horizon, $k/aH \gg 1$, pressure gradients support against gravitational collapse and the over-density oscillates according to

$$\Delta^{(k)} = \Delta_1^{(k)} \sin \left[\sqrt{3} \frac{k}{a_0 H_0} \left(\frac{t}{t_0} \right)^{1/2} \right] + \Delta_2^{(k)} \cos \left[\sqrt{3} \frac{k}{a_0 H_0} \left(\frac{t}{t_0} \right)^{1/2} \right], \quad (3.2.23)$$

where $\Delta_{1,2}^{(k)} = \text{constant}$ and the zero suffix corresponds to a given initial time.

After matter-radiation equality $w = 0 = c_s^2$ and (for $K = 0 = \Lambda$) we have $H = 2/(3t)$ and $\rho = 4/(3t^2)$. In these conditions Eq. (3.2.18) leads to the following scale-independent evolution

$$\Delta = \Delta_1 t^{2/3} + \Delta_2 t^{-1}, \quad (3.2.24)$$

for the density contrast. Accordingly, matter aggregations in the post-recombination universe grow proportionally to the scale factor. Not surprisingly this relativistic result is identical to the one obtained through a Newtonian treatment (e.g. see [139–142]).

Beyond decoupling, the photons can no longer provide pressure support and gravitational attraction is only counterbalanced by ordinary baryonic gas pressure. At the time of recombination the latter is $p^{(b)} \simeq n^{(b)} k_B T_{\text{rec}}$, while its radiation counterpart is given by $p^{(\gamma)} \simeq n^{(\gamma)} k_B T_{\text{rec}}$ (with k_B representing Boltzmann’s constant). Because $n^{(b)} \simeq 10^{-8} n^{(\gamma)}$, the pressure support drops drastically at decoupling and soon after recombination the Jeans mass reduces to

$$M_J \simeq 10^4 \left(\frac{\Omega_b}{\Omega} \right) (\Omega h^2)^{-1/2} M_\odot, \quad (3.2.25)$$

which is close to that of a star cluster [139].

3.2.4. Isocurvature perturbations

One can define as isocurvature perturbations those occurring on hypersurfaces of uniform curvature, namely fluctuations which maintain $D_a \mathcal{R} = 0$ at all times [104]. This should be distinguished from the definition typically found in the literature, where the term isocurvature usually means distortions in multi-component systems with zero perturbation in the total energy–density initially (e.g. see [143, 144]).

Isocurvature fluctuations also require zero vorticity to guarantee the integrability of the 3-D hypersurfaces orthogonal to u_a . Then, the condition for isocurvature perturbations is obtained by linearising the orthogonally projected gradient of (1.3.40). To be precise, using definition (2.2.1), (2.2.3)(a) we arrive at

$$aD_a \mathcal{R} = 2\rho \Delta_a - 4HZ_a. \quad (3.2.26)$$

When $D_a \mathcal{R} = 0$, the right-hand side of the above ensures that linear expansion gradients and those in the fluid are connected by a simple algebraic relation. The projected comoving divergence of the latter translates into the following linear constraint between the associated scalar variables

$$2HZ = \rho \Delta. \quad (3.2.27)$$

Assuming a spatially flat background, this is a self-maintained condition for pressure-free dust, but holds on large scales only when the fluid has non-zero pressure [104]. In particular, propagating (3.2.26) along the observer's worldline and keeping up to linear order terms gives

$$(D_a \mathcal{R})^\cdot = -3HD_a \mathcal{R} + \frac{4c_s^2 H}{a(1+w)} D^2 \Delta_a, \quad (3.2.28)$$

since the vorticity has already been switched off. This result shows that the linear isocurvature condition is self-maintained when the fluid is pressure-free or at long-wavelengths, where the Laplacian terms on the right-hand side of Eq. (3.2.28) can be neglected.

Imposing the isocurvature condition (3.2.27) to the linear system (3.2.14) and (3.2.15), allows us to eliminate the expansion inhomogeneities (i.e. the variable \mathcal{Z}) from Eq. (3.2.14). The latter then acquires the simple form

$$\dot{\Delta}^{(\text{iso})} = -\frac{3}{2} (1-w)H\Delta^{(\text{iso})}, \quad (3.2.29)$$

which monitors the linear evolution of isocurvature scalar/density perturbations on a spatially flat FLRW background filled with a single perfect fluid. Thus, within our scheme, isocurvature perturbations can be defined and treated in single as well as in multi-fluid cosmologies. Solving Eq. (3.2.29) shows that linear isocurvature modes always decay. In particular, we find $\Delta^{(\text{iso})} \propto t^{-1/2}$ during the radiation era and $\Delta^{(\text{iso})} \propto t^{-1}$ after equipartition.

3.2.5. Density vortices

Magnitude preserving changes in the distribution of the density gradient are monitored via the antisymmetric orthogonally projected tensor $\Delta_{[ab]}$. To first order, the latter is directly related to the vorticity tensor according to $\Delta_{[ab]} = -3a^2(1+w)H\omega_{ab}$, given that the orthogonally projected gradients of scalars do not commute in the presence of rotation (see Appendix A.3). Similarly we find that $\mathcal{Z}_{[ab]} = 3a^2\dot{H}\omega_{ab}$, which means that

$$\mathcal{Z}_{[ab]} = -\frac{\dot{H}}{(1+w)H} \Delta_{[ab]}, \quad (3.2.30)$$

at the linear perturbative level. Thus, linearising the skew part of the orthogonally projected gradient of (3.2.1), and then using the above and the background relations (1.4.3)(a) and (1.4.4) we obtain

$$\dot{\mathcal{W}}_a = -\frac{1}{2H} \left[(1-w)\rho - \frac{2(1-3w)K}{a^2} - 2\Lambda w \right] \mathcal{W}_a, \quad (3.2.31)$$

where $\mathcal{W}_a \equiv \varepsilon_{abc} \Delta^{bc}/2$. Accordingly, when $w < 1/3$, positive background 3-curvature tends to increase \mathcal{W}_a , while a negative curvature index will have the opposite effect. Also, for $w > 0$, the effect of a positive cosmological constant is to increase vortex-like density perturbations.

Introducing the vector harmonic functions $\mathcal{Q}_a^{(k)}$ (with $\dot{\mathcal{Q}}_a^{(k)} = 0 = D^a \mathcal{Q}_a^{(k)}$), we may write $\mathcal{W}_a = \Sigma_k \mathcal{W}^{(k)} \mathcal{Q}_a^{(k)}$ with $D_a \mathcal{W}^{(k)} = 0$. Then, the k -th harmonic mode evolves as

$$\dot{\mathcal{W}}^{(k)} = -\frac{1}{2H} \left[(1-w)\rho - \frac{2(1-3w)K}{a^2} - 2\Lambda w \right] \mathcal{W}^{(k)}. \quad (3.2.32)$$

Setting $K = 0 = \Lambda$ we find that $\mathcal{W} \propto t^{-1/2}$ during radiation and $\mathcal{W} \propto t^{-1}$ throughout dust. Therefore, linear vortices in the density distribution of a perturbed FLRW universe decay with time on all scales.

3.2.6. Dissipative effects

So far we have treated the cosmic medium as a perfect fluid, ignored dissipation and have established a physical scale, the Jeans length, that plays an important role in structure formation scenarios. There are other processes, however, which can modify the purely gravitational evolution of matter perturbations. For baryons, the most important interaction is their coupling to the pre-recombination photons. As the latter diffuse from high-density to low-density regions, they drag the baryons along with them, erasing inhomogeneities in their distribution. This process, which is known as Silk damping, can wipe out small-scale fluctuations in the baryonic component [145,146]. In particular, if ℓ_S is the scale associated with this effect, the corresponding Silk mass is given by (e.g. see [139])

$$M_S \propto \rho_b \ell_S^3 \simeq 6.2 \times 10^{12} \left(\frac{\Omega}{\Omega_b} \right)^{3/2} (\Omega h^2)^{-5/4} M_\odot. \quad (3.2.33)$$

The effect of photon free-streaming rises steeply as we approach recombination and then ceases. Thus, before recombination, subhorizon-sized baryonic perturbations on scales below ℓ_S are obliterated by Silk damping, those with masses between M_S and M_J oscillate like acoustic waves, while modes having $M > M_J$ can grow. The latter, however, are of little cosmological interest.

In an analogous way, the free geodesic motion of the collisionless (dark matter) species erases any structure that tries to form in their small-scale distribution. Therefore, the ideal fluid description of the dark component is a good approximation on sufficiently large scales only. The dissipative process is known as *free streaming* (or *Landau damping*) and its full study requires integrating the collisionless Boltzmann equation of the species in question (see Section 4). Nevertheless, one can obtain an estimate of the effect by calculating the maximum distance traveled by a free-streaming particle. Following [139, 142], hot thermal relics (see Section 3.3.5 for a brief discussion on dark matter candidates), have

$$\ell_{FS} \simeq 0.5 \left(\frac{m_{DM}}{1 \text{ keV}} \right)^{-4/3} \left(\Omega_{DM} h^2 \right)^{1/3} \text{ Mpc}, \quad (3.2.34)$$

where m_{DM} is the mass of the species in units of 1 keV. This means that the minimum scale to survive collisionless dissipation depends crucially on the particles' mass. Applied to neutrinos with $m_\nu \simeq 30 \text{ eV}$, the above gives $\ell_{FS} \simeq 28 \text{ Mpc}$ and a corresponding mass-scale of approximately $10^{15} M_\odot$. For much heavier candidate with, say, $m_{DM} \simeq 1 \text{ keV}$, free streaming will wipe out perturbations on scales below $\ell_{FS} \simeq 0.5 \text{ Mpc}$ (or smaller than $M_{FS} \simeq 10^9 M_\odot$). Overall, the lighter the dark matter is, the less power survives on small scales.

Cold thermal relics (CDM—see Section 3.3.5) have very small dispersion velocities and the values of their free-streaming masses are very low. As a result, perturbations in the dark-matter component grow unimpeded by damping processes on all scales of cosmological interest, although they suffer stagnation due to the Meszaros effect (see Section 3.3.4) until the time of matter-radiation equality. After recombination, the potential wells of the collisionless species serve to boost the growth of baryonic perturbations (see Section 3.3.5).

3.3. Multi-fluid perturbations

3.3.1. Linearised evolution equations for the i -th species

In the FLRW background all the members of the multi-component system are perfect fluids sharing the same 4-velocity u_a by construction. As a result, the effective flux terms, which depend on the peculiar velocities of the species vanish to zero order. This ensures that these quantities are gauge invariant first-order perturbations. Then, Eq. (2.4.18) linearises to

$$\begin{aligned} \dot{\Delta}_a^{(i)} &= 3w^{(i)}H\Delta_a^{(i)} - (1+w^{(i)})\mathcal{Z}_a + \frac{3aH}{\rho^{(i)}}(\dot{q}_{(a)}^{(i)} + 4Hq_a^{(i)}) - \frac{a}{\rho^{(i)}}D_a(D^b q_b^{(i)} - \mathcal{I}^{(i)}) \\ &\quad - \frac{1}{\rho^{(i)}}\mathcal{I}^{(i)}\Delta_a^{(i)} + \frac{a}{\rho^{(i)}}\mathcal{I}^{(i)}A_a - \frac{3aH}{\rho^{(i)}}\mathcal{I}_a^{(i)}, \end{aligned} \quad (3.3.1)$$

where $w^{(i)} = p^{(i)}/\rho^{(i)}$ and $q_a^{(i)} = \rho^{(i)}(1+w^{(i)})v_a^{(i)}$, with $v_a^{(i)}$ representing the peculiar velocity of each fluid relative to the u_a -frame. The associated momentum–density conservation law reads

$$a\rho^{(i)}(1+w^{(i)})A_a = -c_s^{2(i)}\rho^{(i)}\Delta_a^{(i)} - p^{(i)}\mathcal{E}_a^{(i)} + a(\dot{q}_{(a)}^{(i)} + 4Hq_a^{(i)}) + a\mathcal{I}_a^{(i)}, \quad (3.3.2)$$

with $c_s^{2(i)} = \dot{p}^{(i)}/\dot{\rho}^{(i)}$ to zero order. Using the above to eliminate the third term in the right-hand side of (3.3.1) and then employing the linear part of (2.4.13) we obtain

$$\begin{aligned} \dot{\Delta}_a^{(i)} &= 3(w^{(i)} - c_s^{2(i)})H\Delta_a^{(i)} - 3w^{(i)}H\mathcal{E}_a^{(i)} - (1+w^{(i)})\mathcal{Z}_a - \frac{a}{\rho^{(i)}}D_a(D^b q_b^{(i)} - \mathcal{I}^{(i)}) - \frac{1}{\rho^{(i)}}\mathcal{I}^{(i)}\Delta_a^{(i)} \\ &\quad + \frac{a}{\rho(1+w)} \left[3(1+w^{(i)})H - \frac{1}{\rho^{(i)}}\mathcal{I}^{(i)} \right] \left(\frac{c_s^2 \rho}{a} \Delta_a + \frac{p}{a} \mathcal{E}_a + \dot{q}_a + 4Hq_a \right), \end{aligned} \quad (3.3.3)$$

where in the FLRW background $w = (1/\rho) \sum_i \rho^{(i)} w^{(i)}$ and $c_s^2 = [1/\rho(1+w)] \sum_i c_s^{2(i)} \rho^{(i)} (1+w^{(i)})$. This is coupled to the linear propagation equation of the expansion gradient, which depends on the total fluid and obeys

$$\dot{\mathcal{Z}}_a = -2H\mathcal{Z}_a - \frac{1}{2}\rho\Delta_a - \frac{3}{2}c_s^2\rho\Delta_a - \frac{3}{2}p\mathcal{E}_a - a \left[3H^2 + \frac{1}{2}(\rho + 3p) - \Lambda \right] A_a + aD_a D^b A_b, \quad (3.3.4)$$

obtained from (2.3.2) by means of decomposition (2.4.12). Using the linearised part of Eq. (2.4.13), the linear commutation laws between the projected gradients of scalars and recalling that $\dot{s} = 0$ in an FLRW spacetime, the above transforms to

$$\begin{aligned} \dot{\mathcal{Z}}_a &= -2H\mathcal{Z}_a - \frac{1}{2}\rho\Delta_a - \frac{c_s^2}{1+w} \left(D^2 \Delta_a + \frac{K}{a^2} \Delta_a \right) - \frac{w}{1+w} \left(D^2 \mathcal{E}_a + \frac{K}{a^2} \mathcal{E}_a \right) \\ &\quad + \frac{3a}{\rho(1+w)} \left[\frac{1}{2}\rho(1+w) - \frac{K}{a^2} \right] (\dot{q}_a + 4Hq_a) - \frac{a}{\rho(1+w)} D_a D^b (\dot{q}_b + 4Hq_b) - 6ac_s^2 H D^b \omega_{ab}. \end{aligned} \quad (3.3.5)$$

Expressions (3.3.3) and (3.3.5) are the basic members of a system of equations that monitors the linear evolution of density inhomogeneities in an almost-FLRW universe filled with several interacting and non-comoving perfect fluids.

3.3.2. Entropy perturbations in a multi-fluid system

In a fluid mixture, inhomogeneities in the effective total energy density (ρ) are related to those in densities of the individual members by

$$\Delta_a = \frac{1}{\rho} \sum_i \rho^{(i)} \Delta_a^{(i)}. \quad (3.3.6)$$

Treating the whole multi-system as an effective single (total) fluid, we may combine (3.3.2) with the linear part of (2.4.13) and the above to obtain

$$p \mathcal{E}_a = \sum_i p^{(i)} \mathcal{E}_a^{(i)} + \sum_i c_s^{2(i)} \rho^{(i)} \Delta_a^{(i)} - c_s^2 \sum_i \rho^{(i)} \Delta_a^{(i)}, \quad (3.3.7)$$

which provides the total effective entropy perturbation of the multi-component system relative to the properties of its members. Note that in the FLRW background all species share the same 4-velocity and the effective total sound speed satisfies the relation

$$c_s^2 = \frac{1}{\rho(1+w)} \sum_i c_s^{2(i)} [\rho^{(i)} (1+w^{(i)})], \quad (3.3.8)$$

to zero order. Using the above we may recast expression (3.3.7) into the following linear expression

$$\mathcal{E}_a = \frac{1}{p} \sum_i p^{(i)} \mathcal{E}_a^{(i)} + \frac{1}{2ph} \sum_{i,j} h^{(i)} h^{(j)} (c_s^{2(i)} - c_s^{2(j)}) \mathcal{J}_a^{(ij)}, \quad (3.3.9)$$

where $h = \rho + p$, $h^{(i)} = \rho^{(i)} + p^{(i)}$ and

$$\mathcal{J}_a^{(ij)} = \frac{\rho^{(i)}}{h^{(i)}} \Delta_a^{(i)} - \frac{\rho^{(j)}}{h^{(j)}} \Delta_a^{(j)} \quad (3.3.10)$$

with $\mathcal{J}_a^{(ij)} = -\mathcal{J}_a^{(ji)}$. Accordingly, in a multi-component fluid the total effective entropy perturbation has one part coming from the intrinsic entropy perturbations of the individual species and another due to their different dynamical behaviour. Following (3.3.9), the latter vanishes if the species share the same sound speed or when $\mathcal{J}_a^{(ij)} = 0$.

3.3.3. Density perturbations in the i -th species

The scalar $\Delta^{(i)} = aD^a \Delta_a^{(i)}$ describes overdensities (or underdensities) in the matter distribution of the i -th fluid. Thus, assuming no interactions and taking the comoving projected divergence of (3.3.3) we obtain

$$\begin{aligned} \dot{\Delta}^{(i)} &= 3(w^{(i)} - c_s^{2(i)}) H \Delta^{(i)} - 3w^{(i)} H \mathcal{E}^{(i)} - (1+w^{(i)}) \mathcal{Z} - \frac{a^2}{\rho^{(i)}} D^2 (D^a q_a^{(i)}) \\ &\quad + \frac{3(1+w^{(i)})H}{\rho(1+w)} (c_s^2 \rho \Delta + p \mathcal{E}) + \frac{3a^2(1+w^{(i)})H}{\rho(1+w)} D^a (\dot{q}_a + 4Hq_a), \end{aligned} \quad (3.3.11)$$

where $\mathcal{E}^{(i)} = aD^a \mathcal{E}_a^{(i)}$ and $\mathcal{E} = aD^a \mathcal{E}_a$. In an analogous way Eq. (3.3.5) leads to the linear expression

$$\begin{aligned} \dot{\mathcal{Z}} &= -2HZ - \frac{1}{2} \rho \Delta - \frac{c_s^2}{1+w} (D^2 \Delta + \frac{3K}{a^2} \Delta) - \frac{w}{1+w} (D^2 \mathcal{E} + \frac{3K}{a^2} \mathcal{E}) \\ &\quad - \frac{a^2}{\rho(1+w)} \left[D^2 D^a (\dot{q}_a + 4Hq_a) + \frac{3K}{a^2} D^a (\dot{q}_a + 4Hq_a) \right] + \frac{3}{2} a^2 D^a (\dot{q}_a + 4Hq_a), \end{aligned} \quad (3.3.12)$$

since $D^a D^b \omega_{ab} = 0$ to first approximation. Also, in deriving the above we have used the first-order auxiliary result $aD^a D^2 \Delta_a = D^2 \Delta + (2K/a^2) \Delta$ and an exactly analogous relation between the Laplacians of \mathcal{E}_a and \mathcal{E} .

Expressions (3.3.11) and (3.3.12) govern the linear evolution of matter aggregations in the density of the i -th species provided the total flux vector $q_a = \sum_i q_a^{(i)} = \sum_i \rho^{(i)} (1+w^{(i)}) v_a^{(i)}$ is specified. Using the energy frame (i.e. setting $q_a = \sum_i q_a^{(i)} = 0$) the system (3.3.11) and (3.3.12) reduces to

$$\dot{\Delta}^{(i)} = 3(w^{(i)} - c_s^{2(i)}) H \Delta^{(i)} - 3w^{(i)} H \mathcal{E}^{(i)} - (1+w^{(i)}) \mathcal{Z} - a(1+w^{(i)}) D^2 v^{(i)} + \frac{3(1+w^{(i)})H}{1+w} (c_s^2 \Delta + w \mathcal{E}) \quad (3.3.13)$$

and

$$\dot{\mathcal{Z}} = -2HZ - \frac{1}{2} \rho \Delta - \frac{c_s^2}{1+w} (D^2 \Delta + \frac{3K}{a^2} \Delta) - \frac{w}{1+w} (D^2 \mathcal{E} + \frac{3K}{a^2} \mathcal{E}), \quad (3.3.14)$$

respectively. The peculiar motion of the species, relative to the u_a -frame, is described by the scalar $v^{(i)} = aD^a v_a^{(i)}$ (not to be confused with the magnitude of the peculiar velocity field – see Section 2.4.1). To linear order, the evolution of the latter is obtained by combining the conservation laws for the momentum density of the i -th component (i.e. Eq. (3.3.2) with

$\mathcal{I}_a^{(i)} = 0$) and that of the effective total fluid (the linear part of (2.4.13)). In particular, recalling that $q_a^{(i)} = \rho^{(i)}(1 + w^{(i)})v_a^{(i)}$, a straightforward algebraic calculation leads to

$$\dot{v}^{(i)} = -\left(1 - 3c_s^{2(i)}\right)Hv^{(i)} - \frac{1}{a(1 + w^{(i)})} \left(c_s^{2(i)}\Delta^{(i)} + w^{(i)}\mathcal{E}^{(i)}\right) - \frac{1}{a(1 + w)} \left(c_s^2\Delta + w\mathcal{E}\right). \quad (3.3.15)$$

3.3.4. A radiation and dust universe

Consider a spatially flat FLRW background containing a mixture of radiation and dust. This is a good approximate description of the pre-decoupling universe, if the radiation is identified with the photons and the neutrinos, $\rho^{(r)} = \rho^{(\gamma)} + \rho^{(\nu)}$, while the pressureless component accounts for the nonrelativistic species, baryonic and cold dark matter, $\rho^{(d)} = \rho^{(b)} + \rho^{(c)}$. Neglecting the photon–baryon interaction terms we have $\rho^{(r)} = \rho_0^{(r)}(a_0/a)^4$ and $\rho^{(d)} = \rho_0^{(d)}(a_0/a)^3$, with the zero suffix indicating a chosen initial time. The total energy of the mixture is $\rho = \rho^{(r)} + \rho^{(d)}$, the total pressure is $p = \rho^{(r)}/3$, the effective total sound speed is $c_s^2 = 4\rho^{(r)}/[3(4\rho^{(r)} + 3\rho^{(d)})]$ and the unperturbed background satisfies the condition $3H^2 = \rho$.

According to expressions (3.3.9) and (3.3.10) in Section 3.3.2, the total effective entropy perturbation of the above described radiation–dust mixture is

$$\mathcal{E}_a = -\frac{4\rho^{(d)}}{3\rho^{(d)} + 4\rho^{(r)}} \left(\Delta_a^{(d)} - \frac{3}{4}\Delta_a^{(r)}\right). \quad (3.3.16)$$

Consequently, imposing the condition of zero effective entropy perturbation corresponds to setting $\Delta_a^{(d)} = 3\Delta_a^{(r)}/4$ and vice-versa. An additional special case emerges when neither $\Delta_a^{(d)}$ nor $\Delta_a^{(r)}$ vanishes, but the total energy density is homogeneously distributed (i.e. $\Delta_a = 0$). Then, Eq. (3.3.6) guarantees that $\rho^{(d)}\Delta_a^{(d)} = -\rho^{(r)}\Delta_a^{(r)}$.

In what follows we will assume that the radiation field has a homogeneous density distribution (i.e. for $\Delta^{(r)} = 0$), which holds, for example, inside the sound horizon after averaging over acoustic oscillations, or on scales that are damped by diffusion. We may then consider aggregations in the dust component only. Following (3.3.13)–(3.3.15) these are monitored by the system

$$\dot{\Delta}^{(d)} = -\mathcal{Z} - aD^2v^{(d)}, \quad \dot{\mathcal{Z}} = -2H\mathcal{Z} - \frac{1}{2}\rho^{(d)}\Delta^{(d)} \quad (3.3.17)$$

and

$$\dot{v}^{(d)} = -Hv^{(d)}, \quad (3.3.18)$$

given that $c_s^2\Delta + w\mathcal{E} = 0$ to first order [117]. The latter is verified by first applying (3.3.9) to our two-component system, which shows that $\mathcal{E} = -(\rho^{(d)}h^{(r)}/3ph)\Delta^{(d)}$. Then, since $\Delta = (\rho^{(d)}/\rho)\Delta^{(d)}$ and using the earlier given expression for the effective sound speed of the total fluid, one can easily ensure that $w\mathcal{E} = -c_s^2\Delta$.

The system (3.3.17) and (3.3.18) monitors the linear evolution of overdensities in the matter distribution of the dust component within a perturbed spatially flat FLRW universe filled with a mixture of radiation and dust, where the radiative component is homogeneously distributed. Taking the time derivative of (3.3.17)(a), and using the linear commutation law $(D^2v^{(d)})^\cdot = D^2\dot{v}^{(d)} - 2HD^2v^{(d)}$, the above system reduces to the scale independent equation [147]

$$\Delta''^{(d)} = -\frac{2 + 3a}{2a(1 + a)}\Delta'^{(d)} + \frac{3}{2a(1 + a)}\Delta^{(d)}. \quad (3.3.19)$$

Here primes indicate differentiation with respect to the scale factor and we have normalised the parameters so that $a = 1$ at the time of matter–radiation equality [117]. The general solution has the form [148]

$$\Delta^{(d)} = \mathcal{C}_1 \left(1 + \frac{3}{2}a\right) - \mathcal{C}_2 \left[\left(1 + \frac{3}{2}a\right) \ln \left(\frac{\sqrt{1+a} + 1}{\sqrt{1+a} - 1} \right) - 3\sqrt{1+a} \right], \quad (3.3.20)$$

which shows that $\Delta^{(d)}$ grows proportionally to the scale factor at late times in agreement with a single-fluid Einstein–de Sitter model. Deep into the radiation era on the other hand, $a \ll 1$ and the density contrast is effectively constant (it grows only logarithmically). This stagnation, or freezing-in, of matter perturbations prior to equality is generic to models with a period of expansion that is dominated by relativistic particles and it is sometimes referred to as the Meszaros effect [147]. Finally, we note that the Newtonian analysis also leads to the same results (e.g. see [139,141,142]).

3.3.5. A CDM and baryon universe

It has long been known that purely baryonic scenarios cannot explain the structure observed in the universe today. The main reason is the tight coupling between the photons and baryons in the pre-recombination era, which washes out perturbations in the baryonic component. Dark matter, on the other hand, interacts only gravitationally and therefore it is not subjected to the photon drag. High energy physics provides a long catalogue of dark matter candidates. Thermal relics, namely those kept in thermal equilibrium with the rest of the universe until the time they decouple, typically classify as Hot

Dark Matter (HDM) and Cold Dark Matter (CDM) species.¹⁶ CDM has small dispersion velocities by the time of decoupling and does not suffer free-streaming dissipation.

Following Section 3.3.4, perturbations in the CDM component grow between equipartition and recombination by a factor of $a_{\text{rec}}/a_{\text{eq}} = T_{\text{eq}}/T_{\text{rec}} \simeq 21 \Omega h^2$, where $h = H/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is the rescaled, dimensionless Hubble parameter. After decoupling the universe becomes effectively transparent to radiation and perturbations in the ordinary matter can start growing, driven by the gravitational potential of the collisionless species. Applied to a mixture of CDM and non-relativistic baryons, the system (3.3.13)–(3.3.15) leads to

$$\dot{\Delta}^{(b)} = -Z - aD^2v^{(b)}, \quad \dot{Z} = -2HZ - \frac{1}{2}\rho\Delta \quad (3.3.21)$$

and

$$\dot{v}^{(b)} = -Hv^{(b)}, \quad (3.3.22)$$

where now $\rho = \rho^{(c)} + \rho^{(b)}$ and $\rho\Delta = \rho^{(c)}\Delta^{(c)} + \rho^{(b)}\Delta^{(b)}$. Proceeding as in the previous section, the time derivative of (3.3.21)(a) leads to

$$\ddot{\Delta}^{(b)} + 2H\dot{\Delta}^{(b)} = \frac{1}{2}\rho^{(c)}\Delta^{(c)}. \quad (3.3.23)$$

Note that the dark component dominates the baryonic one (i.e. $\rho^{(b)} \ll \rho^{(c)}$) and, just after recombination, $\Delta^{(b)} \ll \Delta^{(c)}$. Recalling that $\Delta^{(c)} \propto a$ after recombination (see Section 3.3.4) and also that $\rho^{(c)} \propto a^{-3}$, we find [139,142]

$$\Delta^{(b)} = \Delta^{(c)} \left(1 - \frac{a_{\text{rec}}}{a}\right), \quad (3.3.24)$$

which shows that $\Delta^{(b)} \rightarrow \Delta^{(c)}$ when $a \gg a_{\text{rec}}$. In other words, after decoupling, baryonic fluctuations fall in the potential wells of the collisionless species and quickly catch up with perturbations in the dark-matter component. This result demonstrates how the presence of the non-baryonic species accelerates the gravitational collapse of ordinary matter and therefore the onset of structure formation. Moreover, because the dark-matter perturbation dominates the baryonic one, we expect to see the baryonic fluctuations manifested as a small acoustic peaks in the large-scale correlation function of galaxy surveys. Recent observations seem to confirm this, thus giving further support to dark-matter and to the CDM structure formation scenarios [149,150].

Cold relics have been proposed since the early 1980's in order to reproduce the small-scale structure of the universe (see [151–155] and references therein). Although purely CDM models do not seem to agree with observation, CDM and dark-energy, the latter as an effective cosmological constant, (i.e. Λ CDM scenarios) appear in very good agreement with the current data [156].

3.4. Magnetised perturbations

3.4.1. The Alfvén speed

We assume a weakly magnetised, spatially flat FLRW background containing a sufficiently random magnetic field. This means that $\langle B_a \rangle = 0$, while $\langle B^2 \rangle \neq 0$ and $\langle B^2 \rangle/\rho \ll 1$ on all scales of interest.¹⁷ Therefore, the energy density of the background B -field has only a time dependence (i.e. $\langle B^2 \rangle = B^2(t)$) and decays adiabatically as

$$B^2 = -4HB^2 \Rightarrow B^2 \propto a^{-4}, \quad (3.4.1)$$

according to expression (2.5.7). The relative strength of the field is measured by the dimensionless ratio $\beta = B^2/\rho$, which is used to define the Alfvén speed

$$c_a^2 = \frac{\beta}{1 + w + \beta}. \quad (3.4.2)$$

Provided $w \neq -1$, the above definition satisfies the constraint $c_a^2 < 1$ always and for a weak magnetic field, with $\beta \ll 1$, reduces to $c_a^2 = \beta/(1 + w)$.

¹⁶ Non-thermal relics, such as axions, magnetic monopoles and cosmic strings, remain out of equilibrium throughout their lifetime.

¹⁷ Even a random vector field will introduce a degree of anisotropy to the FLRW background. Nevertheless, it sounds plausible that a sufficiently weak B -field can be adequately studied within almost-FLRW models. This has been verified by studies of perturbed magnetised Bianchi I universes [128].

3.4.2. Magnetised density perturbations

Taking the comoving, orthogonally projected divergence of Eq. (2.5.12), using expression (2.5.10) and linearising we arrive at the following equation [66]

$$\dot{\Delta} = 3w \left(1 - \frac{2}{3} c_a^2 \right) H \Delta - (1+w) \mathcal{Z} + \frac{3}{2} c_a^2 (1+w) H \mathcal{B} - c_a^2 (1+w) H \mathcal{K}, \quad (3.4.3)$$

for the linear evolution of Δ . Note that $\mathcal{K} = a^2 \mathcal{R}$ is the rescaled 3-Ricci tensor of the perturbed spacetime. Also, in deriving the above we have kept up to c_a^2 -order terms, given the weakness of the magnetic field. Finally, we have assumed that $\dot{w} = 0$ to zero order. This means that $w = c_s^2 = \text{constant}$ in the background, as it happens during the radiation and dust eras for example. Following (3.4.3), the field will generally act as a source of density perturbations even when there are no such distortions present initially. Also, the magnetic field's presence has a direct and an indirect effect on Δ . The former results from the pressure part of the Lorentz force (see decomposition (2.5.11) in Section 2.5.2) and carries the effects of the isotropic magnetic pressure. The latter comes from the tension component of the Lorentz force and is triggered by the magnetic coupling to the spatial curvature of the perturbed model. Surprisingly, a positive 3-curvature perturbation causes Δ to decrease, while a negative \mathcal{K} has the opposite effect. This rather counter-intuitive behaviour of the magneto-curvature term in (3.4.3) is a direct consequence of the elasticity of the field lines (see also Eq. (3.4.4)).

Similarly, the linearised, orthogonally-projected, comoving divergences of (2.5.13) and (2.5.14) lead to [66]

$$\begin{aligned} \dot{\mathcal{Z}} = & -2 \left(1 + \frac{2}{3} c_a^2 \right) H \mathcal{Z} - \frac{1}{2} \left(1 - \frac{4}{3} c_a^2 \right) \rho \Delta + \frac{1}{4} c_a^2 (1+w) \rho \mathcal{B} - \frac{1}{2} c_a^2 (1+w) \rho \mathcal{K} \\ & - \frac{c_s^2}{1+w} \left(1 - \frac{2}{3} c_a^2 \right) D^2 \Delta - \frac{1}{2} c_a^2 D^2 \mathcal{B} \end{aligned} \quad (3.4.4)$$

and

$$\dot{\mathcal{B}} = \frac{4}{3(1+w)} \dot{\Delta}, \quad (3.4.5)$$

respectively. The latter is a key linear result, ensuring that perturbations in the magnetic energy density evolve in step with those in the density of the matter [157,158]. Finally, starting from the linear propagation equation of the 3-Ricci scalar we obtain [66]

$$\dot{\mathcal{K}} = -\frac{4}{3} c_a^2 H \mathcal{K} + \frac{4c_s^2}{(1+w)} \left(1 - \frac{2}{3} c_a^2 \right) H \Delta + 2c_a^2 H \mathcal{B}. \quad (3.4.6)$$

The system (3.4.3)–(3.4.6) monitors the linear evolution of scalar inhomogeneities in the density distribution of the matter in a weakly magnetised, spatially flat almost-FLRW universe.

3.4.3. Radiation era

During the radiation epoch the background dynamics are determined by the parameters $w = 1/3 = c_s^2$, $H = 1/(2t)$ and $\rho = 3/(4t^2)$. Also, the weakness of the magnetic field means that $c_a^2 = 3\beta/4$, where $\beta = B^2/\rho = \text{constant} \ll 1$. At the same time, the expansion scale factor is $a \propto t^{1/2}$, which means that the Hubble radius at comoving proper time t is given by $\lambda_H \equiv 1/H = 2t$. Then, recalling that physical wavelengths and comoving wavenumbers are related by $\lambda_k = a/k$, the harmonically decomposed system (3.4.3)–(3.4.6) reads

$$\dot{\Delta}^{(k)} = \frac{1}{2} \left(1 - \frac{1}{2} \beta \right) t^{-1} \Delta^{(k)} - \frac{4}{3} \mathcal{Z}^{(k)} - \frac{1}{2} \beta t^{-1} \mathcal{K}^{(k)} + \frac{3}{4} \beta t^{-1} \mathcal{B}^{(k)}, \quad (3.4.7)$$

$$\begin{aligned} \dot{\mathcal{Z}}^{(k)} = & - \left(1 + \frac{1}{2} \beta \right) t^{-1} \mathcal{Z}^{(k)} - \frac{3}{8} (1-\beta) t^{-2} \left[1 - \frac{1}{6} \left(\frac{k}{aH} \right)^2 \left(1 + \frac{1}{2} \beta \right) \right] \Delta^{(k)} \\ & - \frac{3}{8} \beta t^{-2} \mathcal{K}^{(k)} + \frac{3}{16} \beta t^{-2} \left[1 + \frac{1}{2} \left(\frac{k}{aH} \right)^2 \right] \mathcal{B}^{(k)}, \end{aligned} \quad (3.4.8)$$

$$\dot{\mathcal{K}}^{(k)} = -\frac{1}{2} \beta t^{-1} \mathcal{K}^{(k)} + \frac{1}{2} \left(1 - \frac{1}{2} \beta \right) t^{-1} \Delta^{(k)} + \frac{3}{4} \beta t^{-1} \mathcal{B}^{(k)}, \quad (3.4.9)$$

$$\dot{\mathcal{B}}^{(k)} = \dot{\Delta}^{(k)}. \quad (3.4.10)$$

On super-Hubble scales $k \ll aH$. Then, since $\beta \ll 1$ always, the system (3.4.7)–(3.4.10) is essentially scale-independent and has the power-law solution [66]

$$\Delta = \mathcal{C}_0 + \mathcal{C}_1 t^{1-\frac{1}{3}\beta} + \mathcal{C}_2 t^{-\frac{1}{2}+\frac{3}{8}\beta} + \mathcal{C}_3 t^{-\beta}, \quad (3.4.11)$$

where the \mathcal{C}_i s are constants that depend on the initial conditions. We note that in the absence of the magnetic field (i.e. for $\beta = 0$), we recover the standard evolution of the density contrast in perturbed non-magnetised FLRW models (compare

to solution (3.2.22) in Section 3.2.3). Thus, in the weak-field limit the main magnetic effect is to reduce the growth rate of the dominant density mode. In addition, the field also decreases the rate of the standard decay mode and introduces a new ‘non-adiabatic’ decay mode.

Well below the Hubble radius, $k \gg aH$, the scale-dependent terms inside the brackets of Eq. (3.4.8) become important. Thus, on sub-Hubble lengths the system (3.4.7)–(3.4.10) has the oscillatory solution [66]

$$\Delta^{(k)} = \mathcal{C}_0 + \mathcal{C}_1 \sqrt{\frac{2}{\pi\alpha}} \sin(\alpha t^{1/2}) + \mathcal{C}_2 \sqrt{\frac{2}{\pi\alpha}} \cos(\alpha t^{1/2}), \quad (3.4.12)$$

with $\alpha = c_s(k/a_0H_0)t_0^{-1/2}(1+\beta/2)$. Therefore, well inside the horizon magnetised matter aggregations oscillate like magneto-sonic waves. The field’s presence tends to reduce the amplitude of the oscillation and increase its frequency [157,158]. In both cases the effect of the B -field is proportional to the ratio $\beta = B^2/\rho$ and its relative strength. As pointed out in [159], the increased frequency of $\Delta^{(k)}$ should bring the peaks of short-wavelength oscillations in the density of the radiation component closer. This in turn could produce a potentially observable signature in the CMB. An additional magnetic effect arises from the presence of a constant mode in solution (3.4.12). The latter suggests that, unlike the magnetic-free case (e.g. see [139]), magnetic matter aggregations in the pre-equilibrium universe oscillate around a generally non-zero average value.

3.4.4. Dust era

After the end of the radiation era, the unperturbed background is well approximated by $w = 0 = c_s^2$, $H = 2/(3t)$, $\rho = 4/(3t^2)$. Also, $c_s^2 = \beta$ and it is no longer constant but decreases with time according to $\beta \propto a^{-1} \propto t^{-2/3}$. Applying the usual harmonic decomposition to the perturbation variables, and keeping up to β -order terms, the system (3.4.3)–(3.4.6) reads

$$\dot{\Delta}^{(k)} = -\mathcal{Z}^{(k)} - \frac{2}{3}\beta t^{-1}\mathcal{K}^{(k)} + \beta t^{-1}\mathcal{B}^{(k)}, \quad (3.4.13)$$

$$\dot{\mathcal{Z}}^{(k)} = -\frac{4}{3}\left(1 + \frac{2}{3}\beta\right)t^{-1}\mathcal{Z}^{(k)} - \frac{2}{3}\left(1 - \frac{4}{3}\beta\right)t^{-2}\Delta^{(k)} - \frac{2}{3}\beta t^{-2}\mathcal{K}^{(k)} + \frac{1}{3}\beta\left[1 + \frac{2}{3}\left(\frac{k}{aH}\right)^2\right]t^{-2}\mathcal{B}^{(k)}, \quad (3.4.14)$$

$$\dot{\mathcal{K}}^{(k)} = -\frac{8}{9}\beta t^{-1}\mathcal{K}^{(k)} + \frac{4}{3}\beta t^{-1}\mathcal{B}^{(k)}, \quad (3.4.15)$$

$$\dot{\mathcal{B}}^{(k)} = \frac{4}{3}\dot{\Delta}^{(k)}. \quad (3.4.16)$$

After equipartition the dimensionless parameter β is no longer constant but decays in time. Without the weak and decaying terms, the above has the following late-time solution [158,66]

$$\Delta^{(k)} = \mathcal{C}_+ t^{\alpha_+} + \mathcal{C}_- t^{\alpha_-} + \mathcal{C}_3 t^{-2/3} + \mathcal{C}_4, \quad (3.4.17)$$

with

$$\alpha_{\pm} = -\frac{1}{6}\left[1 \pm 5\sqrt{1 - \frac{32}{75}\beta_0\left(\frac{k}{a_0H_0}\right)^2}\right]. \quad (3.4.18)$$

In the absence of the B -field (i.e. when $\beta_0 = 0$), we immediately recover the standard non-magnetic solution with $\alpha_+ = 2/3$ and $\alpha_- = -1$ (see solution (3.2.24) in Section 3.2.3). Therefore, the main magnetic effect is to reduce the growth rate of density perturbations by an amount proportional to its relative strength (i.e. to the ratio $\beta_0 = (B^2/\rho)_0$). It should be noted that the inhibiting role of the field was first observed in the Newtonian treatment of [160] and later in the relativistic studies of [157,158]. According to (3.4.17) and (3.4.18), the magnetic impact is inversely proportional to the scale in question. Hence, on large scales the introduction of the B -field simply adds the decaying $t^{-2/3}$ mode to the standard magnetic-free result. Note also that Δ describes the directionally averaged gravitational clumping of the matter. Generally, the perturbations will grow at different rates parallel and perpendicular to the magnetic field and so there will also be non-spherical evolution in the shapes of these distortions.

After matter-radiation equality the magnetic field is essentially the sole source of pressure support. The associated magnetic Jeans length is obtained by means of the following wave-like equation

$$\ddot{\Delta}^{(k)} = -\frac{4}{3}\left(1 - \frac{1}{3}\beta\right)t^{-1}\dot{\Delta}^{(k)} + \frac{2}{3}\left\{1 - \frac{8}{3}\beta\left[1 + \frac{1}{6}\left(\frac{k}{aH}\right)^2\right]\right\}t^{-2}\Delta^{(k)}, \quad (3.4.19)$$

obtained by taking the time derivative of (3.4.13), ignoring the 3-curvature effects and setting $\mathcal{B}_0^{(k)} = 4\Delta_0^{(k)}/3$. The latter condition guarantees that $\mathcal{B}^{(k)} = 4\Delta^{(k)}/3$ always (see Eq. (3.4.16)), while dropping the magneto-curvature terms from (3.4.13) means that only the effects of the isotropic magnetic pressure are accounted for. When the factor in braces in the

last term of Eq. (3.4.19) is positive, gravity prevails and the density contrast grows. This happens on scales larger than the magnetic Jeans length

$$\lambda_{J_B} = \sqrt{\frac{4\beta}{9 - 24\beta}} \lambda_H. \quad (3.4.20)$$

The latter is considerably smaller than the corresponding Hubble radius since $\beta \ll 1$, but for B -fields with current strengths of the order of 10^{-7} G is intriguingly close to the size of a galaxy cluster (i.e. $\lambda_{J_B} \sim 1$ Mpc—see [66]).

3.5. Scalar-field perturbations

3.5.1. The effective fluid characteristics

In line with the fluid description of scalar fields, we may introduce the familiar dimensionless w -parameter. Thus, based on Eq. (2.6.10) we define

$$w_\varphi = \frac{p^{(\varphi)}}{\rho^{(\varphi)}} = \frac{\dot{\varphi}^2 - 2V(\varphi)}{\dot{\varphi}^2 + 2V(\varphi)}. \quad (3.5.1)$$

Then, provided that $\rho^{(\varphi)} > 0$, we have $\dot{\varphi}^2/2 + V(\varphi) > 0$ and the familiar condition $-1 \leq w_\varphi \leq 1$ follows from $V(\varphi) \geq 0$. Taking the time derivative and using the Klein–Gordon equation, as seen in expression (2.6.12) we arrive at

$$\dot{w}_\varphi = -3H(1 + w_\varphi) \left(\frac{\dot{p}^{(\varphi)}}{\dot{\rho}^{(\varphi)}} - w_\varphi \right). \quad (3.5.2)$$

The above agrees with the evolution law of the w -parameter in conventional perfect-fluid cosmologies (see Eq. (3.2.2) in Section 3.2.1), although in the scalar-field case the ratio

$$\frac{\dot{p}^{(\varphi)}}{\dot{\rho}^{(\varphi)}} = 1 + \frac{2V'(\varphi)}{3H\dot{\varphi}}, \quad (3.5.3)$$

no longer represents an associated thermodynamic sound speed. According to expression (3.5.2), for $H \neq 0$ and $w_\varphi \neq -1$, the $w^{(\varphi)}$ -parameter is time-invariant when $\dot{p}^{(\varphi)}/\dot{\rho}^{(\varphi)} = w^{(\varphi)}$. In that case, relations (3.5.1) and (3.5.3) combine to give the evolution law $\dot{V}(\varphi) = -3H(1 + w^{(\varphi)})V(\varphi)$ for the scalar-field potential. In the latter case it is straightforward to show that $\dot{p}^{(\varphi)}/\dot{\rho}^{(\varphi)} = w^{(\varphi)} = 1 - 2V(\varphi)/\rho^{(\varphi)}$ is constant. As expected, $\dot{p}^{(\varphi)}/\dot{\rho}^{(\varphi)} = w^{(\varphi)} \rightarrow 1$ when $V(\varphi) \rightarrow 0$ and $\dot{p}^{(\varphi)}/\dot{\rho}^{(\varphi)} = w^{(\varphi)} \rightarrow -1$ for $\rho^{(\varphi)} \rightarrow V(\varphi)$ (i.e. for $\dot{\varphi} \rightarrow 0$).

3.5.2. Density perturbations

Aggregations in the effective energy density of a minimally coupled scalar field are described by the comoving divergence $\Delta^{(\varphi)} = aD^a \Delta_a^{(\varphi)}$, which to linear order is given by $\Delta^{(\varphi)} = (a^2/\rho^{(\varphi)})D^2\rho^{(\varphi)} = (a^2\dot{\varphi}/\rho^{(\varphi)})D^2\dot{\varphi}$. At this perturbative level, the comoving divergences of (2.6.16) and (2.6.17) read

$$\dot{\Delta}^{(\varphi)} = 3w^{(\varphi)}H\Delta^{(\varphi)} - (1 + w^{(\varphi)})\mathcal{Z}, \quad (3.5.4)$$

and

$$\dot{\mathcal{Z}} = -2H\mathcal{Z} - \left[\frac{1}{2}\rho^{(\varphi)} + \frac{3K}{a^2(1 + w_\varphi)} \right] \Delta^{(\varphi)} - \frac{1}{1 + w_\varphi} D^2\Delta^{(\varphi)}, \quad (3.5.5)$$

respectively. Note that in deriving (3.5.5) and in order to express the 4-acceleration and its divergence with respect to the density gradients, we have used expression (2.6.8). The latter ensures that $D^a A_a = -[1/a^2(1 + w_\varphi)]\Delta$ to first approximation. According to Eq. (3.5.4), Δ decouples from \mathcal{Z} when $w^{(\varphi)} \rightarrow -1$ (see Section 3.5.3). For $w^{(\varphi)} \neq -1$, on the other hand, the time derivative of (3.5.4) gives the following wavelike equation for the k -th harmonic mode [135]

$$\ddot{\Delta}_{(k)}^{(\varphi)} = -2 \left(1 - 3w_\varphi + \frac{3\dot{p}^{(\varphi)}}{2\dot{\rho}^{(\varphi)}} \right) H\dot{\Delta}_{(k)}^{(\varphi)} + \frac{3}{2} \left(1 + 8w_\varphi - 3w_\varphi^2 - 6\frac{\dot{p}^{(\varphi)}}{\dot{\rho}^{(\varphi)}} \right) H^2\Delta_{(k)}^{(\varphi)} + \frac{1}{a^2} \left[\frac{9}{2}(1 - w_\varphi^2)K - k^2 \right] \Delta_{(k)}^{(\varphi)}, \quad (3.5.6)$$

where $K = 0, \pm 1$. The above governs the evolution of linear inhomogeneities in the density distribution of a minimally coupled scalar field in a perturbed FLRW universe. It becomes clear that to a very large extent this evolution is determined by the effective equation of state of the φ -field, which in turn depends on the latter's kinetic energy and potential.

3.5.3. Standard slow-roll inflation

Standard inflation corresponds to approximately exponential de Sitter expansion, with H and $\rho^{(\varphi)}$ nearly constants. This is achieved when the scalar field is rolling very slowly down its potential. For $\dot{\varphi}^2 \ll V(\varphi)$ and $|\dot{\varphi}| \ll H|\varphi|$, we have $w^{(\varphi)} = p^{(\varphi)}/\rho^{(\varphi)} \rightarrow -1$ (see Section 2.6.3). This means that as we approach the de Sitter regime, Eq. (3.5.6) no longer depends on the background spatial curvature and the linear evolution of $\Delta_{(k)}^{(\varphi)}$ is governed by

$$\ddot{\Delta}_{(k)}^{(\varphi)} = -5H\dot{\Delta}_{(k)}^{(\varphi)} - 6H^2 \left[1 + \frac{1}{6} \left(\frac{k}{aH} \right)^2 \right] \Delta_{(k)}^{(\varphi)}, \quad (3.5.7)$$

where $H \simeq \text{constant}$. Thus, once the mode has crossed the Hubble radius and $k \ll aH$, the solution is (e.g. see [139])

$$\Delta^{(\varphi)} = \mathcal{C}_1 e^{-2Ht} + \mathcal{C}_2 e^{-3Ht}. \quad (3.5.8)$$

Following Section 1.4.4, this means that during a phase of exponential, de Sitter-type expansion $\Delta^{(\varphi)} \propto a^{-2}$ (e.g. see [139]). Therefore, any overdensities (or underdensities) that may exist in the spatial distribution of the inflaton field will decay exponentially irrespective of their scale and the background curvature.¹⁸

We note that the above analysis applies to a slowly rolling scalar field, with $w^{(\varphi)} \simeq -1$. When $w^{(\varphi)} = -1$, on the other hand, the φ -field has zero kinetic energy and $p^{(\varphi)} = -\rho^{(\varphi)} = -V(\varphi)$. In that case it is straightforward to show that $D_a p^{(\varphi)} = -D_a \rho^{(\varphi)} = -V'(\varphi)D_a \varphi = 0$, which means that there are no inhomogeneities in the effective energy density and pressure of the field. This result is a direct consequence of our spacetime slicing, which guarantees that $D_a \varphi = 0$.

3.5.4. Coasting universe

When $w^{(\varphi)} = -1/3$ the effective gravitational mass of the φ -field vanishes. This leads to a ‘coasting’ FLRW universe with $a \propto t$ and $H = 1/t$. During this phase of ‘minimal’ inflation Eq. (3.5.6) reduces to

$$\ddot{\Delta}_{(k)}^{(\varphi)} = -3H\dot{\Delta}_{(k)}^{(\varphi)} + \frac{1}{a^2} (4K - k^2) \Delta_{(k)}^{(\varphi)} \quad (3.5.10)$$

and has the solution

$$\Delta_{(k)}^{(\varphi)} = \mathcal{C}_+ t^{\alpha_+} + \mathcal{C}_- t^{\alpha_-}, \quad (3.5.11)$$

with

$$\alpha_{\pm} = -1 \pm \frac{1}{a_0} \sqrt{a_0^2 + (4K - k^2)t_0^2} \quad (3.5.12)$$

and $a_0 = a(t_0)$. Recall that when dealing with a coasting FLRW universe we have $\Omega = \Omega_0$ constant and $a_0/t_0 = \sqrt{K/(\Omega_0 - 1)}$ (see Section 1.4.4). The nature of the above given solution is determined by the sign of the sum $a_0^2 + (4K - k^2)t_0^2$, which in turn depends on the background 3-curvature and the scale of the inhomogeneity [135]. Assuming a spatially flat unperturbed model, expression (3.5.12) gives

$$\alpha_{\pm} = -1 \pm \sqrt{1 - \left(\frac{k}{a_0 H_0} \right)^2}. \quad (3.5.13)$$

Following this, we find $\alpha_+ \simeq 0$ and $\alpha_- \simeq -2$ on super-Hubble lengths, which implies that superhorizon-sized inhomogeneities remain constant. Well inside the horizon, on the other hand, solution (3.5.11) and (3.5.13) takes the oscillatory form

$$\Delta_{(k)}^{(\varphi)} = \mathcal{C} t^{-1} \cos \left[\left(\frac{k}{a_0 H_0} \right) \ln t \right], \quad (3.5.14)$$

meaning that small-scale perturbations in the density of the φ -field fluctuate with decreasing amplitude. Finally, at the $k = aH$ threshold we find that $\Delta_{(k)}^{(\varphi)}$ decays as t^{-1} .

3.6. Gravitational wave perturbations

Gravitational waves are propagating fluctuations in the geometry of the spacetime fabric, usually described as weak perturbations of the background metric. Alternatively, one can monitor gravity-wave distortions covariantly by means of the electric and magnetic components of the Weyl tensor [162], which describe the free gravitational field (see Section 1.2.2).

¹⁸ The covariant variable corresponding to the coordinate-based canonical variable for quantisation of scalar field fluctuations, is given by [161]

$$v_a = \frac{a\dot{\varphi}}{3H} \left(\int D_a \theta d\tau - D_a \int \theta d\tau \right). \quad (3.5.9)$$

This gradient variable corresponds to the variable $v = aQ$, where Q is the Mukhanov-Sasaki variable.

3.6.1. Isolating tensor modes

Gravitational waves are covariantly described by the transverse degrees of freedom in the electric (E_{ab}) and magnetic (H_{ab}) parts of the Weyl tensor. The transversality is necessary to ensure that the pure tensor modes of the locally free gravitational field have been isolated. The same condition is also imposed on the shear and any other orthogonally projected, traceless, second-rank tensor that might be present. Thus, when studying the propagation of gravitational radiation in perturbed FLRW models with perfect-fluid matter, we demand that (see Sections 1.2.2 and 1.3.1)

$$D^b E_{ab} = \frac{1}{3} D_a \rho = 0, \quad (3.6.1)$$

$$D^b H_{ab} = \rho(1+w)\omega_a = 0, \quad (3.6.2)$$

$$D^b \sigma_{ab} = \frac{2}{3} D_a \Theta + \text{curl } \omega_a = 0, \quad (3.6.3)$$

to linear order and at all times [163]. In our cosmological environment, this is achieved by switching the vorticity off and by setting $D_a \rho = 0 = D_a p = D_a \Theta$ (for a barotropic medium it suffices to ensure that $D_a \rho = 0 = D_a \Theta$). These constraints, which are self-consistent (i.e. preserved in time) at the linear perturbative level, guarantee that the 4-acceleration also vanishes to first approximation. Then, the only nontrivial linear constraints left are

$$H_{ab} = \text{curl } \sigma_{ab} \quad \text{and} \quad \mathcal{R}_{(ab)} = -H\sigma_{ab} + E_{ab}, \quad (3.6.4)$$

where $\text{curl } \sigma_{ab} = \varepsilon_{cd(a} D^c \sigma_{b)}^d$ (see Eqs. (1.3.8) and (1.3.39) in Sections 1.3.1 and 1.3.5 respectively). Note that, according to (3.6.4)(b), the linear condition $D^b E_{ab} = 0 = D^b \sigma_{ab}$ guarantees that $D^b \mathcal{R}_{(ab)} = 0$ as well. When allowing for anisotropic pressure, one needs to impose the additional constraint

$$D^b \pi_{ab} = 0. \quad (3.6.5)$$

In the case of a magnetised, highly conductive environment, for example, we demand that $D^b \Pi_{ab} = (1/3)D_a B^2 - B^b D_b B_a = 0$ at all times [164,165].

3.6.2. Covariant description of the gravitational-wave energy density

In a perturbed FLRW universe, the energy density of gravitational radiation is determined by the pure tensor part ($\mathcal{H}_{\alpha\beta}^{TT}$, with $\alpha, \beta = 1, 2, 3$) of the metric perturbation, according to (e.g. see [166])

$$\rho_{GW} = \frac{(\mathcal{H}_{\alpha\beta}^{TT})' (\mathcal{H}_{TT}^{\alpha\beta})'}{2a^2}, \quad (3.6.6)$$

where a prime indicates differentiation with respect to conformal time (recall that $c = 1 = 8\pi G$ throughout this review). In a comoving frame, with $u^a = \delta_0^a u^0$, we have [167,168]

$$\sigma_{\alpha\beta} = a(\mathcal{H}_{\alpha\beta}^{TT})' \quad \text{and} \quad \sigma^{\alpha\beta} = a^{-3} (\mathcal{H}_{TT}^{\alpha\beta})', \quad (3.6.7)$$

so that [165]

$$\rho_{GW} = \sigma^2, \quad (3.6.8)$$

which provides a simple covariant expression for the energy density of gravitational-wave distortions in almost FLRW universes.

3.6.3. Evolution of gravitational waves

In a FLRW spacetime the Weyl tensor vanishes identically, which means that E_{ab} and H_{ab} provide a covariant and gauge-invariant description of perturbations in the free-gravitational field. Once the pure tensor modes have been isolated, we can proceed to linearise the propagation equations (1.3.42) and (1.3.43) of Section 1.3.6. Around a Friedmannian background filled with a single perfect fluid, the latter reduce to

$$\dot{E}_{ab} = -3HE_{ab} - \frac{1}{2} \rho(1+w)\sigma_{ab} + \text{curl } H_{ab} \quad (3.6.9)$$

and

$$\dot{H}_{ab} = -3HH_{ab} - \text{curl } E_{ab}, \quad (3.6.10)$$

respectively. Because the magnetic part of the Weyl tensor satisfies the constraint (3.6.4)(a) the linear evolution of H_{ab} is determined by that of the shear, which propagates according to (see expression (1.3.4) in Section 1.3.1)

$$\dot{\sigma}_{ab} = -2H\sigma_{ab} - E_{ab}. \quad (3.6.11)$$

Furthermore, on using the commutation law between the orthogonally projected gradients of spacelike tensors (see Appendix A.3) and the zero-order expression $\mathcal{R}_{abcd} = (K/a^2)(h_{ac}h_{bd} - h_{ad}h_{bc})$ for the 3-Riemann tensor (see Section 1.3.5), constraint (3.6.4)(a) leads to the auxiliary relation $\text{curl} H_{ab} = (3K/a^2)\sigma_{ab} - D^2\sigma_{ab}$. The latter transforms Eq. (3.6.9) into

$$\dot{E}_{ab} = -3HE_{ab} - \frac{1}{2}\rho(1+w)\sigma_{ab} + \frac{3K}{a^2}\sigma_{ab} - D^2\sigma_{ab}, \quad (3.6.12)$$

which together with (3.6.11) monitors the linear evolution of gravitational waves in perturbed FLRW universes.

Taking the time derivative of (3.6.11), using Eq. (3.6.12), the background Raychaudhuri and Friedmann formulae (see (1.4.3)(a), (1.4.4) respectively) and keeping terms of up to linear order only, we arrive at the following wave-like equation for the gravitationally induced shear

$$\ddot{\sigma}_{ab} = -5H\dot{\sigma}_{ab} - \frac{1}{2}\rho(1-3w)\sigma_{ab} + \frac{K}{a^2}\sigma_{ab} + D^2\sigma_{ab}, \quad (3.6.13)$$

with $K = 0, \pm 1$. The above is no longer coupled to the propagation equation of the electric Weyl tensor, which means that the shear wave-equation alone can describe the propagation of gravitational radiation in perturbed, perfect-fluid FLRW cosmologies. We proceed by introducing the standard tensor harmonics $\mathcal{Q}_{ab}^{(k)}$, with $\mathcal{Q}_{ab}^{(k)} = \mathcal{Q}_{(ab)}^{(k)}$, $\dot{\mathcal{Q}}_{ab}^{(k)} = 0 = D^b\mathcal{Q}_{ab}^{(k)}$ and $D^2\mathcal{Q}_{ab}^{(k)} = -(k/a)^2\mathcal{Q}_{ab}^{(k)}$. Then, setting $\sigma_{ab} = \sum_k \sigma_{(k)}\mathcal{Q}_{ab}^{(k)}$, with $D_a\sigma_{(k)} = 0$, expression (3.6.13) provides the following wave equation for the k -th shear mode.

$$\ddot{\sigma}_{(k)} = -5H\dot{\sigma}_{(k)} - \left[\frac{1}{2}\rho(1-3w) - \frac{1}{a^2}(K-k^2) \right] \sigma_{(k)}. \quad (3.6.14)$$

It should be noted here that, in order to account for the different polarisation states of gravitational radiation, one expands the pure-tensor perturbations in terms of the electric and the magnetic parity harmonics (see [169] and references therein). Nevertheless, the coupling between the two states means that Eq. (3.6.14) still holds.

Assuming a spatially flat background and a radiation-dominated universe, we have $w = 1/3$, $H = 1/(2t)$ and $\rho = 3/(4t^2)$. Then, (3.6.14) simplifies to

$$\ddot{\sigma}_{(k)} = -\frac{5}{2t}\dot{\sigma}_{(k)} - \left(\frac{k}{a}\right)^2\sigma_{(k)}, \quad (3.6.15)$$

with the last term vanishing on super-Hubble lengths (asymptotically). The above admits the solution

$$\sigma_{(k)} = t^{-3/2} \left(\frac{k}{aH}\right) \left[\mathcal{C}_1 \sin\left(\frac{k}{aH}\right) + \mathcal{C}_2 \cos\left(\frac{k}{aH}\right) \right] + t^{-3/2} \left[\mathcal{C}_1 \cos\left(\frac{k}{aH}\right) - \mathcal{C}_2 \sin\left(\frac{k}{aH}\right) \right], \quad (3.6.16)$$

where $k/aH \propto t^{1/2}$. Therefore, the amplitude of gravitational waves on small scales in a radiation-dominated and spatially-flat FLRW universe decays as a^{-2} . On super-Hubble scales, on the other hand, $k/aH \ll 1$ and (3.6.16) is approximated by the power law

$$\sigma_{(k)} = \mathcal{C}_1 t^{-3/2} \left[1 - \frac{1}{3} \frac{\mathcal{C}_2}{\mathcal{C}_1} \left(\frac{k}{a_0 H_0}\right)^3 \left(\frac{t}{t_0}\right)^{3/2} \right] = \sigma_0 + \frac{2}{3} \dot{\sigma}_0 t_0 \left[1 - \left(\frac{t_0}{t}\right)^{3/2} \right]. \quad (3.6.17)$$

Consequently, when $\dot{\sigma}_0 = 0$ we have $\sigma_{(k)} = \text{constant}$. On the other hand, for $\dot{\sigma}_0 = -3\sigma_0/2t_0$ we find $\sigma_{(k)} \propto t^{-3/2} \propto a^{-3}$.

After equality, we have $w = 0$, $H = 2/(3t)$ and $\rho = 4/(3t^2)$. In this environment, Eq. (3.6.14) takes the form

$$\ddot{\sigma}_{(k)} = -\frac{10}{3t}\dot{\sigma}_{(k)} - \frac{2}{3t^2} \left[1 + \frac{2}{3} \left(\frac{k}{aH}\right)^2 \right] \sigma_{(k)}. \quad (3.6.18)$$

Hence, on superhorizon scales,

$$\sigma_{(k)} = \mathcal{C}_1 t^{-1/3} + \mathcal{C}_2 t^{-2}, \quad (3.6.19)$$

implying that after equality large-scale gravitational wave perturbations decay as $a^{-1/2}$.

3.7. Perturbed non-Friedmannian cosmologies

Current observational data strongly support the homogeneous and isotropic FLRW spacetimes as the best model for our universe on very large scales. On the other hand, the presence of nonlinear structures, in the form of galaxies, galaxy clusters and superclusters, shows that on the relevant scales the universe is neither homogeneous nor isotropic. For this reason and also because we have no real information about the very early stages of our universe, it may be unwise to exclude a priori all the non-FLRW cosmologies from our studies. In what follows we will consider the stability of such unconventional universes against linear perturbations.

3.7.1. The Bianchi I universe

Bianchi models have been studied by several authors in an attempt to achieve better understanding of the observed small amount of anisotropy in the universe. The same models have also been used to examine the role of certain anisotropic sources during the formation of the large-scale structure we see in the universe today. Some Bianchi cosmologies, for example, are natural hosts of large-scale magnetic fields and therefore their study can shed light on the implications of cosmic magnetism for galaxy formation. The simplest Bianchi family that contains the flat FLRW universe as a special case are the type-I spacetimes (see Table 1).

Using covariant techniques, the stability of the Bianchi I models against linear perturbations was studied in [170], for the case of a single perfect fluid, and also in [128], where a magnetic field was also present. Assuming non-magnetised dust for simplicity, Eqs. (2.3.1) and (2.3.2) linearise to

$$\dot{\Delta}_a = -Z_a - \sigma_{ab}\Delta^b \quad (3.7.1)$$

and

$$\dot{Z}_a = -\frac{2}{3}\Theta Z_a - \frac{1}{2}\rho\Delta_a - \sigma_{ab}Z^b - 2aD_a\sigma^2, \quad (3.7.2)$$

respectively. The above monitor the linear evolution of density inhomogeneities in the dust component on a Bianchi I background. The system is obviously not closed and one also needs the propagation formula of the shear and the shear constraint (see expressions (1.3.4) and (1.3.6) respectively). To first order and in the absence of fluid pressure, these are respectively given by

$$\dot{\sigma}_{ab} = -\frac{2}{3}\Theta\sigma_{ab} - \sigma_{c(a}\sigma^c{}_{b)} - E_{ab} \quad (3.7.3)$$

and

$$D^b\sigma_{ab} = \frac{2}{3a}Z_a + \text{curl}\omega_a. \quad (3.7.4)$$

The above immediately bring into play the electric component of the Weyl field, the linear evolution of which depends on its magnetic counterpart (see Eq. (1.3.42) in Section 1.3.6), as well as the vorticity. The complexity of the mathematics means that analytical progress can be achieved in special cases only. In [170], for example, the perturbed model is assumed both irrotational and axially symmetric. It was then possible to obtain analytical solutions for the two independent components of the density contrast in the form of power series. Confining to early times, when $t \rightarrow 0$, the latter converge to

$$\Delta_1 = C_1 t^{-2/3} + C_2 t^{-5/3} \quad \text{and} \quad \Delta_{2,3} = C_3 t^{10/3} + C_4 t^{1/3}, \quad (3.7.5)$$

with the C s representing the integration constants. This result argues for growth along two of the directions that is faster than the standard Einstein–de Sitter rate (compare to Eq. (3.2.24) in Section 3.2.3) and agrees qualitatively with the analysis of [171].

In the magnetised study of [128] the anisotropic pressure of the B -field (see expression (2.5.3) in Section 2.5.1) adds further complications to the system of (3.7.1) and (3.7.2). Treating the magnetic field as the sole source of anisotropy, makes B_a a shear eigenvector and considerably simplifies the mathematics. Then, for magnetic strengths compatible with the high isotropy of the CMB spectrum, one can obtain analytical solutions and establish the corrections to the FRW-related results of Section 3.4. Following [128], the anisotropy of the background model brings about the tension properties of the B -field. This happens through the general relativistic coupling between the field and the spatial curvature of the perturbed model and makes the overall magnetic effect sensitive to the amount of the 3-curvature distortion.

3.7.2. The Einstein static universe

The possibility that our universe might have started out in an asymptotically static state, reminiscent of the Eddington–Lemaître cosmology, has been discussed by several authors (e.g. see [172–176] and references therein). It is therefore useful to investigate the stability of the family of the Einstein static spacetimes, which has long been known to be unstable against homogeneous and isotropic perturbations [177]. In fact, the instability of this model is very well established among the community, despite later work showing that the issue is not as clear cut as the Newtonian intuition may suggest [178].

In an Einstein static background $H = 0 = \dot{H}$ and $K = +1$. Then, formulae (1.4.3) and (1.4.4) in Section 1.4.2 ensure that ρ (and therefore p , w) is time invariant, $\Lambda = (1 + 3w)\rho/2$ and $1/a^2 = (1 + w)\rho/2$. In this environment expression (3.2.16) reduces to [179]

$$\ddot{\Delta} = \frac{1}{2}(1 + w)(1 + 3c_s^2)\rho\Delta + c_s^2 D^2\Delta, \quad (3.7.6)$$

which monitors the linear evolution of scalar/density perturbations in a perturbed Einstein static universe. By employing the standard scalar harmonics (see Section 3.2.3 for more details), the above decomposes into

$$\ddot{\Delta}^{(k)} = \frac{1}{2}(1 + w)\left[1 + (3 - k^2)c_s^2\right]\rho\Delta^{(k)}, \quad (3.7.7)$$

and has a (neutrally) stable, oscillatory solution as long as $(k^2 - 3)c_s^2 > 1$. Recalling that in spatially closed models the eigenvalue of a given mode and its comoving wavenumber are discrete and related by $k^2 = \nu(\nu + 2)$, with $\nu = 1, 2, 3, \dots$, we conclude that a perturbed mode is stable against gravitational collapse if [179]

$$[\nu(\nu + 2) - 3]c_s^2 > 1, \quad (3.7.8)$$

to linear order. The first inhomogeneous mode has $\nu = 1$ and therefore it does not satisfy the stability condition. Nevertheless, this is simply a gauge mode reflecting our freedom to change the fundamental 4-velocity vector. Hence, all physical modes have $\nu \geq 2$ and for them linear stability is guaranteed as long as $c_s^2 > 1/5$.¹⁹

The physical explanation for this rather unexpected stability lies in the Jeans length of the Einstein static model [178]. Although there are always unstable modes in flat spaces, namely wavelengths larger than the associated Jeans scale, a closed universe sets an upper limit to the allowed wavelengths. The above analysis shows that, for a sufficiently large sound speed, all physical wavelengths fall below the Jeans scale. Noting that the linear stability of the Einstein static model does not guarantee its overall stability [181], we point out that current data seem to favour a slightly closed universe [156]. Also, spatially closed static models violate the inflationary singularity theorems of [182–184] and therefore can avoid the quantum-gravity era [185,172].

3.7.3. The Gödel universe

Gödel's universe is an exact solution of the Einstein field equations that is both stationary and spatially homogeneous [186]. The Gödel spacetime is also rotationally symmetric about each point and well known for its unusual properties. The most intriguing among them is the existence of closed timelike curves, which violates global causality and makes time travel a theoretical possibility.²⁰ Although Gödel's world is not a realistic model of the universe we live in, it has been widely used to study and illustrate the effects of global vorticity within the realm of general relativity (e.g. see [189–191] and references therein).

Relative to a timelike 4-velocity field, tangent to the world lines of the fundamental observers, the kinematics of the Gödel spacetime are covariantly described by [41,45]

$$\Theta = 0 = A_a = \sigma_{ab} = H_{ab} \quad \text{and} \quad \omega_a, E_{ab} \neq 0, \quad (3.7.9)$$

where $\nabla_b \omega_a = 0 = \dot{E}_{ab} = D_c E_{ab}$ and $\nabla_c E_{ab} \neq 0$. Also, the stationary nature of the model means that all the propagation equations have been transformed into constraints, some of which are trivial. In particular, the Gödel analogues of the continuity, the Raychaudhuri and the Friedmann formulae (see Sections 1.3.1, 1.3.3 and 1.3.5) read

$$\dot{\rho} = 0, \quad \frac{1}{2} \rho(1 + 3w) - 2\omega^2 - \Lambda = 0 \quad (3.7.10)$$

and

$$\rho - \omega^2 + \Lambda = 0, \quad (3.7.11)$$

respectively (see also [192] for more technical details). We note that expressions (3.7.10)(b) and (3.7.11) combine to give $\rho(1 + w) = 2\omega^2$ and $p - \rho = 2\Lambda$. The former of these relations shows that the vorticity provides a direct measure of the model's inertial mass. The latter ensures that the value of the cosmological constant depends on the equation of state of the fluid that fills the Gödel spacetime. The amount of rotation also determines the radius, $R_G = \sqrt{2} \ln(1 + \sqrt{2})/\omega$, of the observers causal region. Using the above given relations, one can linearise the nonlinear formulae of Section 2.3 around a Gödel background. More specifically, Eqs. (2.3.1) and (2.3.2) reduce to the system [192]

$$\dot{\Delta}_a = -(1 + w)Z_a + \omega_{ab}\Delta^b, \quad (3.7.12)$$

$$\dot{Z}_a = -\frac{1}{2}\rho\Delta_a - \frac{3}{2}aD_a p + aD_a D^b A_b + 2aD_a \omega^2 + \omega_{ab}Z_b, \quad (3.7.13)$$

with $\rho(1 + w)A_a = -D_a p$. Assuming rigid rotation, we may ignore the second-last term in the right-hand side of (3.7.13). Then, the orthogonally projected derivatives of the above combine to give the following wavelike equation [192]

$$\ddot{\Delta}^{(k)} = \left[\omega^2(1 + 3c_s^2) - \left(\frac{kc_s}{a} \right)^2 \right] \Delta^{(k)}, \quad (3.7.14)$$

for the linear evolution of the k -th harmonic mode. Thus, for dust there is no pressure support and matter aggregations grow unimpeded.²¹ Otherwise, there is an effective Jeans length, equal to $\lambda_{Jc} = c_s/\omega\sqrt{1 + 3c_s^2}$, below which the inhomogeneities

¹⁹ The same condition for the linear stability of the Einstein static universe was also obtained in [178] in the restricted case of conformal metric perturbations. The stability of the radiation-filled model and the instability of the one containing pressureless dust was demonstrated in [180].

²⁰ For additional examples of spacetimes with closed timelike curves the reader is referred to [187,188].

²¹ Gödel's solution was originally given for dust (i.e. $\rho \neq 0$, $p = 0$ and $\Lambda \neq 0$). Nevertheless, by introducing the transformation $\rho \rightarrow \rho' = \rho + p$ and $\Lambda \rightarrow \Lambda' = \Lambda + p$, we can reinterpret the Gödel spacetime as a perfect-fluid model.

oscillate [192]. The same type of neutral stability was also claimed for gravitational wave perturbations, while the evolution of linear rotational distortions was found to depend on the amount of the shear anisotropy (see [192] for further discussion).

We point out that the value of the Jeans length is comparable to the radius of the smallest closed timelike curve (see above). This means that the causal regions in Gödel models with nonzero pressure are stable against linear matter aggregations. Also, expression (3.7.14) demonstrates how the vorticity of the Gödel universe contributes to the overall gravitational pull. This result suggests that rotational energy has ‘weight’ and seems to favour the de Felice [193] and the Barrabès et al. [194] interpretation of the Abramowicz–Lasota ‘centrifugal-force reversal’ effect [195].

The above analysis does not specifically address the stability of Gödel’s closed timelike curves, a issue recently considered in [196]. The appearance of closed timelike curves in general relativistic Gödel-type cosmologies, namely in rigidly rotating homogeneous spacetimes, can be avoided by introducing extra matter sources, higher-order terms in the Laplacian or string-theory corrections [197–199]. This does not seem to be the case, however, in Gödel-type brane models [200].

3.7.4. The Lukash Bianchi VII_h universe

The most general Bianchi universes that contain the spatially open FLRW model as a special subcase, are those of type VII_h. The late-time asymptotes of the non-tilted Bianchi VII_h cosmologies evolve towards a vacuum plane-wave solution known as the Lukash universe [201,202]. These spacetimes describe the most general effects of spatially homogeneous perturbations on open Friedmann universes and the Lukash metric plays a guiding role in these investigations (see [80] and references therein).

In the absence of matter, the plane-wave attractors of the Bianchi VII_h models are covariantly characterised by

$$A_a = 0 = \omega_a \quad \text{and} \quad \Theta, \sigma_{ab}, E_{ab}, H_{ab} \neq 0, \tag{3.7.15}$$

with $E_{ab}E^{ab} = H_{ab}H^{ab}$ and $E_{ab}H^{ab} = 0$ [90,91]. Then, the Lukash analogues of the Raychaudhuri and Friedmann equations read

$$\dot{\Theta} = -\frac{1}{3}\Theta^2 - 2\sigma^2 \quad \text{and} \quad \mathcal{R} = -\frac{2}{3}\Theta^2 + 2\sigma^2, \tag{3.7.16}$$

respectively. We note that the 3-curvature scalar \mathcal{R} is negative, thus guaranteeing the hyperbolic spatial geometry of the model (e.g. see [90,91]). Also, despite the absence of matter fields, the Lukash metric is not Ricci flat and the associated Gauss–Codacci formula is

$$\mathcal{R}_{(ab)} = -\frac{1}{3}\Theta\sigma_{ab} + \sigma_{c(a}\sigma^c_{b)} + E_{ab}. \tag{3.7.17}$$

Using the dimensionless parameter $\Sigma = 3\sigma^2/\Theta^2$ to measure the expansion anisotropy, one can also recast the Raychaudhuri equation into the following alternative expression [91]

$$\dot{\Theta} = -\frac{1}{3}\Theta^2(1 + 2\Sigma), \tag{3.7.18}$$

where $0 < \Sigma < 1$ in accord with $\mathcal{R} < 0$ in (3.7.16)(b). Then, the average scale factor obeys a simple power-law evolution with $a \propto t^{1/(1+2\Sigma)}$. In the absence of shear effects, we recover the familiar $a \propto t$ evolution of the Milne universe. As $\Sigma \rightarrow 1$, on the other hand, we approach the Kasner vacuum solution (i.e. $a \propto t^{1/3}$). Note that for maximum shear the spatial curvature of the model vanishes, while only its isotropic part survives at the opposite end.

The nature of the Lukash solution makes it a good testing ground for studying the final stages of ever-expanding FLRW cosmologies with positive gravitational mass (i.e. $\rho + 3p > 0$). Allowing for a low-density, pressure-free matter component, we may identify our fundamental 4-velocity field with that of the fluid. Then, linearising Eq. (1.3.4) around the Lukash background and using (1.3.41), we find that shear perturbations are monitored by the system [91]

$$\dot{\bar{S}} = -\frac{4}{3}\bar{\Theta}(1 - \bar{S})\Sigma - 3\bar{\Theta}^{-2}\mathcal{R}_{(ab)}\bar{\sigma}^{ab} \tag{3.7.19}$$

and

$$\dot{\bar{S}} = \frac{4}{3}\bar{\Theta}(1 - \bar{S})S + 3\bar{\Theta}^{-2}\mathcal{R}_{(ab)}\bar{\sigma}^{ab}, \tag{3.7.20}$$

where the overbars indicate zero order quantities. Here $S = 1 - \Sigma$ by definition and overbars indicate quantities of zero perturbative order. Expression (3.7.19) allows for a gauge-independent description of linear shear anisotropies at the $\bar{S} \rightarrow 0$ limit, while Eq. (3.7.20) does the same as $\bar{S} \rightarrow 1$ (i.e. when $\bar{S} \rightarrow 0$). In the first instance the last term in the right-hand side of (3.7.19) vanishes and therefore $\Sigma \propto a^{-4}$, meaning that any kinematic anisotropies that may occur will quickly disperse. When $\bar{S} \rightarrow 1$, on the other hand, the shear can increase further and therefore force the Lukash solution away from the Bianchi VII_h family. In other words, the linear stability of the Lukash universe appears to depend on the amount of the background shear anisotropy, something also seen in the study of linear vortices and gravitational wave perturbations (see [91] and also [203]).

4. Kinetic theory and the cosmic neutrino and microwave backgrounds

Kinetic theory provides a self-consistent description of a gas of particles and naturally includes the free-streaming limit, where collisions are negligible, and the hydrodynamic limit, where collisions maintain tight-coupling and the gas approaches fluid-like behaviour. In cosmology, only the free-streaming limit is appropriate for the neutrino background for all times after decoupling (at temperature $\sim 10^{11}$ K). However, for the cosmic microwave background (CMB), Compton scattering off electrons in the electron–baryon plasma prior to recombination (at temperature ~ 3000 K) makes a fluid description adequate on comoving scales larger than 30 Mpc. On smaller scales, perturbations in the density of the radiation are damped due to photon diffusion [146]. Kinetic theory provides a seamless description of the associated transitions from ideal fluid behaviour, through that of an imperfect fluid to a free-streaming (collisionless) gas.

In this section we review the 1 + 3-covariant formulation of relativistic kinetic theory and its application to the (massive) neutrino background and the anisotropies and polarization of the CMB. Our emphasis is on the physics involved, but we do include brief discussions of current observations where this is helpful. For textbook discussions on relativistic kinetic theory, see Refs. [204–206].

4.1. Distribution functions and the Liouville equation

We consider a gas of identical particles each of mass m . The four-momentum of a particle p^a can be decomposed into an energy and three-momentum with respect to the velocity field u^a as

$$p^a = Eu^a + \lambda^a = Eu^a + \lambda e^a, \quad (4.1.1)$$

where $\lambda = \sqrt{E^2 - m^2}$ is the magnitude of the three-momentum and the projected vector e^a is the propagation direction. For massless particles, like photons, $\lambda = E$. In situations where the polarization (or helicity) is not important, we can describe the gas by a scalar-valued one-particle distribution function $f(x^a, p^a)$. The number of particles in a proper phase space element $d^3\mathbf{x}d^3\mathbf{p}$ is then $fd^3\mathbf{x}d^3\mathbf{p}$. For a given set of particles, their phase space volume is both Lorentz invariant (i.e. the same for all observers) and, in the absence of collisions, constant along their path. Introducing an affine parameter τ , normalised such that $p^a = dx^a/d\tau$, the collisionless evolution of the gas is described by the Liouville equation

$$\frac{df}{d\tau} = 0, \quad (4.1.2)$$

where the derivative is along the path in phase space. In the presence of collisions, the right-hand side should be replaced by the appropriate Lorentz-invariant collision operator $C[f]$, which is a functional of f , to give a Boltzmann equation.

The 4-momentum of a free particle is parallel-transported so that $p^b\nabla_b p^a = 0$. We can use this to find how the energy (or λ) and direction propagate by projecting along and perpendicular to u^a respectively. For λ ,

$$\frac{d\lambda}{d\tau} = -E^2 A_a e^a - E\lambda \left(\sigma_{ab} e^a e^b + \frac{1}{3} \Theta \right). \quad (4.1.3)$$

In the Robertson–Walker limit, $d\lambda/d\tau = -E\lambda H$ where $H = \Theta/3$ is the Hubble parameter. The momentum thus redshifts as the inverse of the scale factor, $1/a$. For the direction,

$$\frac{de^{(a)}}{d\tau} = -\frac{E^2}{\lambda} s^a{}_b A^b - E \left(\omega^a{}_b e^b + s^{ab} \sigma_{bc} e^c \right), \quad (4.1.4)$$

where $s_{ab} \equiv h_{ab} - e_a e_b$ is the screen-projection tensor which projects into the two-dimensional screen perpendicular to the propagation direction e^a in the local rest-space of u^a . Note that the derivative on the left of Eq. (4.1.4) is the covariant derivative along the particle's path. The normalisation condition $e^a e_a = 1$ is preserved since $de^{(a)}/d\tau$ is perpendicular to e^a . In the Robertson–Walker limit, $de^{(a)}/d\tau = 0$, i.e. e^a is as constant as the constraint $e^a u_a = 0$ allows. In the real universe this is no longer so and Eq. (4.1.4) then describes the action of gravitational lensing (see e.g. Ref. [207] for a recent review).

The distribution function depends on spacetime position, energy and direction. For the direction dependence, it is convenient to expand in spherical multipoles as [208,209]

$$f(x^a, p^a) = \sum_{l=0}^{\infty} F_{A_l}(x^a, E) e^{A_l} = F(E) + F_a(E) e^a + F_{ab}(E) e^a e^b + \dots, \quad (4.1.5)$$

where the tensors $F_{A_l}(E) = F_{(a_1 \dots a_l)}(E)$ are projected (orthogonal to u^a) symmetric and trace-free so are irreducible under the action of three-dimensional rotations. The expansion (4.1.5) is equivalent to an expansion in spherical harmonics, but has the advantage of being fully covariant. We can invert the expansion as follows:

$$F_{A_l}(E) = \frac{1}{\Delta_l} \int f e_{(A_l)} d\Omega \quad \text{where } \Delta_l \equiv \frac{4\pi 2^l (l!)^2}{(2l+1)!}, \quad (4.1.6)$$

and we have used

$$\int e_{(A_l)} e^{(B_l)} d\Omega = \Delta_l h_{(A_l)}^{(B_l)} \delta_{ll} = \Delta_l h_{(a_1}^{(b_1} \dots h_{a_l)}^{b_l)} \delta_{ll}. \quad (4.1.7)$$

For further details on the covariant multipole expansion, see e.g. Ref. [210].

The propagation equations for the multipole moments follows from substituting Eq. (4.1.5) into the Boltzmann equation, using Eqs. (4.1.3) and (4.1.4), and extracting the irreducible terms. The result is

$$\begin{aligned} E\dot{F}_{(A_l)} - \lambda^2 \frac{\Theta}{3} \frac{\partial F_{A_l}}{\partial E} + \frac{l+1}{2l+3} \lambda D^a F_{aA_l} + \lambda D_{(a_l} F_{A_{l-1})} + l E F_{a(A_{l-1}} \omega_{a_l)}^a - \left[\lambda E \frac{\partial F_{(A_{l-1}}}{\partial E} - (l-1) \frac{E^2}{\lambda} F_{(A_{l-1}} \right] A_{a_l)} \\ - \frac{l+1}{2l+3} \left[(l+2) \frac{E^2}{\lambda} F_{aA_l} + \lambda E \frac{\partial F_{aA_l}}{\partial E} \right] A^a - \frac{l}{2l+3} \left[3 E F_{a(A_{l-1}} + 2 \lambda^2 \frac{\partial F_{a(A_{l-1}}}{\partial E} \right] \sigma_{a_l)}^a \\ - \frac{(l+1)(l+2)}{(2l+3)(2l+5)} \left[(l+3) E F_{abA_l} + \lambda^2 \frac{\partial F_{abA_l}}{\partial E} \right] \sigma^{ab} - \left[\lambda^2 \frac{\partial F_{(A_{l-2}}}{\partial E} - (l-2) E F_{(A_{l-2}} \right] \sigma_{a_{l-1} a_l)} = C_{A_l}[f], \end{aligned} \quad (4.1.8)$$

where $C_{A_l}[f]$ are the multipoles of the invariant collision term. The spacetime derivatives are all taken at fixed E (or λ). This equation was first obtained in Ref. [209] (and in Ref. [208] for the massless case), but the form given here benefits from the streamlined notation introduced in Ref. [46]. The original Boltzmann equation contains a spacetime derivative along the particle path which, when split into multipoles, connects the derivative of $F_{A_l}(E)$ along u^a with the projected derivatives of the $l-1$ and $l+1$ moments. This describes the generation of anisotropy from spatial inhomogeneities in the distribution function. For a fluctuation of characteristic size a/k , the timescale for propagating anisotropy through $\Delta l = 1$ is $\sim a/(kv)$ where $v = \lambda/E$ is the magnitude of the particle's 3-velocity. This is simply the time taken for a particle to traverse the inhomogeneity. The terms in Eq. (4.1.8) containing derivatives with respect to E arise from the redshifting of the particle's energy. We see from Eq. (4.1.3) that the isotropic expansion, acceleration and shear source anisotropy at multipole l from $l, l \pm 1$ and both $l \pm 2$ and l respectively. The remaining terms in Eq. (4.1.8) arise from the evolution of the particle's direction: vorticity, acceleration and shear source anisotropy at l from $l, l \pm 1$ and both $l \pm 2$ and l respectively.

4.1.1. Bulk properties

The stress–energy tensor of the gas is determined from the one-particle distribution function by

$$T_{ab} = \int f p_a p_b \frac{d^3 \mathbf{p}}{E}. \quad (4.1.9)$$

Note that $d^3 \mathbf{p}/E$ is the Lorentz-invariant volume element on the positive-energy mass shell $p^a p_a = -m^2$. Decomposing T_{ab} into energy and momentum densities, isotropic pressure and anisotropic stress, as in Eq. (1.2.11), we have

$$\rho = \Delta_0 \int_0^\infty d\lambda \lambda^2 E F, \quad (4.1.10)$$

$$q_a = \Delta_1 \int_0^\infty d\lambda \lambda^2 E (\lambda/E) F_a, \quad (4.1.11)$$

$$\pi_{ab} = \Delta_2 \int_0^\infty d\lambda \lambda^2 E (\lambda/E)^2 F_{ab}, \quad (4.1.12)$$

$$p = \frac{\Delta_0}{3} \int_0^\infty d\lambda \lambda^2 E (\lambda/E)^2 F, \quad (4.1.13)$$

where F is the monopole of the distribution function. The equations of motion for these quantities follow from integrating Eq. (4.1.8) with $\lambda^2 d\lambda$ and appropriate powers of the velocity-weight λ/E . For the energy and momentum densities,

$$\dot{\rho} = -\Theta(\rho + p) - D^a q_a - 2A^a q_a - \sigma^{ab} \pi_{ab} + \Delta_0 \int_0^\infty d\lambda \lambda^2 C_0[f] \quad (4.1.14)$$

$$\dot{q}_{(a)} = -\frac{4}{3} \Theta q_a + (\rho + p) A_a - D_a p - D^b \pi_{ab} - (\omega_a{}^b + \sigma_a{}^b) q_b - A^b \pi_{ab} + \Delta_1 \int_0^\infty d\lambda \frac{\lambda^3}{E} C_a[f], \quad (4.1.15)$$

where $C_0[f]$ is the monopole of the collision term. The gas can exchange energy and momentum through interactions with external particles and/or fields, and these processes are described by the final terms in Eqs. (4.1.14) and (4.1.15). In their presence, the stress-tensor is not conserved but instead has divergence

$$\nabla^b T_{ab} = u_a \Delta_0 \int_0^\infty d\lambda \lambda^2 C_0[f] + \Delta_1 \int d\lambda \frac{\lambda^3}{E} C_a[f] \quad (4.1.16)$$

$$= \int C[f] p_a \frac{d^3 \mathbf{p}}{E}, \quad (4.1.17)$$

which is manifestly a 4-vector.

The propagation equations for the energy and momentum densities do not form a closed system even when there are no interactions. For massive particles, the pressure is not simply related to ρ and so the equation of state is dynamical.

Moreover, a propagation equation for the anisotropic stress is also needed to close the system. The required information is, of course, contained in the original Boltzmann equation which can be recast as a two-dimensional, infinite closed hierarchy for the moments of f integrated over energy with positive (integer) velocity-weights [209,211]. The integrated moments contain all those that appear in the stress-tensor as a subset. The two-dimensional hierarchy simplifies in a number of important special cases. For relativistic matter, $\lambda \sim E$, the hierarchy becomes one-dimensional and this is appropriate for the cosmological neutrino background at temperatures $T \gg m$. For tightly-coupled collisional matter, such as the CMB in the pre-recombination era when Thomson scattering is efficient (see Section 4.3), anisotropies at multipole l are suppressed by $(v_* k t_{\text{coll}}/a)^l$, where a/k is the scale of inhomogeneity, t_{coll} is the collision time and v_* is a typical particle speed. Finally, for non-relativistic matter the hierarchy can be truncated at low velocity weight²² in which case only a small number of moments need be propagated. The latter truncation scheme is used to study the effect of velocity dispersion on linear structure formation in Refs. [211,212].

The other important bulk properties of the gas are the particle and entropy fluxes. We shall only consider the former here; for relativistic thermodynamics, see e.g. Ref. [213]. The particle flux is given in terms of the distribution function by

$$N_a = \int f p_a \frac{d^3 \mathbf{p}}{E}. \quad (4.1.18)$$

The number density and particle drift evaluate to

$$n = \Delta_0 \int_0^\infty d\lambda \lambda^2 F, \quad (4.1.19)$$

$$\mathcal{N}_a = \Delta_1 \int_0^\infty d\lambda \lambda^2 (\lambda/E) F_a. \quad (4.1.20)$$

The propagation equation for n follows from integrating the $l = 0$ moment of Eq. (4.1.8) over $\lambda^2 d\lambda$:

$$\dot{n} = -\Theta n - D^a \mathcal{N}_a - A^a \mathcal{N}_a + \Delta_0 \int_0^\infty d\lambda \frac{\lambda^2}{E} C_0[f], \quad (4.1.21)$$

so that the divergence of the particle flux is

$$\nabla^a N_a = \Delta_0 \int_0^\infty d\lambda \frac{\lambda^2}{E} C_0[f] \quad (4.1.22)$$

$$= \int C[f] \frac{d^3 \mathbf{p}}{E}. \quad (4.1.23)$$

As for the stress-tensor, Eq. (4.1.21) is part of a two-dimensional infinite hierarchy of integrated moments [209]. With an approximate truncation, this can be solved to determine the evolution of the particle flux.

We end by noting that if we chose the frame u^a such that $q_a = 0$ (the energy frame), the particle flux will generally not vanish. Hence the energy and particle frame generally differ. An important exception is for linear CMB fluctuations, where the energy-dependence of the dipole $F_a(E)$ (and higher multipoles) factorises; see Section 4.3.

4.1.2. Linearisation around FLRW cosmologies

An important result due to Ehlers, Geren and Sachs [214] follows from the exact multipole equation (4.1.8): if there exists a family of free-falling observers who measure freely-propagating self-gravitating radiation to be exactly isotropic in some domain of a dust-dominated universe, the spacetime is *exactly* FLRW in that region. Of course, the CMB is not exactly isotropic but the result can be shown to be stable in the sense that if a family of observers sees almost isotropic radiation, the universe is close to FLRW [215] (but see Ref. [216] for a critique of the technical assumptions involved). Then, if we accept the (spacetime) Copernican assumption, the currently observed isotropy of the CMB implies that the geometry of the universe is well described by an almost-FLRW model (at least since recombination).

For an FLRW model, isotropy demands that the distribution function be isotropic and homogeneous, i.e. the only non-zero multipole is the monopole, and this has vanishing projected gradient at fixed energy. According to the discussion in Section 2.2.1, the $l \geq 1$ multipoles and the projected derivative of the monopole are, therefore, gauge-invariant measures of perturbations in the distribution function about an FLRW model. As for most covariant and gauge-invariant perturbations, the variables do, however, depend on the choice of frame u^a . This dependence is discussed further for massless particles in Section 4.3.2. If we consider small departures from FLRW, the covariant and gauge-invariant variables will themselves be small and we can safely ignore products between small quantities.

In an FLRW background, the Liouville equation (4.1.2) for collisionless matter reduces to

$$\frac{\partial f}{\partial t} - H\lambda \frac{\partial f}{\partial \lambda} = 0. \quad (4.1.24)$$

²² This is accurate provided that the typical free-streaming distance per Hubble time is small compared to the size of the inhomogeneity.

This is solved by $f = f(a\lambda)$ where a is the Robertson–Walker scale factor, as follows also from noting that $a\lambda$ is conserved along the particle path. In the perturbed universe, it is convenient to introduce the comoving momentum $q \equiv a\lambda$ and energy $\epsilon \equiv aE$. It is q that is conserved in the background, while $\epsilon^2 = q^2 + a^2 m^2$ exactly. We can then write the distribution function as $f(x^a, q, \epsilon^a)$ or in terms of angular multipoles $F_{A_l}(x^a, q)$.

Before proceeding, it is worth making the following remarks about the scale factor a in the perturbed universe. The scale factor is defined by integrating $\dot{a}/a = \Theta/3$ and, for a given choice of u^a , is only defined up to a hypersurface. The initial hypersurface on which $a = \text{const.}$ should be chosen physically so as to ensure that $h_a \equiv D_a a = 0$ if the model is FLRW. The exact propagation equation for h_a follows from commuting the space and time derivatives to find

$$\dot{h}_{(a)} = \frac{1}{3} a (\Theta A_a + D_a \Theta) - \sigma_a^b h_b + \omega_a^b h_b. \quad (4.1.25)$$

It follows that \dot{h}_a is well-defined at first-order, despite the first-order hypersurface ambiguity in both h_a and a .

The multipole form of the Boltzmann equation (4.1.8) contains spacetime derivatives taken at fixed E (or λ). If instead we take the derivative at fixed q , we have

$$\nabla_a F_{A_l}|_\lambda = \nabla_a F_{A_l}|_q + \left(\frac{1}{a} h_a - \frac{1}{3} \Theta a u_a \right) q \frac{\partial F_{A_l}}{\partial q}. \quad (4.1.26)$$

Using this result in Eq. (4.1.8), and dropping terms that are second-order, we obtain the linearised multipole equations [211]

$$\dot{F}_{(A_l)} + \frac{q}{\epsilon} \frac{l+1}{2l+3} D^b F_{bA_l} + \frac{q}{\epsilon} D_{(a_l} F_{A_{l-1})} + \delta_{l1} \left(\frac{1}{a} h_{a_1} - \frac{\epsilon}{q} A_{a_1} \right) q \frac{\partial F}{\partial q} - \delta_{l2} \sigma_{a_1 a_2} q \frac{\partial F}{\partial q} = \frac{a}{\epsilon} C_{A_l}[f], \quad (4.1.27)$$

where all spacetime derivatives are at fixed q . The $l = 1$ equation contains the first-order combination

$$\mathcal{V}_a(q) \equiv a D_a F + h_a q \frac{\partial F}{\partial q} = a D_a F|_\lambda, \quad (4.1.28)$$

which removes the hypersurface ambiguity of $D_a F$ and h_a . The propagation equation for $\mathcal{V}_a(q)$ follows from Eq. (4.1.27) for $l = 0$. Taking the projected derivative and commuting time and spatial derivatives gives

$$\dot{\mathcal{V}}_{(a)} = -\frac{a}{3} \frac{q}{\epsilon} D_a D^b F_b + \dot{h}_{(a)} q \frac{\partial F}{\partial q} + a D_a \left(\frac{C_0[f]}{E} \right) \Big|_\lambda + a A_a \frac{C_0[f]}{E}, \quad (4.1.29)$$

with all spacetime derivatives at fixed q except where stated otherwise. In the absence of collisions, this equation, together with the $l > 0$ multipole equations, form a closed system (once supplemented with the usual kinematic equations) with which one can propagate the perturbations to the distribution function directly. The collisionless form of these equations is well known in the synchronous and Newtonian gauge [217] and are what is usually used to propagate massive neutrino perturbations numerically [217–219].

The general strategy for solving Eq. (4.1.27) is to decompose the spatial dependence of F_{A_l} into scalar, vector and tensor parts which evolve independently in linear theory. The essential idea is that the scalar component of F_{A_l} is obtained by taking the PSTF part of l projected derivatives of some scalar field, the vector component by the PSTF part of $l - 1$ projected derivatives of a (projected) divergence-free vector field, and the tensor component from the PSTF part of $l - 2$ derivatives of a PSTF, divergence-free rank-2 tensor field. The decomposition is unique, although the tensor potentials are generally not, provided the tensor field being expanded satisfies appropriate boundary conditions (for non-compact spaces) [107]. In general, for $l > 2$, a rank- l PSTF tensor can have higher-rank tensor contributions. However, in linear theory there are no gravitational source terms for the higher-rank contributions and so, starting from an early epoch when interactions are efficient in maintaining isotropy, the higher-rank contributions will not be present.

The tensor potentials for the scalar, vector and tensor contributions can be expanded in terms of sets of appropriate harmonic functions that are complete in FLRW spaces. Given FLRW symmetry, a convenient choice is the eigenfunctions of the (comoving) projected Laplacian $a^2 D^2$. These can be chosen to be orthogonal, and this property will be inherited by the tensors derived from them. In the following paragraphs we briefly summarise the properties of these harmonic functions, bringing together and extending results in Refs. [167,220,221].

Scalar perturbations. For these we expand in terms of scalar-valued eigenfunctions satisfying

$$a^2 D^2 \mathcal{Q}^{(0)} + k^2 \mathcal{Q}^{(0)} = 0, \quad (4.1.30)$$

with $\dot{\mathcal{Q}}^{(0)} = 0$. These equations hold only at zero-order, i.e. the harmonic functions are defined on the FLRW background. The superscript (0) denotes scalar perturbations, and, to avoid clutter, in this section we suppress the index (k). The allowed eigenvalues k^2 depend on the spatial curvature of the background model. Defining $\nu^2 = (k^2 + K)/|K|$, where $6K/a^2$ is the curvature scalar of the FLRW spatial sections, the regular, normalisable eigenfunctions have $\nu \geq 0$ for open and flat models ($K \leq 0$). In Euclidean space, this implies all $k^2 \geq 0$. The $k = 0$ solutions are homogeneous and, therefore, do not appear in the expansion of first-order tensors, for example $\Delta_a \equiv a D_a \rho / \rho$. In open models, the modes with $\nu \geq 0$ form a complete set for expanding square-integrable functions, but they necessarily have $k \geq \sqrt{|K|}$ and so cannot describe correlations longer than the curvature scale [71]. Super-curvature solutions (with $-1 < \nu^2 < 0$) can be constructed by analytic continuation

and Ref. [71] argues that these should be included in an expansion of a general random field. A super-curvature mode is generated in some models of open inflation [222]. In closed models ν is restricted to integer values ≥ 1 [220,221] and there are ν^2 linearly-independent modes per ν . The mode with $\nu = 1$ cannot be used to construct perturbations (its projected gradient vanishes globally), while the modes with $\nu = 2$ can only describe perturbations where all perturbed tensors with rank > 1 vanish [103].

For the scalar contribution to a rank- l tensor, such as the l -th multipole of the distribution function F_{A_l} , we expand in rank- l PSTF tensors $\mathcal{Q}_{A_l}^{(0)}$ derived from the $\mathcal{Q}^{(0)}$ via [210,223]

$$\mathcal{Q}_{A_l}^{(0)} = \left(\frac{-a}{k}\right)^l D_{(a_1} \dots D_{a_l)} \mathcal{Q}^{(0)}. \quad (4.1.31)$$

The recursion relation for the $\mathcal{Q}_{A_l}^{(0)}$,

$$\mathcal{Q}_{A_l}^{(0)} = -\frac{a}{k} D_{(a_l} \mathcal{Q}_{A_{l-1}}^{(0)}, \quad (4.1.32)$$

follows directly. The factor a^l in the definition of the $\mathcal{Q}_{A_l}^{(0)}$ ensures that $\dot{\mathcal{Q}}_{A_l}^{(0)} = 0$ at zero-order. The multipole equation (4.1.27) also involves the divergence of $\mathcal{Q}_{A_l}^{(0)}$, for which we need the result [114,210]

$$D^{a_l} \mathcal{Q}_{A_l}^{(0)} = \frac{k}{a} \frac{l}{(2l-1)} \left[1 - (l^2 - 1) \frac{K}{k^2}\right] \mathcal{Q}_{A_{l-1}}^{(0)}. \quad (4.1.33)$$

When we discuss CMB polarization in Section 4.3, we shall also require the result that [224]

$$\text{curl } \mathcal{Q}_{A_l}^{(0)} = 0, \quad (4.1.34)$$

where the curl of a general rank- l PSTF tensor is defined by

$$\text{curl } S_{A_l} = \varepsilon_{bc(a_l} D^b S_{A_{l-1})}{}^c. \quad (4.1.35)$$

In closed models, the $\mathcal{Q}_{A_l}^{(0)}$ vanish for $l \geq \nu$, so only modes with $\nu > l$ contribute to rank- l tensors.

The decomposition of the distribution function into angular multipoles F_{A_l} , and the subsequent expansion in the $\mathcal{Q}_{A_l}^{(0)}$, combine to give a normal mode expansion which involves the objects $D_{(A_l)} \mathcal{Q}^{(0)} e^{A_l}$. For $K = 0$, with the $\mathcal{Q}^{(0)}$ taken to be Fourier modes, this is equivalent to the usual Legendre expansion $P_l(\hat{\mathbf{k}} \cdot \mathbf{e})$ where $\hat{\mathbf{k}}$ is the Fourier wavevector (e.g. Ref. [217]). In non-flat models, the expansion is equivalent to the Legendre tensor approach, first introduced by Wilson [225]. The advantage of handling the angular and scalar harmonic decompositions separately is that the former can be applied quite generally for an arbitrary cosmological model. Furthermore, extending the normal-mode expansions to cover polarization and vector and tensor modes in non-flat models is then rather trivial.

Vector perturbations. For these we use the PSTF rank-1 eigenfunctions of the Laplacian,

$$a^2 D^2 \mathcal{Q}_a^{(\pm 1)} + k^2 \mathcal{Q}_a^{(\pm 1)} = 0, \quad (4.1.36)$$

that are divergence-free, $D^a \mathcal{Q}_a^{(\pm 1)} = 0$, and have vanishing time derivative, $\dot{\mathcal{Q}}_a^{(\pm 1)} = 0$. The superscript (± 1) labels the two possible parities (electric and magnetic) of the vector harmonics, e.g. Ref. [220]. We can always choose the parity states so that

$$\text{curl } \mathcal{Q}_a^{(\pm 1)} = \frac{k}{a} \sqrt{1 + \frac{2K}{k^2}} \mathcal{Q}_a^{(\mp 1)}, \quad (4.1.37)$$

which ensures that both parities have the same normalisation. For vector modes we define $\nu^2 = (k^2 + 2K)/|K|$. The regular, normalisable eigenmodes have $\nu \geq 0$ for flat and open models, while for closed models ν is an integer ≥ 2 .

We now differentiate the $\mathcal{Q}_a^{(\pm 1)}$ vectors $l - 1$ times to form PSTF tensors:

$$\mathcal{Q}_{A_l}^{(\pm 1)} \equiv \left(\frac{-a}{k}\right)^{l-1} D_{(A_{l-1}} \mathcal{Q}_{a_l)}^{(\pm 1)}, \quad (4.1.38)$$

which are constant in time. They satisfy the same recursion relation (4.1.32) as the scalar harmonics. For the projected divergences, we have [226]

$$D^{a_l} \mathcal{Q}_{A_l}^{(\pm 1)} = \frac{k}{a} \frac{(l^2 - 1)}{l(2l - 1)} \left[1 - (l^2 - 2) \frac{K}{k^2}\right] \mathcal{Q}_{A_{l-1}}^{(\pm 1)}. \quad (4.1.39)$$

Finally, for polarization we shall require the result

$$\text{curl } \mathcal{Q}_{A_l}^{(\pm 1)} = \frac{1}{l} \frac{k}{a} \sqrt{1 + \frac{2K}{k^2}} \mathcal{Q}_{A_l}^{(\mp 1)}. \quad (4.1.40)$$

As for scalar perturbations, the $\mathcal{Q}_{A_l}^{(\pm 1)}$ vanish for $l \geq \nu$ in closed models.

Tensor perturbations. Here we use the PSTF rank-2 eigenfunctions of the Laplacian,

$$a^2 D^2 \mathcal{Q}_{ab}^{(\pm 2)} + k^2 \mathcal{Q}_{ab}^{(\pm 2)} = 0, \quad (4.1.41)$$

that are transverse, $D^b \mathcal{Q}_{ab}^{(\pm 2)} = 0$, and have vanishing time derivative, $\dot{\mathcal{Q}}_{ab}^{(\pm 2)} = 0$. The superscript (± 2) labels the two possible parity states for the tensor harmonics [169,220,227]. For our purposes, the states can be conveniently chosen so that

$$\text{curl } \mathcal{Q}_{ab}^{(\pm 2)} = \frac{k}{a} \sqrt{1 + \frac{3K}{k^2}} \mathcal{Q}_{ab}^{(\mp 2)}. \quad (4.1.42)$$

For tensor modes we define $\nu^2 = (k^2 + 3K)/|K|$. The regular, normalisable eigenmodes have $\nu \geq 0$ for flat and open models, while for closed models ν is an integer ≥ 3 .

Following our treatment of scalar perturbations, we form rank- l PSTF tensors $\mathcal{Q}_{A_l}^{(\pm 2)}$ by differentiation:

$$\mathcal{Q}_{A_l}^{(\pm 2)} \equiv \left(\frac{-a}{k}\right)^{l-2} D_{(A_{l-2}} \mathcal{Q}_{a_{l-1}a_l}^{(\pm 2)}. \quad (4.1.43)$$

The $\mathcal{Q}_{A_l}^{(\pm 2)}$ satisfy the same recursion relation (4.1.32) as the scalar harmonics but their projected divergences are [169]

$$D^{a_l} \mathcal{Q}_{A_l}^{(\pm 2)} = \frac{k}{a} \frac{(l^2 - 4)}{l(2l - 1)} \left[1 - (l^2 - 3) \frac{K}{k^2}\right] \mathcal{Q}_{A_{l-1}}^{(\pm 2)}. \quad (4.1.44)$$

Finally, for polarization we require

$$\text{curl } \mathcal{Q}_{A_l}^{(\pm 2)} = \frac{2k}{l} \frac{1}{a} \sqrt{1 + \frac{3K}{k^2}} \mathcal{Q}_{A_l}^{(\mp 2)}. \quad (4.1.45)$$

As before, in closed models, the $\mathcal{Q}_{A_l}^{(\pm 2)}$ vanish for $l \geq \nu$.

Combining the angular and spatial expansions gives a set of normal-mode functions going like $D_{(A_{l-2}} \mathcal{Q}_{a_{l-1}a_l}^{(\pm 2)} e^{A_l}$ for tensor modes. This generalises Wilson's approach [225] for scalar perturbations to tensor modes. For the special case of a flat universe, and working with circularly-polarized Fourier modes,²³ the normal modes reduce to $m = \pm 2$ spherical harmonics, $Y_{l\pm 2}(\mathbf{e})$, when the Fourier wavevector lies along the z -axis. Similarly, for vector modes $D_{(A_{l-1}} \mathcal{Q}_{a_l}^{(\pm 1)} e^{A_l}$ reduces to $m = \pm 1$ spherical harmonics, $Y_{l\pm 1}(\mathbf{e})$. In this limit, the normal-mode expansions are equivalent to those in the total angular momentum method of Hu & White [228].

Quite generally, the normal-mode functions for constructing anisotropy from rank- m perturbations are $\mathcal{Q}_{A_l}^{(m)} e^{A_l}$. The evolution of these quantities along the line of sight from some point R is governed by the recursion [225,229]

$$\frac{1}{k} \frac{d}{d\chi} \mathcal{Q}_{A_l}^{(m)} e^{A_l} = \mathcal{Q}_{A_{l+1}}^{(m)} e^{A_{l+1}} - \frac{(l^2 - m^2)}{(2l + 1)(2l - 1)} \left[1 - (l^2 - m - 1) \frac{K}{k^2}\right] \mathcal{Q}_{A_{l-1}}^{(m)} e^{A_{l-1}}, \quad (4.1.46)$$

where we have used the recursion relation (4.1.32) and the divergence result (4.1.33) suitably generalised to rank- m perturbations. Here, χ is the comoving radial distance along the line of sight, so $\chi = 0$ at R . The coupling of the l th normal mode to $l \pm 1$ induces a similar coupling between the anisotropy at multipoles l and $l \pm 1$ (c.f. the advective coupling in Eq. (4.1.8)). The solution for the $l = m$ normal mode along the line of sight can be written in terms of the $l \geq m$ modes at R as

$$\mathcal{Q}_{A_m}^{(m)} e^{A_m} \Big|_{\chi} = 4\pi \sum_{l=m}^{\infty} \frac{1}{\Delta_l} \frac{l!}{(l-m)!} \frac{\nu(\nu^2 + m + 1)^{(l-m)/2}}{\prod_{n=0}^l \sqrt{\nu^2 + n^2}} \frac{\Phi_l^\nu(x)}{\sinh^m x} \left(\mathcal{Q}_{A_l}^{(m)} e^{A_l}\right)_{\chi=0}, \quad (4.1.47)$$

where $\nu^2 = [k^2 + (m+1)K]/|K|$, $x = \sqrt{|K|}\chi$ and the $\Phi_l^\nu(x)$ are the ultra-spherical Bessel functions (see e.g. Refs. [221]).²⁴ They are the generalisation of spherical Bessel functions to $K \neq 0$ spaces – $\Phi_l^\nu(x) \rightarrow j_l(k\chi)$ in the limit of a flat model – and we see that $\Phi_l^\nu(x)/\sinh^m x$ give the radial dependence of the rank- m eigenfunctions $\mathcal{Q}_{A_m}^{(m)}$.²⁵ We have written Eq. (4.1.47) in a form appropriate for $K < 0$; for closed models $\nu^2 + n$ should be replaced by $\nu^2 - n$, where n is an integer, and the hyperbolic functions by their trigonometric counterparts. For $K > 0$ the sum over l truncates at $l = \nu - 1$. Note that Eq. (4.1.47) is independent of the specific representation of the $\mathcal{Q}_{A_m}^{(m)}$ that we choose. Physically, it describes the linearised anisotropy pattern generated at R from a first-order source in the Boltzmann equation of the form $\mathcal{Q}_{A_m}^{(m)} e^{A_m}$ acting at radial distance χ . For a given rank- m perturbation, the response to sources with $l > m$ local angular dependence, i.e. of the form $\mathcal{Q}_{A_l}^{(m)} e^{A_l}$, follows from differentiating Eq. (4.1.47) and using the recursion (4.1.46). In FLRW models, the sources for the scalar distribution function are always of this form. We shall see meet examples of these projections in the integral solutions for the CMB anisotropies in Sections 4.3.4–4.3.6.

²³ Note that the (complex) Fourier modes are not parity states and the curl relations (4.1.37) and (4.1.42) do not hold directly.

²⁴ Note that $\nu x \rightarrow k\chi$ in the limit $K \rightarrow 0$.

²⁵ This agrees with the explicit constructions in Refs. [220,221] for the $m \leq 2$ cases.

4.2. Cosmic neutrino background

In this subsection we discuss briefly the perturbations of the cosmic neutrino background from the perspective of the $1 + 3$ -covariant kinetic theory developed above.

Flavour oscillations imply that at least two of the three neutrino mass eigenstates have non-zero masses. The inferred squared-mass differences $\Delta m_{21}^2 = 7.9 \times 10^{-5} \text{ eV}^2$ and $|\Delta m_{31}^2| = 2.2 \times 10^{-3} \text{ eV}^2$ imply that the summed masses $\sum_i m_i \geq 0.056 \text{ eV}$ for the normal hierarchy ($\Delta m_{31}^2 > 0$) and $\sum_i m_i \geq 0.095 \text{ eV}$ in the inverted hierarchy. All other things being equal, the effect of non-zero neutrino masses is to suppress the matter perturbations on scales below the neutrino Jeans scale (see e.g. Ref. [230] for a recent review and Refs. [231], [232] for early pioneering work). Combining their three-year data with large-scale structure data, the WMAP team find $\sum_i m_i < 0.9 \text{ eV}$ (95% confidence) via this route in a flat universe [1]. Cosmology therefore places an important constraint on the absolute neutrino mass scale. Better constraints can be obtained by including smaller scale measurements of the matter power spectrum from the (Lyman- α) absorption spectra of distant quasars [233,234], although there are some apparent inconsistencies between the CMB and Lyman- α data that may invalidate the conclusions on neutrino masses.

Neutrinos decouple at temperatures $\sim 1 \text{ MeV}$ when they are ultra-relativistic. Since they were in thermal equilibrium before this time, the distribution function in the FLRW limit is

$$f(q) \approx \left[\exp\left(\frac{q}{k_B T_d a_d}\right) + 1 \right]^{-1}, \quad (4.2.1)$$

where we have set the chemical potential to zero. Here, T_d is the temperature at neutrino decoupling when the scale factor is a_d . Note that when expressed in terms of the comoving momentum, the distribution function is time independent. Once the neutrinos become non-relativistic (typical momentum $\lambda \ll m$), the distribution function starts to depart from Fermi-Dirac form since it is the momentum not energy that redshifts with the expansion of the universe. Defining the neutrino temperature $T_\nu \equiv a_d T_d / a$, at redshift zero T_ν is related to the CMB temperature by $T_\nu = (4/11)^{1/3} T_{\text{CMB}} = 1.96 \text{ K}$; the CMB temperature is higher due to photon production at electron-positron annihilation.

The current neutrino temperature and the inferred mass differences imply that at least two of the mass eigenstates are non-relativistic at the present epoch. Noting that the average momentum

$$\langle \lambda \rangle = \frac{7\pi^4}{180\zeta(3)} k_B T_\nu \approx 3.15 k_B T_\nu, \quad (4.2.2)$$

the current upper limit of 0.3 eV for the mass of any eigenstate implies that neutrinos were relativistic at the time of hydrogen recombination. For this reason, light neutrino masses only affect the CMB indirectly through changes to the angular-diameter distance to last scattering. Their effect on the growth of small-scale matter perturbations is significant though [235] and is the mechanism by which cosmological observations constrain the (summed) absolute neutrino mass.

In the following subsections we discuss the dynamics of scalar and tensor-mode neutrino perturbations. We include the latter since it has the potentially observable consequence of damping sub-horizon gravitational waves during radiation domination [236].

4.2.1. Scalar perturbations

For scalar perturbations we expand the multipoles $F_{A_l}(q)$ in terms of the $\mathcal{Q}_{A_l}^{(0)}$:

$$F_{A_l}(q) = -\frac{\pi}{\Delta_l} \frac{dF(q)}{d \ln q} \sum_k F_l^{(0)}(q) \mathcal{Q}_{A_l}^{(0)}, \quad l \geq 1, \quad (4.2.3)$$

where $F(q)$ is the zero-order monopole of the distribution function, Eq. (4.2.1). Here, we have left the harmonic index (k) implicit on the scalar harmonic functions and their coefficients. The momentum-dependent prefactor is chosen so that for massless particles the $F_l^{(0)}$ are independent of q (see Eq. (4.2.8)) and

$$q_a = \rho \sum_k F_1^{(0)} \mathcal{Q}_a^{(0)}, \quad (m = 0), \quad (4.2.4)$$

$$\pi_{ab} = \rho \sum_k F_2^{(0)} \mathcal{Q}_{ab}^{(0)}, \quad (m = 0). \quad (4.2.5)$$

In the massless case, $\sum_k F_l^{(0)} \mathcal{Q}_{A_l}^{(0)}$ are proportional to the multipoles of the neutrino temperature anisotropy. For the gradient of the monopole we use

$$\mathcal{V}_a(q) = \frac{k}{4} \frac{dF(q)}{d \ln q} \sum_k F_0^{(0)}(q) \mathcal{Q}_a^{(0)}, \quad (4.2.6)$$

so that, on integrating over q ,

$$\Delta_a = \frac{a}{\rho} D_a \rho = - \sum_k k F_0^{(0)} \mathcal{Q}_a^{(0)}, \quad (m = 0), \quad (4.2.7)$$

and hence $\sum_k F_0^{(0)} \mathcal{Q}^{(0)}$ is essentially $\delta\rho/\rho$ on hypersurfaces orthogonal to u^a (up to a constant).

After they decouple, the neutrino multipoles satisfy Eqs. (4.1.27) and (4.1.29) with vanishing collision terms. Expanding in harmonics, we find [211]

$$\begin{aligned} \dot{F}_l^{(0)}(q) + \frac{k}{a\epsilon} \left\{ \frac{l+1}{2l+1} \left[1 - \frac{K}{k^2} \right] F_{l+1}^{(0)}(q) - \frac{l}{2l+1} F_{l-1}^{(0)}(q) \right\} \\ + 4\delta_{l0}\dot{h} + \delta_{l1} \frac{4k}{3a} \left(\frac{q}{\epsilon} h + \frac{\epsilon}{q} A \right) + \delta_{l2} \frac{8k}{15a} \sigma = 0, \end{aligned} \quad (4.2.8)$$

where the kinematic quantities are

$$h_a = - \sum_k kh \mathcal{Q}_a^{(0)}, \quad A_a = \sum_k \frac{k}{a} A \mathcal{Q}_a^{(0)}, \quad \sigma_{ab} = \sum_k \frac{k}{a} \sigma \mathcal{Q}_{ab}^{(0)}. \quad (4.2.9)$$

When most of the neutrinos are relativistic, we can set $q = \epsilon$ in Eq. (4.2.8) and we recover massless dynamics. On sub-horizon scales, neutrinos are unable to cluster and free-streaming excites multipoles higher than $l = 1$. At late times, for those species that are non-relativistic, free-streaming effectively turns off for $k/a \ll H\epsilon/q$ which defines the neutrino Jeans length $\lambda_J \sim v_*/H$ where v_* is the typical thermal velocity. When free-streaming is not operating, the $F_l^{(0)}(q)$ are constant for $l > 2$. During matter domination, the comoving Jeans length falls (as $a^{-1/2}$) so that neutrinos start to cluster again at late times for modes that are sub-Hubble at the non-relativistic transition once their thermal velocities have redshifted sufficiently. For modes that are inside the Jeans length today, the clustered baryons and CDM have never felt the gravity of clustered neutrinos. All other things being equal, increasing the neutrino mass increases the expansion rate by a constant factor when neutrinos are very non-relativistic, and so structure formation is slowed down on small scales and the matter power spectrum is suppressed by a constant factor proportional to the summed neutrino masses [237]. There is no such effect on large scales since the enhanced expansion rate is mitigated by neutrino clustering. A comprehensive review of neutrino scalar perturbations is given in Ref. [230].

We have focussed on the effects of massive neutrinos on matter clustering. However, if in the future it were possible to detect the cosmic neutrino background directly, its anisotropies would open up a rich new source of cosmological information. The power spectra of the anisotropies can easily be computed from the solution of Eq. (4.2.8) using the methods described for the CMB in Section 4.3; see Ref. [238] for further details.

4.2.2. Tensor perturbations

For tensor perturbations, we expand the $F_{A_l}(q)$ as

$$F_{A_l}(q) = - \frac{\pi}{\Delta l} \frac{dF(q)}{d \ln q} \sum_k F_l^{(\pm 2)}(q) \mathcal{Q}_{A_l}^{(\pm 2)}, \quad l \geq 2, \quad (4.2.10)$$

where the sum over modes \sum_k includes the two parity states labelled with a superscript (± 2) . Expanding Eq. (4.1.27) in harmonics gives [211]

$$\dot{F}_l^{(\pm 2)}(q) + \frac{k}{a\epsilon} \left\{ \frac{(l+3)(l-1)}{(2l+1)(l+1)} \left[1 - \frac{K}{k^2} \right] F_{l+1}^{(\pm 2)}(q) - \frac{l}{2l+1} F_{l-1}^{(\pm 2)}(q) \right\} + \delta_{l2} \frac{8k}{15a} \sigma^{(\pm 2)} = 0, \quad (4.2.11)$$

where the shear $\sigma_{ab} = \sum_k (k/a) \sigma^{(\pm 2)} \mathcal{Q}_{ab}^{(\pm 2)}$. The physics of Eq. (4.2.11) is straightforward: the anisotropic expansion due to the shear of gravitational waves continually sources quadrupole anisotropy; advection moves this power up to higher l with $\Delta l = 1$ taking the typical thermal crossing time across the wavelength of the perturbation. For perturbations outside the neutrino Jeans scale, advection is ineffective within an expansion time and the $l > 2$ multipoles become approximately constant.

An important application of Eq. (4.2.11) is the damping of gravitational waves by neutrino free-streaming [236,239]. The evolution of cosmological gravitational waves was considered in Section 3.6 in terms of the shear. However, to make contact with the metric-based literature, it is convenient to work with the transverse, trace-free metric perturbation $\mathcal{H}_{ab} = a^2 \mathcal{H}_{ij}(\mathbf{dx}^i)_a(\mathbf{dx}^j)_b$, where

$$ds^2 = a^2 \left[-d\eta^2 + (\gamma_{ij} + \mathcal{H}_{ij}) dx^i dx^j \right] \quad (4.2.12)$$

and γ_{ij} the background (conformal) spatial metric. Note that \mathcal{H}_{ij} is gauge-invariant [103]. In terms of the metric, the shear and the electric part of the Weyl tensor are (e.g. Ref. [107])

$$\sigma_{ab} = \frac{1}{2} \dot{\mathcal{H}}_{ab} \quad (4.2.13)$$

$$E_{ab} = - \frac{1}{4} \frac{1}{a^2} \mathcal{H}_{ab}'' + \frac{K}{2a^2} \mathcal{H}_{ab} - \frac{1}{4} D^2 \mathcal{H}_{ab}, \quad (4.2.14)$$

where primes denote the action of $au^a \nabla_a$. The equation of motion for \mathcal{H}_{ab} follows from the shear propagation equation; we find

$$\mathcal{H}_{ab}'' + 2 \frac{a'}{a} \mathcal{H}_{ab}' - (a^2 D^2 - 2K) \mathcal{H}_{ab} = 2a^2 \pi_{ab}. \quad (4.2.15)$$

After neutrino decoupling, but during radiation domination, the neutrino anisotropic stress is dynamically important for sub-Hubble scales. Before decoupling, scattering keeps the neutrinos isotropic, while after matter-radiation equality, the role of the neutrino stress is suppressed by the ratio f_ν of neutrino energy density to the total energy density. During radiation domination, $f_\nu \approx 0.405$ is constant and the harmonic expansion of \mathcal{H}_{ab} evolves with conformal time as

$$\mathcal{H}^{(\pm 2)''} + \frac{2}{\eta} \mathcal{H}^{(\pm 2)'} + (k^2 + 2K) \mathcal{H}^{(\pm 2)} = \frac{6f_\nu}{\eta^2} F_2^{(\pm 2)} \quad (4.2.16)$$

for massless neutrinos. (We have used the tensor-mode version of Eq. (4.2.5) to relate $F_2^{(\pm 2)}$ to the neutrino anisotropic stress.) In the absence of anisotropic stress, the regular solution of Eq. (4.2.16) is $\mathcal{H}^{(\pm 2)} \propto j_0(k\eta)$ in a flat universe. The metric perturbation is thus constant outside the horizon but then decays as a after horizon crossing.

Quite generally, Eqs. (4.2.11) and (4.2.15) can be recast as an integro-differential using the following integral solution for the $F_i^{(\pm 2)}$ in the massless limit in a flat universe,

$$F_i^{(\pm 2)} = -4i(l-1) \int_0^\eta d\eta' k \sigma^{(\pm 2)}(\eta') \frac{j_l(k\Delta\eta)}{(k\Delta\eta)^2}, \quad (4.2.17)$$

where $\Delta\eta \equiv \eta - \eta'$.²⁶ The lower limit of integration should be the end of neutrino decoupling, but it is harmless to approximate this as $\eta = 0$. Combining the quadrupole solution with the harmonic expansion of Eq. (4.2.16) for $K = 0$, gives [236]

$$\mathcal{H}^{(\pm 2)''} + \frac{2a'}{a} \mathcal{H}^{(\pm 2)'} + k^2 \mathcal{H}^{(\pm 2)} = -24f_\nu(\eta) \left(\frac{a'}{a}\right)^2 \int_0^\eta d\eta' \mathcal{H}^{(\pm 2)'}(\eta') \frac{j_2(k\Delta\eta)}{(k\Delta\eta)^2}, \quad (4.2.18)$$

where we have used $\mathcal{H}^{(\pm 2)'} = 2k\sigma^{(\pm 2)}$. Modes that enter the horizon well before matter-radiation equality, but well after neutrino decoupling, have their amplitude damped by a factor ≈ 0.803 with essentially no phase shift, i.e. $\mathcal{H}^{(\pm 2)}(\eta) \approx 0.803 \mathcal{H}^{(\pm 2)}(0) j_0(k\eta)$ during radiation domination [236]. However, the observational consequences of such modes is limited since their wavelengths are too short to affect the CMB and too long for direct detection with laser interferometers. CMB polarization is dominated by modes entering the horizon around recombination and the universe is not fully matter-dominated at this epoch. The asymptotic results of Ref. [236] are not applicable in this limit and a full numerical solution of Eq. (4.2.18) is required for accurate results. Both the phase shift and damping of the amplitude turn out to be important and their interplay can both enhance or reduce the polarization power by a few percent depending on scale.²⁷

4.3. Cosmic microwave background

Since the detection of anisotropies in the temperature of CMB radiation by COBE [240], the CMB has played a major role in establishing quantitative constraints on the cosmological model. The small $O(10^{-5})$ amplitude of these fluctuations means they are well described by linear perturbation theory and the physics of the CMB is thus very well understood. For detailed recent reviews, see Refs. [241–244].

The prediction of angular variations in the temperature of the radiation, due to the propagation of photons through an inhomogeneous universe [102], followed shortly after the (definitive) discovery of the CMB in 1965 by Penzias and Wilson [245]. Shortly after, polarization in the CMB was predicted in models with anisotropy in the expansion around the time of recombination [246]. The detailed physics of CMB fluctuations in almost-FLRW models was essentially understood by the early 1970s [146,247–250] for models with only baryonic matter; cold-dark-matter models were considered a decade later [151,251]. Further important developments included the effect of spatial curvature [225], polarization [252,251] and gravitational waves [253,254]. All of these works used the standard metric-based approach to cosmological perturbation theory, but the physics of the CMB has also been studied extensively in the $1+3$ -covariant approach [114,169,210,215,223,224,226,255–263]. This brings to the CMB the benefits described in Section 2, and in particular: (i) clarity in the definition of the variables employed; (ii) covariant and gauge-invariant perturbation theory around a variety of background models; (iii) provision of a sound basis for studying non-linear effects; and (iv) freedom to employ any coordinate system or tetrad.

In this subsection our main focus is on the linear theory of CMB anisotropies, but we present most of the basic $1+3$ -covariant framework in a non-perturbative manner. The approach therefore provides a convenient starting point for non-linear treatments of CMB fluctuations [259]. There are several non-linear effects that are expected to affect the CMB at an important level on scales below a few arcminutes. These include gravitational lensing [264–267] (see, also, Ref. [268] for a comprehensive review), various scattering effects either during [269] or after the universe reionized [270–272] (see e.g. Ref. [273] for a review) and gravitational redshifting effects [274]. For polarization, the gravitational lensing effect may even be the dominant contribution to the B -mode component [275,276] on all angular scales [277]. A complete non-linear computation of the CMB anisotropies in almost-FLRW models is still lacking, but there has been some progress made recently at second-order in perturbation theory [278,279]. (However, note that non-perturbative effects are known to be important for gravitational lensing of the CMB [280]).

²⁶ Unsurprisingly, the integral solution for massless neutrinos is the same as for the CMB in the absence of scattering; see Section 4.3.4.

²⁷ For further discussion of this point, see the unpublished notes by Antony Lewis that accompany the CAMB code; they are available at <http://cosmologist.info/notes/CAMB.ps.gz>.

4.3.1. CMB observables

We are interested in both the total intensity and polarization properties of the CMB. Therefore, we describe the CMB photons by a one-particle distribution function that is tensor-valued: $f_{bc}(x^a, p^a)$ [281]. It is a Hermitian tensor defined so that the expected number of photons contained in a proper phase-space element $d^3\mathbf{x}d^3\mathbf{p}$, and with polarization state ϵ^a is $\epsilon^{a*}f_{ab}\epsilon^b d^3\mathbf{x}d^3\mathbf{p}$. The complex polarization 4-vector ϵ^a is orthogonal to the photon momentum, $\epsilon^a p_a = 0$ (adopting the Lorentz gauge), and is normalised as $\epsilon_a^* \epsilon^a = 1$. The distribution function is also defined to be orthogonal to p^a so $f_{ab}p^a = 0$. We can make a 1 + 3-covariant decomposition of the photon 4-momentum as

$$p^a = E(u^a + e^a), \quad (4.3.1)$$

where, now, the magnitude of the momentum λ equals the energy E since photons are massless. For a photon in a pure polarization state ϵ^a , the direction of the electric field relative to u^a is $s^a_b \epsilon^b$ where, recall, $s_{ab} \equiv h_{ab} - e_a e_b$ is the screen projection tensor.

The (Lorentz-gauge) polarization 4-vector is only unique up to constant multiples of p^a , reflecting the remaining gauge freedom, but the *observed* polarization vector $s^a_b \epsilon^b$ is unique. As the residual gauge freedom also affects the distribution function f_{ab} , it is sometimes convenient to work directly with the (screen-)projected polarization tensor,

$$P_{ab} \propto E^3 s_a^c s_b^d f_{cd} \quad (4.3.2)$$

which governs the observable properties of the radiation field from the perspective of u^a . It is unaffected by the residual electromagnetic gauge freedom. The factor E^3 is included in the definition of P_{ab} so that it relates simply to the observed Stokes brightness parameters for the radiation field. Decomposing P_{ab} into its irreducible components,

$$P_{ab}(E, e^d) = \frac{1}{2}I(E, e^d)s_{ab} + \mathcal{P}_{ab}(E, e^d) + \frac{1}{2}iV(E, e^d)\varepsilon_{abc}e^c, \quad (4.3.3)$$

defines the total intensity brightness, I , the circular polarization, V , and the linear polarization tensor \mathcal{P}_{ab} which is PSTF and transverse to e^a . The projected polarization tensor can, alternatively, be interpreted in terms of classical electromagnetic fields: for quasi-monochromatic radiation with electric field $\Re[E^a(t) \exp(-i\omega t)]$, where ω is the angular frequency and the complex representative E^a varies little over a wave period,

$$P^{ab} \propto \langle E^a E^{b*} \rangle. \quad (4.3.4)$$

Here, the angle brackets denote time averaging. The linear polarization is often described in terms of Stokes brightness parameters Q and U (e.g. Ref. [282]) which, operationally, measure the difference in intensity between radiation transmitted by a pair of orthogonal polarizers (for Q), and the same but after a right-handed rotation of the polarizers by 45 degrees about the propagation direction e^a (for U). If we introduce a pair of orthogonal polarization vectors $(e_1)^a$ and $(e_2)^a$, which are perpendicular to u^a and e^a , i.e. $s_b^a(e_i)^b = (e_i)^a$ for $i = 1$ and 2 , and are oriented so that $\{u^a, (e_1)^a, (e_2)^a, e^a\}$ form a right-handed orthonormal tetrad, we have

$$\mathcal{P}_{ab}(e_i)^a(e_j)^b = \frac{1}{2} \begin{pmatrix} Q & U \\ U & -Q \end{pmatrix}. \quad (4.3.5)$$

The invariant $2\mathcal{P}^{ab}\mathcal{P}_{ab} = Q^2 + U^2$ is the magnitude (squared) of the linear polarization.

Since $I(E, e^c)$ and $V(E, e^c)$ are scalar functions on the sphere $e^a e_a = 1$ at a point in spacetime, their local angular dependence can be handled by an expansion in PSTF tensor-valued multipoles, as in Eq. (4.1.5):

$$I(E, e^c) = \sum_{l=0}^{\infty} I_{A_l}(E) e^{A_l}, \quad (4.3.6)$$

$$V(E, e^c) = \sum_{l=0}^{\infty} V_{A_l}(E) e^{A_l}. \quad (4.3.7)$$

For \mathcal{P}_{ab} , we use the fact that any STF tensor on the sphere can be written in terms of angular derivatives of two scalar potentials, P_E and P_B , as (e.g. Ref. [275])

$$\mathcal{P}_{ab} = {}^{(2)}\nabla_{(a} {}^{(2)}\nabla_{b)} P_E + \epsilon^c_{(a} {}^{(2)}\nabla_{b)} {}^{(2)}\nabla_c P_B, \quad (4.3.8)$$

where ${}^{(2)}\nabla_a$ and $\epsilon_{ab} = \varepsilon_{abc}e^c$ are the covariant derivative and alternating tensor on the two-sphere. The scalar fields P_E and P_B are even and odd under parity respectively, and define the electric and magnetic parts of the linear polarization. Expanding P_E and P_B in PSTF multipoles in the usual way, and evaluating the angular derivatives, we can write [224,227]

$$\mathcal{P}_{ab}(E, e^c) = \sum_{l=2}^{\infty} [\varepsilon_{abc_{l-2}}(E) e^{c_{l-2}}]^{TT} - \sum_{l=2}^{\infty} [e_{d_1} \varepsilon^{d_1 d_2} {}_{(a} \mathcal{B}_{b)d_2 c_{l-2}}(E) e^{c_{l-2}}]^{TT}. \quad (4.3.9)$$

Here, TT denotes the transverse (to e^a), trace-free part: for a general second-rank tensor F_{ab}

$$[F_{ab}]^{TT} = s_a^c s_b^d F_{cd} - \frac{1}{2} s_{ab} s^{cd} F_{cd}. \quad (4.3.10)$$

Eq. (4.3.9) can be inverted to determine the PSTF tensors \mathcal{E}_{A_l} and \mathcal{B}_{A_l} as

$$\mathcal{E}_{A_l}(E) = M_l^2 \Delta_l^{-1} \int d\Omega e_{(A_{l-2}} \mathcal{P}_{a_{l-1}a_l)}(E, e^c), \quad (4.3.11)$$

$$\mathcal{B}_{A_l}(E) = M_l^2 \Delta_l^{-1} \int d\Omega e_b \epsilon^{bd}{}_{(a_l} e_{A_{l-2}} \mathcal{P}_{a_{l-1})d}(E, e^c), \quad (4.3.12)$$

where $M_l \equiv \sqrt{2l(l-1)/[(l+1)(l+2)]}$. The multipole expansion in Eq. (4.3.9) is the coordinate-free version of the tensor spherical harmonic expansion introduced to the analysis of CMB polarization in Ref. [275]. An alternative expansion, whereby $Q \pm iU$ is expanded in spin-weighted spherical harmonics (e.g. Ref. [283]), is also commonly employed [276]. The expansion (4.3.9) is also equivalent to that introduced in the 1970s by Dautcourt & Rose [284].

In observational cosmology, the CMB anisotropy is usually expressed in terms of thermodynamic equivalent temperature, i.e. the distribution function $I(E, e^a)/E^3$ divided by $\partial f_{\text{Pl}}/\partial T_{\text{CMB}}$, where $f_{\text{Pl}}(E/T_{\text{CMB}})$ is the Planck function at the average CMB temperature.²⁸ Ignoring spectral distortions, the linear-theory CMB *anisotropy* and polarization brightness are independent of energy when expressed as thermodynamic temperatures. This is because the *linear* perturbations in f_{ab} inherit the spectral dependence of $\epsilon \partial f_{\text{Pl}}/\partial \epsilon$.²⁹ For the CMB, it follows that we can integrate over energy without loss of information so we define bolometric multipoles

$$I_{A_l} = \Delta_l \int_0^\infty dE I_{A_l}(E) \quad (4.3.13)$$

for $l \geq 0$. The normalisation is chosen so that the three lowest multipoles give the radiation energy and momentum densities and anisotropic stress respectively (c.f. Eqs. (4.1.10)–(4.1.12)):

$$I = \rho, \quad I_a = q_a, \quad I_{ab} = \pi_{ab}^{(\gamma)}. \quad (4.3.14)$$

We define \mathcal{E}_{A_l} , \mathcal{B}_{A_l} and V_{A_l} similarly. The fractional anisotropy in the CMB temperature, $\delta_T(e^a)$, is then related to the I_{A_l} by

$$\delta_T(e^c) = \frac{\pi}{I} \int_0^\infty dE I'(E, e^a) \quad (4.3.15)$$

$$= \frac{\pi}{I} \sum_{l=1}^\infty \Delta_l^{-1} I_{A_l} e^{A_l} \quad (4.3.16)$$

to first order, where $I'(E, e^a)$ is the brightness anisotropy.

For theories, such as single-field inflation, that predict initial perturbations that are very close to being Gaussian distributed (see Ref. [286] for a recent review), the CMB fluctuations should also be Gaussian distributed where linear theory applies. If we further assume that the statistical properties of the fluctuations are invariant under the isometries of the background cosmological model (i.e. translations and rotations for FLRW), the CMB power spectra fully characterise the statistics of the CMB anisotropies and polarization.³⁰ The temperature power spectrum is defined in terms of the I_{A_l} by [210]:

$$\left(\frac{\pi}{I}\right)^2 \langle I_{A_l} I^{B_l} \rangle = \Delta_l C_l^T \delta_l^f h_{(A_l)}^{(B_l)}. \quad (4.3.17)$$

where $h_{(A_l)}^{(B_l)} \equiv h_{(a_1}^{(b_1)} \dots h_{a_l)}^{(b_l)}$. The angle brackets denote a statistical average over the ensemble of fluctuations. Eq. (4.3.17) is entirely equivalent to the usual definition of the anisotropy power spectrum in terms of the variance of a_{lm} with $\delta_T(e) = \sum_{l>0} a_{lm} Y_{lm}(e)$. The temperature correlation function evaluates to

$$\langle \delta_T(e^c) \delta_T(e'^c) \rangle = \sum_{l=1}^\infty \frac{(2l+1)}{4\pi} C_l^T P_l(\cos \theta), \quad (4.3.18)$$

where θ is the angle between the directions e and e' , and P_l is a Legendre polynomial. In deriving Eq. (4.3.18) we used the result $e^{(A_l)} e'_{(A_l)} = (2l+1) \Delta_l P_l(\cos \theta)/(4\pi)$.

The power spectra for the polarization are defined similarly, but, since we choose our conventions for the power spectra to conform with Ref. [276], it is necessary to include an additional factor of $M_l/\sqrt{2}$ for each factor of the polarization. For example, for E -modes

$$\left(\frac{\pi}{I}\right)^2 \langle \mathcal{E}_{A_l} \mathcal{E}^{B_l} \rangle = \frac{l(l-1)}{(l+1)(l+2)} \Delta_l C_l^E \delta_l^f h_{(A_l)}^{(B_l)}. \quad (4.3.19)$$

²⁸ $T_{\text{CMB}} = 2.725$ K [285].

²⁹ The spectral dependence follows from, for example, the form of the source terms in Eq. (4.1.27) for massless particles, and from the energy dependence of the linearised scattering term in Eq. (4.3.41).

³⁰ We shall not be concerned here with spectral distortions in the CMB, which are an important probe of the energetics of the universe. For a recent review of this topic, and future prospects, see e.g. Ref. [287].

Unfortunately, the definitions of the polarization power spectra given in Ref. [276] differ from those in Ref. [275] by factors of $\sqrt{2}$; see Ref. [288] for details. For a parity-symmetric ensemble, B -mode polarization is uncorrelated with the E -mode and the temperature anisotropies. The correlation functions of linear polarization are most simply expressed in terms of the Stokes parameters. For two propagation directions e and e' , define Stokes parameter \bar{Q} , \bar{U} and \bar{Q}' and \bar{U}' using the direction of the tangent to the spherical geodesic connecting e and e' to define the $(e_1)^a$ basis vectors at the two points. Then the non-zero correlation functions assuming parity symmetry are [288–290]

$$\begin{aligned} \langle \bar{Q}\bar{Q}' \rangle &= \frac{1}{2} \sum_l \frac{2l+1}{4\pi} [C_l^E(d_{22}^l + d_{2-2}^l)(\theta) + C_l^B(d_{22}^l - d_{2-2}^l)(\theta)] \\ \langle \bar{U}\bar{U}' \rangle &= \frac{1}{2} \sum_l \frac{2l+1}{4\pi} [C_l^E(d_{22}^l - d_{2-2}^l)(\theta) + C_l^B(d_{22}^l + d_{2-2}^l)(\theta)], \end{aligned} \quad (4.3.20)$$

where d_{mn}^l are the reduced Wigner functions and the Stokes parameters are expressed in thermodynamic temperature, i.e. $Q(e^a) = (\pi/l) \int dE Q(E, e^a)$. The correlation properties of the linear polarization tensor $\mathcal{P}_{ab}(e^c) \equiv (\pi/l) \int dE \mathcal{P}_{ab}(E, e^c)$ are easily expressed in terms of these results. For example,³¹

$$2\langle \mathcal{P}_{ab}(e^c) \mathcal{P}^{ab}(e'^c) \rangle = \frac{1}{2} (1 + \cos^2 \theta) \langle \bar{Q}\bar{Q}' \rangle + \cos \theta \langle \bar{U}\bar{U}' \rangle. \quad (4.3.21)$$

The geometric factors $(1 + \cos^2 \theta)$ and $\cos \theta$ appear due to the contractions between the $(e_1)^a$ basis vectors at the two points. These would not be present if, instead, we parallel-transported $\mathcal{P}_{ab}(e^c)$ on the 2-sphere to the point e' and performed the contraction there.

4.3.2. Transformation properties under change of frame

The phase-space volume element $d^3\mathbf{x}d^3\mathbf{p}$ is Lorentz-invariant and so $\epsilon^{a*}f_{ab}\epsilon^b$ is a Lorentz scalar for any polarization 4-vector e^a . It follows that f_{ab} is properly covariant under Lorentz transformations. To see the implications of this for the observed polarization, consider a new velocity field $\tilde{u}^a = \gamma(u^a + v^a)$, where v^a is the projected relative velocity in the u^a frame and γ is the associated Lorentz factor. For a given photon with 4-momentum p^a , the energy and propagation directions in the \tilde{u}_a frame are given by the Doppler and aberration formulae:

$$\tilde{E} = \gamma E (1 - e^a v_a), \quad (4.3.22)$$

$$\tilde{e}^a = [\gamma(1 - e^b v_b)]^{-1} (u^a + e^a) - \gamma(u^a + v^a). \quad (4.3.23)$$

Note that \tilde{e}^a is a projected vector relative to \tilde{u}_a . The screen projection tensor for a given null direction transforms to

$$\tilde{s}_{ab} = s_{ab} + \frac{2\gamma}{E} p_{(a} s_{b)c} v^c + \frac{\gamma^2}{E^2} p_a p_b s_{cd} v^c v^d. \quad (4.3.24)$$

For any vector orthogonal to p^a , for example the 4-polarization e^a ,

$$\tilde{s}_c^a s_b^c e^b = \tilde{s}_b^a e^b \quad (p^a \epsilon_a = 0). \quad (4.3.25)$$

From this it follows that the screen-projected direction of polarization transforms as

$$s_b^a e^b \mapsto \tilde{s}_b^a e^b = \tilde{s}_c^a (s_b^c e^b), \quad (4.3.26)$$

and the observed polarization tensor by [224]

$$\tilde{E}^{-3} \tilde{P}_{ab}(\tilde{E}, \tilde{e}^c) = E^{-3} \tilde{s}_a^d \tilde{s}_b^{d2} P_{d_1 d_2}(E, e^c). \quad (4.3.27)$$

Under this transformation law the intensity, circular polarization and linear polarization do not mix. Moreover, $I(E, e^c)/E^3$ and $V(E, e^c)/E^3$ are frame-invariant for a given null direction, and the transformation law for $\mathcal{P}_{ab}(E, e^c)$ follows that for $P_{ab}(e^c)$. The degree of linear polarization $[\mathcal{L}\mathcal{P}_{ab}(E, e^c)\mathcal{P}^{ab}(E, e^c)]^{1/2}/I(E, e^c)$ is invariant under changes of frame. The transformation law for $P_{ab}(E, e^c)$ ensures that the tetrad components of $P_{ab}(E, e^c)/E^3$, and hence the Stokes parameters divided by E^3 , are invariant if the polarization basis vectors are transformed as

$$(\tilde{e}_i)^a = \tilde{s}_b^a (e_i)^b, \quad (4.3.28)$$

for $i = 1, 2$. An alternative way of viewing Eq. (4.3.28) is to note that, in terms of components on an orthonormal tetrad in the u^a frame and its Lorentz-boosted counterpart in the \tilde{u}^a frame, the basis vector \mathbf{e}_1 at direction \mathbf{e} is parallel-propagated along the geodesic connecting \mathbf{e} and $\tilde{\mathbf{e}}$ on the 2-sphere to obtain $\tilde{\mathbf{e}}_1$ [291].

³¹ This result corrects Eq. 2.20 of Ref. [224].

Under changes of frame, multipoles with different l mix because of Doppler and beaming effects. Using the invariance of $I(E, e^a)/E^3$, the multipoles $I_{A_l}(E)$ and the energy-integrated I_{A_l} transform as

$$\tilde{I}_{A_l}(\tilde{E}) = \Delta_l^{-1} \sum_{\mathcal{V}} \int d\Omega \gamma (1 - e^b v_b) I_{B_{\mathcal{V}}} [\gamma^{-1} \tilde{E} / (1 - e^c v_c)] e^{B_{\mathcal{V}}} \tilde{e}_{(A_l)}, \quad (4.3.29)$$

$$\tilde{I}_{A_l} = \sum_{\mathcal{V}} \Delta_{\mathcal{V}}^{-1} I_{B_{\mathcal{V}}} \int d\Omega [\gamma (1 - e^b v_b)]^2 e^{B_{\mathcal{V}}} \tilde{e}_{(A_l)}, \quad (4.3.30)$$

where the PSTF $\tilde{e}_{(A_l)}$ is with respect to \tilde{u}^a . The transformation of the energy-dependent multipoles is non-local in energy and therefore a little messy; expansions in terms of v^a can be found in Ref. [291]. The transformation of the energy-integrated multipoles are somewhat simpler: to first order in the relative velocity [224],

$$\tilde{I}_{A_l} = \tilde{h}_{(A_l)}^{(B_l)} I_{B_l} + (l-2)v^b I_{bA_l} - \frac{l(l+3)}{(2l+1)} v_{(a_l} I_{A_{l-1})}. \quad (4.3.31)$$

An equivalent result holds for V_{A_l} . The linear polarization is more complicated since E and B mix. To first-order in v^a ,

$$\tilde{\mathcal{E}}_{A_l} = \tilde{h}_{(A_l)}^{(B_l)} \mathcal{E}_{B_l} - \frac{l(l+3)}{(2l+1)} v_{(a_l} \mathcal{E}_{A_{l-1})} + \frac{(l-2)(l-1)(l+3)}{(l+1)^2} v^b \mathcal{E}_{bA_l} + \frac{6}{(l+1)} v_b \mathcal{E}^{bc}{}_{(a_l} \mathcal{B}_{A_{l-1})c}, \quad (4.3.32)$$

$$\tilde{\mathcal{B}}_{A_l} = \tilde{h}_{(A_l)}^{(B_l)} \mathcal{B}_{B_l} - \frac{l(l+3)}{(2l+1)} v_{(a_l} \mathcal{B}_{A_{l-1})} + \frac{(l-2)(l-1)(l+3)}{(l+1)^2} v^b \mathcal{B}_{bA_l} - \frac{6}{(l+1)} v_b \mathcal{E}^{bc}{}_{(a_l} \mathcal{E}_{A_{l-1})c}. \quad (4.3.33)$$

In an almost-FLRW model the polarization is a first-order quantity, as are physically-defined relative velocities. It follows that the E and B -mode multipoles are frame-invariant in linear theory.

4.3.3. Radiative transfer

The phase-space volume element $d^3\mathbf{x}d^3\mathbf{p}$ is conserved along a photon path and the polarization 4-vector ϵ^a is parallel-transported. It follows that f_{ab} is also parallel-propagated in phase space if there are no collisions (see e.g. Ref. [292] for a rigorous discussion). The observed polarization tensor is given by Eq. (4.3.2); multiplying by E^3 and taking the derivative in phase space gives [224]

$$\mathcal{L}[E^{-3}P_{ab}(E, e^c)] = 0, \quad (4.3.34)$$

where the Liouville operator \mathcal{L} acts on transverse tensors $A_{ab} = [A_{ab}]^{\text{TT}}$, like P_{ab} , as

$$\mathcal{L}[A_{ab}(E, e^c)] \equiv s_a^{d_1} s_b^{d_2} p^e \nabla_e A_{d_1 d_2}(E, e^c). \quad (4.3.35)$$

Physically, this means the observed polarization tensor is propagated as parallel as its projection properties allow. If we now include collisions, we obtain the exact Boltzmann equation

$$\mathcal{L}[E^{-3}P_{ab}(E, e^c)] = K_{ab}(E, e^c), \quad (4.3.36)$$

where $K_{ab}(E, e^c)$ is the fully projected (with s_{ab}) form of the invariant collision tensor. If we change frame to $\tilde{u}^a \equiv \gamma(u^a + v^a)$, the projected collision tensor transforms like P_{ab}/E^3 , i.e.

$$\tilde{K}_{ab}(\tilde{E}, \tilde{e}^c) = \tilde{s}_a^{d_1} \tilde{s}_b^{d_2} K_{d_1 d_2}(E, e^c). \quad (4.3.37)$$

This result is useful since the scattering tensor is often simplest to evaluate in some preferred frame, picked out by the physics of the scattering process. The scattering tensor in a general frame then follows from the transformation (4.3.37).

The Liouville operator \mathcal{L} preserves the irreducible decomposition of the polarization tensor (Eq. (4.3.3)), so that

$$\mathcal{L}[E^{-3}P_{ab}(E, e^c)] = \frac{1}{2} \frac{d}{d\lambda} [E^{-3}I(E, e^c)] s_{ab} + \mathcal{L}[E^{-3}\mathcal{P}_{ab}(E, e^c)] + \frac{1}{2} i \frac{d}{d\lambda} [E^{-3}V(E, e^c)] \epsilon_{abd} e^d. \quad (4.3.38)$$

If we propagate the (transverse) polarization basis vectors $(e_i)^a$ according to

$$s_{ab} p^c \nabla_c (e_i)^b = 0, \quad (4.3.39)$$

then all four Stokes brightness parameters (divided by E^3) are constant along the photon path when there are no collisions.

For the CMB, the dominant collisional process over the epochs of interest for the formation of anisotropies and polarization (around recombination and reionization) is Compton scattering. To an excellent approximation we can ignore the effects of Pauli blocking, induced scattering, and electron recoil in the rest frame of the scattering electron, so that the scattering may be approximated by classical Thomson scattering in the electron rest frame with no change in the photon energy. Furthermore, we can neglect the small velocity dispersion of the electrons due to their finite temperature and so

consider scattering off a cold gas of electrons with proper number density \tilde{n}_e in the electron rest frame.³² Denoting the rest frame by \tilde{u}^a , the *exact* projected collision tensor in the Thomson limit in that frame is [224]³³ (see also Ref. [296] for Klein–Nishina corrections)

$$\begin{aligned} \tilde{E}^2 \tilde{K}_{ab}(\tilde{E}, \tilde{e}^c) = & \tilde{n}_e \sigma_T \left\{ \frac{1}{2} \tilde{s}_{ab} \left[-\tilde{I}(\tilde{E}, \tilde{e}^c) + \tilde{I}(\tilde{E}) + \frac{1}{10} \tilde{I}_{d_1 d_2}(\tilde{E}) \tilde{e}^{d_1} \tilde{e}^{d_2} - \frac{3}{5} \tilde{\mathcal{E}}_{d_1 d_2}(\tilde{E}) \tilde{e}^{d_1} \tilde{e}^{d_2} \right] \right. \\ & \left. + \left[-\tilde{\mathcal{P}}_{ab}(\tilde{E}, \tilde{e}^c) - \frac{1}{10} [\tilde{I}_{ab}(\tilde{E})]^{\text{TT}} + \frac{3}{5} [\tilde{\mathcal{E}}_{ab}(\tilde{E})]^{\text{TT}} \right] + \frac{1}{2} i \tilde{\varepsilon}_{abd_1} \tilde{e}^{d_1} \left[-\tilde{V}(\tilde{E}, \tilde{e}^c) + \frac{1}{2} \tilde{V}_{d_2}(\tilde{E}) \tilde{e}^{d_2} \right] \right\}, \end{aligned} \quad (4.3.40)$$

where σ_T is the Thomson cross section. This expression for the scattering tensor follows from inserting the multipole decomposition of the polarization tensor into the Kernel for Thomson in-scattering (e.g. Ref. [282]), and integrating over scattering directions. An expression equivalent to Eq. (4.3.40) appears to have been first derived in Ref. [284].

Some general observations follow from Eq. (4.3.40) [224,228,284]. It is written in irreducible form with the first set of terms on the right affecting the total intensity, the second set the linear polarization and the third the circular polarization. In each case, scattering out of the phase-space element is described by $-\tilde{n}_e \sigma_T P_{ab}$. For I , in-scattering couples to the monopole and quadrupole in total intensity, and to the E -mode quadrupole. Comparison with Eq. (4.1.14) and (4.1.15) show that there is no change in energy density in the electron rest frame due to Thomson scattering, but there is momentum exchange if the radiation has a dipole moment. Linear polarization is generated by in-scattering of the quadrupoles in total intensity and E -mode polarization. Comparison with Eq. (4.3.9) shows that in the electron rest-frame, the polarization is *generated* purely as an E -mode quadrupole. The transformations (4.3.32) and (4.3.33) show that this is generally not true in some other frame. Finally, we see that circular polarization is fully decoupled from total intensity and linear polarization. In consequence, in any frame the circular polarization will remain exactly zero if it is initially.

The scattering tensor (4.3.40) is first-order in small quantities about an FLRW background. Transforming to a general frame u^a , and keeping only first-order terms, we find

$$\begin{aligned} E^2 K_{ab}(E, e^c) = & n_e \sigma_T \left\{ \frac{1}{2} s_{ab} \left[-I(E, e^c) + I(E) - e^c v_c E^4 \frac{\partial}{\partial E} \left(\frac{I(E)}{E^3} \right) + \frac{1}{10} I_{d_1 d_2}(E) e^{d_1} e^{d_2} - \frac{3}{5} \mathcal{E}_{d_1 d_2}(E) e^{d_1} e^{d_2} \right] \right. \\ & \left. + \left[-\mathcal{P}_{ab}(E, e^c) - \frac{1}{10} [I_{ab}(E, e^c)]^{\text{TT}} + \frac{3}{5} [\mathcal{E}_{ab}(E, e^c)]^{\text{TT}} \right] + \frac{1}{2} i \varepsilon_{abd_1} e^{d_1} \left[-V(E, e^c) + \frac{1}{2} V_{d_2}(E) e^{d_2} \right] \right\}, \end{aligned} \quad (4.3.41)$$

where n_e is the electron density relative to u^a .

The Boltzmann equation (4.3.36) can be written in multipole form by expressing $P_{ab}(E, e^c)$ as a multipole expansion using Eqs. (4.3.6), (4.3.7) and (4.3.9), and decomposing the resulting equation into multipoles. This leads to four sets of multipole hierarchies for $I_{A_l}(E)$, $\mathcal{E}_{A_l}(E)$, $\mathcal{B}_{A_l}(E)$, and $V_{A_l}(E)$. In linear theory, those for $I_{A_l}(E)$ and $\mathcal{E}_{A_l}(E)$ are coupled by Thomson scattering, and, generally, those for $\mathcal{E}_{A_l}(E)$ and $\mathcal{B}_{A_l}(E)$ are coupled by advection (see below). The exact equations for total intensity and circular polarization follow Eq. (4.1.8) but with, for example, $F_{A_l}(\lambda)$ replaced by $I_{A_l}(E)/E^3$, $E = \lambda$ and the collision multipoles

$$E^2 C_{A_l}[f] = n_e \sigma_T \left[-I_{A_l}(E) + I(E) \delta_{l0} - v_{a_1} E^4 \frac{\partial}{\partial E} \left(\frac{I(E)}{E^3} \right) \delta_{l1} + \left(\frac{1}{10} I_{a_1 a_2}(E) - \frac{3}{5} \mathcal{E}_{a_1 a_2}(E) \right) \delta_{l2} + \dots \right]. \quad (4.3.42)$$

The multipole equations for $\mathcal{E}_{A_l}(E)$ and $\mathcal{B}_{A_l}(E)$ are more involved than those for the intensity or circular polarization and only the linearised equations (about FLRW) have been calculated to date [224]. (The $l = 2$ equation was given in an orthonormal tetrad in Ref. [284] under the assumption that the higher multipoles vanish.) In linear form, for the E -mode polarization

$$\begin{aligned} \dot{\mathcal{E}}_{A_l}(E) - \frac{1}{3} \Theta E^4 \frac{\partial}{\partial E} [E^{-3} \mathcal{E}_{A_l}(E)] + D_{(a_1} \mathcal{E}_{A_{l-1})}(E) + \frac{(l+3)(l-1)}{(l+1)(2l+3)} D^b \mathcal{E}_{b A_l}(E) - \frac{2}{(l+1)} \text{curl } \mathcal{B}_{A_l}(E) \\ = -n_e \sigma_T \mathcal{E}_{A_l}(E) - \frac{1}{10} n_e \sigma_T [I_{a_1 a_2}(E) - 6 \mathcal{E}_{a_1 a_2}(E)] \delta_{l2}, \end{aligned} \quad (4.3.43)$$

and for the B -mode

$$\begin{aligned} \dot{\mathcal{B}}_{A_l}(E) - \frac{1}{3} \Theta E^4 \frac{\partial}{\partial E} [E^{-3} \mathcal{B}_{A_l}(E)] + D_{(a_1} \mathcal{B}_{A_{l-1})}(E) + \frac{(l+3)(l-1)}{(l+1)(2l+3)} D^b \mathcal{B}_{b A_l}(E) + \frac{2}{(l+1)} \text{curl } \mathcal{E}_{A_l}(E) \\ = -n_e \sigma_T \mathcal{B}_{A_l}(E). \end{aligned} \quad (4.3.44)$$

The E and B -mode multipoles are coupled by the curl terms where, recall, the curl of a PSTF tensor is defined by Eq. (4.1.35). In a general almost-FLRW cosmology, B -mode polarization is generated only by advection of the E -mode. This does not happen

³² An important exception of relevance for secondary anisotropies in the CMB is Compton scattering in hot intra-cluster gas – the Sunyaev–Zel’dovich effect [270] (see Ref. [273] for a review) – where the electron temperatures can be $\sim O(10)$ keV. The low-frequency decrement in total intensity of the CMB in the direction of clusters is proportional to the electron temperature. Finite temperature effects for polarization are considered in Refs. [293–295] and can give up to $\sim O(10)\%$ corrections to the results for cold intra-cluster gas at the spectral peak of the signal.

³³ This corrects two sign errors in the right-hand side of Eq. 3.7 of Ref. [224]. There, the term involving $\tilde{\mathcal{E}}_{ab}$ in the first line and \tilde{I}_{ab} in the second have the wrong sign.

if the perturbations about FLRW are curl-free, as is the case for scalar perturbations. We thus have the important result that linear scalar perturbations do not generate B -mode polarization [275,276]; see also Sections 4.3.4 and 4.3.5.

The equations for the energy-integrated multipoles follow from integrating Eqs. (4.1.8), (4.3.43) and (4.3.44) over energy. In linear form,

$$\begin{aligned} \dot{\mathcal{E}}_{A_l} + \frac{4}{3}\Theta\mathcal{E}_{A_l} + \frac{(l+3)(l-1)}{(l+1)^2}D^b\mathcal{E}_{bA_l} + \frac{l}{(2l+1)}D_{(a_l}\mathcal{E}_{A_{l-1})} - \frac{2}{(l+1)}\text{curl}\mathcal{B}_{A_l} \\ = -n_e\sigma_T\left[\mathcal{E}_{A_l} + \left(\frac{1}{10}I_{a_1a_2} - \frac{3}{5}\mathcal{E}_{a_1a_2}\right)\delta_{l2}\right], \end{aligned} \quad (4.3.45)$$

$$\dot{\mathcal{B}}_{A_l} + \frac{4}{3}\Theta\mathcal{B}_{A_l} + \frac{(l+3)(l-1)}{(l+1)^2}D^b\mathcal{B}_{bA_l} + \frac{l}{(2l+1)}D_{(a_l}\mathcal{B}_{A_{l-1})} + \frac{2}{(l+1)}\text{curl}\mathcal{E}_{A_l} = -n_e\sigma_T\mathcal{B}_{A_l}. \quad (4.3.46)$$

For the circular polarization,

$$\dot{V}_{A_l} + \frac{4}{3}\Theta V_{A_l} + D^bV_{bA_l} + \frac{l}{(2l+1)}D_{(a_l}V_{A_{l-1})} = -n_e\sigma_T\left(V_{A_l} - \frac{1}{2}V_{a_1}\delta_{l1}\right), \quad (4.3.47)$$

and for the total intensity

$$\begin{aligned} \dot{I}_{A_l} + \frac{4}{3}\Theta I_{A_l} + D^bI_{bA_l} + \frac{l}{(2l+1)}D_{(a_l}I_{A_{l-1})} + \frac{4}{3}IA_{a_1}\delta_{l1} + \frac{8}{15}I\sigma_{a_1a_2}\delta_{l2} \\ = -n_e\sigma_T\left[I_{A_l} - I\delta_{l0} - \frac{4}{3}IV_{a_1}\delta_{l1} - \left(\frac{1}{10}I_{a_1a_2} - \frac{3}{5}\mathcal{E}_{a_1a_2}\right)\delta_{l2}\right]. \end{aligned} \quad (4.3.48)$$

The monopole moment does not vanish in a homogeneous background so we use its projected gradient to characterise the perturbation in the radiation energy density. Defining $\Delta_a \equiv aD_aI/l$ for the radiation, the projected gradient of the $l = 0$ moment of Eq. (4.3.48) gives

$$\dot{\Delta}_a + \frac{a}{l}D_aD^bI_b + 4\dot{h}_a = 0, \quad (4.3.49)$$

where, to linear order, $3\dot{h}_a = a(3HA_a + D_a\Theta)$ from Eq. (4.1.25). The above equation also follows from integrating Eq. (4.1.29) with $\lambda^3 d\lambda$ and noting that the linear Thomson collision term has no monopole.

Eq. (4.3.45)–(4.3.49) provide a complete description of the linear evolution of the CMB anisotropies and polarization in general almost-FLRW models. In particular, they are valid for all types of perturbation since no harmonic expansion has been made. We also see that the highest rank of the source terms is two so that only scalar, vector and tensor modes can be excited. In the following subsections we give integral solutions of the multipole equations for scalar, vector and tensor modes and briefly discuss the physics of each. We shall not consider circular polarization any further since it is not generated by Thomson scattering.

4.3.4. Scalar perturbations

Following the discussion in Section 4.1.2, we expand the PSTF multipoles in the harmonic tensors $\mathcal{Q}_{A_l}^{(0)}$ defined in Eq. (4.1.31):

$$I_{A_l} = I \sum_k \left(\prod_{n=0}^l \kappa_n^{(0)} \right)^{-1} I_l^{(0)} \mathcal{Q}_{A_l}^{(0)}, \quad l \geq 1, \quad (4.3.50)$$

$$\mathcal{E}_{A_l} = I \sum_k \left(\prod_{n=0}^l \kappa_n^{(0)} \right)^{-1} \mathcal{E}_l^{(0)} \mathcal{Q}_{A_l}^{(0)}, \quad l \geq 2, \quad (4.3.51)$$

where, for later convenience, we have introduced

$$\kappa_l^{(m)} \equiv [1 - (l^2 - 1 - m)K/k^2]^{1/2}, \quad l \geq m, \quad (4.3.52)$$

with $\kappa_0^{(0)} = 1$. It follows from Eq. (4.1.34) that $\text{curl}\mathcal{E}_{A_l} = 0$ and so the B -mode of polarization is not excited by advection (or, as always, Thomson scattering) for scalar perturbations so we need not include it. For the projected gradient of the radiation energy density I , we expand as

$$\Delta_a \equiv \frac{aD_aI}{l} = - \sum_k k l_0^{(0)} \mathcal{Q}_a^{(0)}. \quad (4.3.53)$$

In scalar harmonic form, the linearised multipole equations now become

$$\begin{aligned} \dot{l}_l^{(0)} + \frac{k}{a} \left[\frac{(l+1)}{(2l+1)} \kappa_{l+1}^{(0)} l_{l+1}^{(0)} - \frac{l}{(2l+1)} \kappa_l^{(0)} l_{l-1}^{(0)} \right] + 4\dot{h}\delta_{l0} + \frac{4}{3} \frac{k}{a} A\delta_{l1} + \frac{8}{15} \frac{k}{a} \kappa_2^{(0)} \sigma\delta_{l2} \\ = -n_e\sigma_T \left[l_l^{(0)} - I_0^{(0)}\delta_{l0} - \frac{4}{3}v\delta_{l1} - \left(\frac{1}{10}l_2^{(0)} - \frac{3}{5}\mathcal{E}_2^{(0)} \right) \delta_{l2} \right] \end{aligned} \quad (4.3.54)$$

for $l \geq 0$ and, for E -mode polarization,

$$\dot{\xi}_l^{(0)} + \frac{k}{a} \left[\frac{(l+3)(l-1)}{(2l+1)(l+1)} \kappa_{l+1}^{(0)} \xi_{l+1}^{(0)} - \frac{l}{(2l+1)} \kappa_l^{(0)} \xi_{l-1}^{(0)} \right] = -n_e \sigma_T \left[\xi_l^{(0)} + \left(\frac{1}{10} I_2^{(0)} - \frac{3}{5} \mathcal{E}_2^{(0)} \right) \delta_{l2} \right]. \quad (4.3.55)$$

In deriving these, we have used Eqs. (4.1.31) and (4.1.33). The kinematic quantities that enter Eq. (4.3.54) have been mode-expanded following Eq. (4.2.9) with $v_a = \sum v \mathcal{Q}_a^{(0)}$ for the relative velocity of the electron–baryon plasma. These multipole equations hold for a general FLRW model and are fully equivalent to those obtained in Ref. [229] using the total-angular momentum method.

Note that in closed models, the $\mathcal{Q}_{A_l}^{(m)}$ necessarily vanish globally for $l \geq \nu$, so the same is true of the I_{A_l} and \mathcal{E}_{A_l} from a given harmonic mode. Power thus streams up the hierarchy as far as the $l = \nu - 1$ multipole, but is then reflected back down. This is enforced in Eqs. (4.3.54) and (4.3.55) by $\kappa_\nu^{(m)} = 0$. That there is a maximum multipole, hence minimum angular scale, that can arise from projection of a given harmonic mode is due to the focusing of geodesics in closed FLRW models.

Early computer codes to compute the CMB anisotropy integrated a carefully truncated version of the multipole equations directly. A major advance was made in Ref. [218] where the Boltzmann hierarchy was formally integrated thus allowing a very efficient solution for the CMB anisotropy. This procedure was implemented in the CMBFAST code³⁴ and, later, in parallelised derivatives such as CAMB [262].³⁵

The integral solution for the total intensity for general spatial curvature is [224,229,297]

$$I_l^{(0)} = 4 \int^{t_R} dt e^{-\tau} \left\{ \left[-\frac{k}{a} \sigma_k + \frac{3}{16} n_e \sigma_T (\kappa_2^{(0)})^{-1} (I_2^{(0)} - 6 \mathcal{E}_2^{(0)}) \right] \left[\frac{1}{3} \Phi_l^\nu(x) + \frac{1}{(\nu^2 + 1)} \frac{d^2}{dx^2} \Phi_l^\nu(x) \right] - \left(\frac{k}{a} A - n_e \sigma_T v \right) \frac{1}{\sqrt{\nu^2 + 1}} \frac{d}{dx} \Phi_l^\nu(x) - \left[\dot{h} - \frac{1}{4} n_e \sigma_T I_l^{(0)} \right] \Phi_l^\nu(x) \right\}. \quad (4.3.56)$$

Here, $\tau \equiv \int n_e \sigma_T dt$ is the optical depth back along the line of sight and, recall, $x = \sqrt{|K|} \chi$ with χ the comoving radial distance (or, equivalently, conformal look-back time) along the line of sight and $\Phi_l^\nu(x)$ are the ultra-spherical Bessel functions with $\nu^2 = (k^2 + K)/|K|$ for scalar perturbations. The geometric factors $\Phi_l^\nu/3 + (\nu^2 + 1)^{-1} d^2 \Phi_l^\nu/dx^2$ and $(\nu^2 + 1)^{-1/2} d \Phi_l^\nu/dx$ arise from the projections of $\mathcal{Q}_{ab}^{(0)} e^a e^b$ and $\mathcal{Q}_a^{(0)} e^a$ respectively, at x back along the line of sight. Source terms of these forms enter the Boltzmann equation through the shear and the quadrupole dependence of Thomson scattering, for $\mathcal{Q}_{ab}^{(0)} e^a e^b$, and the acceleration and baryon velocity, for $\mathcal{Q}_a^{(0)} e^a$. Their angular projections follow from Eq. (4.1.46) and (4.1.47).

The integral solution for the E -mode polarization is [224,229,297]

$$\xi_l^{(0)} = -\frac{l(l-1)}{(\nu^2 + 1)} \int^{t_R} dt n_e \sigma_T e^{-\tau} (\kappa_2^{(0)})^{-1} \left(\frac{3}{4} I_2^{(0)} - \frac{9}{2} \mathcal{E}_2^{(0)} \right) \frac{\Phi_l^\nu(x)}{\sinh^2 x}. \quad (4.3.57)$$

The geometric term $\Phi_l^\nu(x)/\sinh^2 x$ now arises from the projection of source terms that go like $[\mathcal{Q}_{ab}^{(0)}]^{TT}$; see the form of the linear-polarization source terms in the scattering tensor (4.3.41). Eqs. (4.3.56) and (4.3.57) are valid in an open universe. For closed models one should replace the hyperbolic functions by their trigonometric counterparts, and $\nu^2 + n$ by $\nu^2 - n$ where n is an integer.

Since we are working in linear theory, the coefficients $I_l^{(0)}$ and $\xi_l^{(0)}$ will depend linearly on the primordial perturbation $\phi_{(k)}$. Introducing transfer functions, $T_l^I(k)$ and $T_l^E(k)$, we can write

$$I_l^{(0)} = T_l^I(k) \phi_{(k)}, \quad \xi_l^{(0)} = \frac{M_l}{\sqrt{2}} T_l^E(k) \phi_{(k)}, \quad (4.3.58)$$

where the normalisation for polarization is to account for the additional l -dependent factors in Eq. (4.3.19). Note that the symmetry of the background model ensures that transfer functions depend only on the magnitude of the wavenumber k . The choice of $\phi_{(k)}$ is one of convention. For the adiabatic, growing-mode initial conditions that follow from single-field inflation, the convenient choice is the (constant) curvature perturbation $\mathcal{R}_{(k)}$ on comoving hypersurfaces. For models with isocurvature fluctuations, the relative entropy gradient is appropriate. More generally, in models with mixed initial conditions having N degrees of freedom per harmonic mode, the transfer functions generalise to N functions per l and k .

We shall express the CMB power spectra in terms of the power spectrum of the $\phi_{(k)}$ and, since we discuss higher-rank perturbations later, we shall sketch the derivation for rank- m perturbations in which case the $\phi_{(k)}^{(\pm m)}$ are the coefficients of the expansion of the PSTF tensor ϕ_{A_m} in terms of the $\mathcal{Q}_{A_m}^{(\pm m)}$.³⁶ It is always possible to choose the $\mathcal{Q}_{A_m}^{(m)}$ such that

$$\langle \phi_{(k)}^{(\pm m)} \phi_{(k')}^{(\pm m)*} \rangle = f_\phi(k) \delta_{kk'}, \quad (4.3.59)$$

³⁴ <http://www.cmbfast.org>.

³⁵ <http://camb.info/>.

³⁶ The derivation here follows that given by Antony Lewis in the unpublished CAMB notes available as <http://cosmologist.info/notes/CAMB.ps.gz>.

with no correlations between the opposite parity modes, for perturbations that are statistically isotropic, homogeneous and parity-invariant; see Ref. [298] for further details, and, e.g. Ref. [224] for specific constructions for scalar and tensor perturbations. Here, the symbolic delta-function enforces $\sum_{k'} A_{(k')}^{(\pm m)} \delta_{kk'} = A_{(k)}^{(\pm m)}$, and $f_\phi(k)$ depends only on the magnitude of the wavenumber k . The power spectrum of ϕ_{A_m} is proportional to $f_\phi(k)$ and here we define it so that

$$\begin{aligned} \langle \phi_{A_m} \phi^{A_m} \rangle &= \sum_k f_\phi(k) \mathcal{Q}_{A_m}^{(\pm m)} (\mathcal{Q}^{(\pm m)A_m})^* \\ &= \sum_k f_\phi(k) \frac{1}{V} \int dV \mathcal{Q}_{A_m}^{(\pm m)} (\mathcal{Q}^{(\pm m)A_m})^* \\ &\equiv \int \frac{v dv}{(v^2 + 1)} \mathcal{P}_\phi(k). \end{aligned} \quad (4.3.60)$$

Here, we have used statistical homogeneity to replace the correlator by an integral over a spatial volume V that we shall let tend to all space. Note that for scalar modes the integration measure $v dv / (v^2 + 1) = d \ln k$, but this is not the case for higher rank perturbations. For non-flat models there is no universal convention for the definition of $\mathcal{P}_\phi(k)$ for $m > 0$ but, of course, any ambiguity is removed when a definite physical model for the generation of fluctuations is considered. For example, for $m = 2$ and using the (gauge-invariant) metric perturbation \mathcal{H}_{ab} for ϕ_{ab} ,

$$\mathcal{P}_{\mathcal{H}}(k) \propto \frac{(v^2 + 4)}{v^2} \tanh(\pi v/2) \quad (4.3.61)$$

for the minimal scale-invariant open inflation model of Ref. [299]. Forming the power spectra from Eq. (4.3.17) and (4.3.19), we now have

$$\begin{aligned} \frac{(2l+1)\Delta_l}{\pi^2} C_l^{XY} &= \sum_k \left(\prod_{n=m}^l \kappa_n^{(m)} \right)^{-2} T_l^X(k) T_l^Y(k) f_\phi(k) \mathcal{Q}_{A_l}^{(\pm m)} (\mathcal{Q}^{(\pm m)A_l})^* \\ &= \sum_k \left[\left(\prod_{n=m}^l \kappa_n^{(m)} \right)^{-2} T_l^X(k) T_l^Y(k) f_\phi(k) \left(\frac{-a}{k} \right)^{l-m} \frac{1}{V} \int dV D_{A_{l-m}} \mathcal{Q}_{a_{l-m+1} \dots a_l}^{(\pm m)} (\mathcal{Q}^{(\pm m)A_l})^* \right] \\ &= \sum_k \left[\left(\prod_{n=m}^l \kappa_n^{(m)} \right)^{-2} T_l^X(k) T_l^Y(k) f_\phi(k) \left(\frac{a}{k} \right)^{l-m} \frac{1}{V} \int dV \mathcal{Q}_{a_{l-m+1} \dots a_l}^{(\pm m)} D_{A_{l-m}} (\mathcal{Q}^{(\pm m)A_l})^* \right] \\ &= \frac{2^{l-m} (l+m)! (l-m)!}{(2l)!} \sum_k \left[(\kappa_m^{(m)})^{-2} T_l^X(k) T_l^Y(k) f_\phi(k) \frac{1}{V} \int dV \mathcal{Q}_{A_m}^{(\pm m)} (\mathcal{Q}^{(\pm m)A_m})^* \right], \end{aligned} \quad (4.3.62)$$

where we have integrated by parts and made repeated use of the divergence relations for the $\mathcal{Q}_{A_l}^{(m)}$, e.g. Eq. (4.1.33). Here, XY is equal to TT , TE , EE and BB only, by parity. If we now compare Eqs. (4.3.60) and (4.3.62), we see that

$$C_l^{XY} = \frac{\pi}{4} \frac{(l+m)! (l-m)!}{2^m (l!)^2} \int \frac{v dv}{(v^2 + 1)} (\kappa_m^{(m)})^{-2} T_l^X(k) T_l^Y(k) \mathcal{P}_\phi(k), \quad (4.3.63)$$

where, recall, $\kappa_0^{(0)} = 1$. In closed models, we replace $v^2 + 1$ by $v^2 - 1$ and the integral becomes a discrete sum over integer v .

The simple physics of CMB anisotropies for scalar perturbations is more apparent if we integrate the shear and acceleration terms in Eq. (4.3.56) by parts. We then find

$$\begin{aligned} \frac{1}{4} I_l^{(0)} - \left(\frac{a}{k} \dot{\sigma} + A \right) \delta_{l0} - \frac{1}{3} \sigma \delta_{l1} &= \int^{t_R} dt e^{-\tau} \left\{ n_e \sigma_T \left[\frac{1}{4} I_0^{(0)} - \left(\frac{a}{k} \dot{\sigma} + A \right) \right] \Phi_l^v(x) + n_e \sigma_T (v - \sigma) \frac{1}{\sqrt{v^2 + 1}} \frac{d}{dx} \Phi_l^v(x) \right. \\ &\quad \left. + \frac{3}{16} n_e \sigma_T (\kappa_2^{(0)})^{-1} (I_2^{(0)} - 6\mathcal{E}_2^{(0)}) \left[\frac{1}{3} \Phi_l^v(x) + \frac{1}{(v^2 + 1)} \frac{d^2}{dx^2} \Phi_l^v(x) \right] + 2 \dot{\Phi}_l^v(x) \right\}, \end{aligned} \quad (4.3.64)$$

where Φ is the scalar potential for the Weyl tensor,

$$E_{ab} = \sum_k \left(\frac{k}{a} \right)^2 \Phi \mathcal{Q}_{ab}^{(0)}, \quad (4.3.65)$$

which plays the role of the (conformally-invariant) Newtonian potential. We have written Eq. (4.3.64) in such a form that the combination of terms on the left-hand side (which are evaluated at R) are independent of the choice of u^a , as are each set on the right-hand side. In deriving Eq. (4.3.64) we have used the linearised $\dot{\sigma}_{ab}$, \dot{E}_{ab} equations and the shear constraint of Section 1.3.1. As we can evaluate Eq. (4.3.64) in any frame, it is convenient to choose the conformal Newtonian gauge in which the velocity field u^a has zero shear and vorticity. In this frame, the linearised $\dot{\sigma}_{ab}$ equation becomes a constraint that determines the acceleration:

$$D_{(a} A_{b)} = E_{ab} - \frac{1}{2} \pi_{ab}. \quad (4.3.66)$$

In harmonic form, this is $A = -\Phi + \rho a^2 \pi / 2$, where $\pi_{ab} = \sum_k \rho \pi \mathcal{Q}_{ab}^{(0)}$. In terms of the usual conformal Newtonian metric (e.g. Ref. [217])

$$ds^2 = a^2 [-(1 + 2\psi_N)d\eta^2 + (1 + 2\phi_N)\gamma_{ij}dx^i dx^j], \quad (4.3.67)$$

where γ_{ij} is the conformal spatial metric in the FLRW background, $A_a = D_a \psi_N$ or, in harmonics, $A = -\psi_N$. Note also that the conformally-invariant Weyl potential $2\Phi = -\phi_N + \psi_N$ and $\phi_N = -\psi_N$ in the absence of anisotropic stress. If we further approximate to the case of sharp last scattering, which is valid for $k^{-1} \gg 50$ Mpc, and ignore reionization and the anisotropic nature of Thomson scattering, the temperature anisotropy in the Newtonian gauge reduces to

$$\left[\frac{1}{4} I_l^{(0)} + \psi_N \delta_{l0} \right]_R = \left(\frac{I_0^{(0)}}{4} + \psi_N \right) \phi_l^v(x_*) + v_N \frac{1}{\sqrt{v^2 + 1}} \frac{d}{dx} \phi_l^v(x) \Big|_{x_*} + 2 \int_{t_*}^{t_R} \dot{\phi} \phi_l^v(x) dt, \quad (4.3.68)$$

where, e.g. t_* is the time at last scattering. The temperature anisotropy can now be interpreted in terms of projections of three terms at last scattering: (i) the intrinsic temperature variations $I_0^{(0)}/4$ (recall, $I_0^{(0)}$ is related to the projected gradient of the energy density of the CMB); (ii) the Newtonian potential ψ_N , which appears because of gravitational redshifting; and (iii) Doppler shifts of the form $e_a v_N^a$, where v_N^a is the baryon velocity relative to the zero-shear u^a . The integrated (Sachs–Wolfe) term in Eq. (4.3.68) contributes when the Weyl potential evolves in time, such as when dark energy starts to dominate the expansion dynamics at low redshift, and arises because of the net blueshift accrued as a photon traverses a decaying potential well.

On comoving scales ~ 30 Mpc or greater, photon diffusion due to the finite mean-free path to Thomson scattering can be ignored. In this limit, the dynamics of the source terms in Eq. (4.3.68) can be reduced to that of a driven oscillator [300]. To see this, note from Eq. (4.3.48) that in the limit of tight-coupling, $I_a = 4lv_a/3$ and that $I_{A_l} = \mathcal{E}_{A_l} = 0$ for $l \geq 2$. The CMB is therefore isotropic in the baryon rest-frame and the linearised momentum evolution for the combined photon–baryon fluid gives

$$\dot{v}_a + \frac{HR}{(1+R)} v_a + \frac{1}{4(1+R)} \frac{\Delta_a}{a} + A_a = 0, \quad (4.3.69)$$

where we have ignored baryon pressure. Here, $R \equiv 3\rho^{(b)}/(4\rho^{(\nu)})$, which scales as a , and Δ_a is the fractional comoving density gradient for the CMB. The evolution of Δ_a was given in Eq. (4.3.49) which now becomes

$$\dot{\Delta}_a + 4\dot{h}_a + \frac{4}{3} a D_a D^b v_b = 0. \quad (4.3.70)$$

Switching to conformal time and combining Eqs. (4.3.69) and (4.3.70), we find

$$\Delta''_a + \frac{\mathcal{H}R}{(1+R)} \Delta'_a - \frac{1}{3(1+R)} a^2 D_a D^b \Delta_b = -4h''_a - \frac{4\mathcal{H}R}{(1+R)} h'_a + \frac{4}{3} a^3 D_a D^b A_b, \quad (4.3.71)$$

where $\mathcal{H} = aH$ is the conformal Hubble parameter. This equation is valid in any frame; it describes a driven oscillator whose free oscillations are at frequency kc_s , where the sound speed $c_s^2 = 1/[3(1+R)]$, and are damped by the expansion of the universe acting on the baryon velocity. If we now specialise to the Newtonian frame, we can express the driving terms on the right in terms of ϕ_N and ψ_N . We use Eq. (4.1.25) and the zero-shear \dot{E}_{ab} equation and shear constraint to find

$$a D_{(a} \dot{h}_{b)} = (a^2 D_{(a} D_{b)} \phi_N)', \quad (4.3.72)$$

or $\dot{h} = \dot{\phi}_N$ in harmonics. Finally, we recover the standard harmonic form of the oscillator equation in the Newtonian frame:

$$\Delta'' + \frac{\mathcal{H}R}{(1+R)} \Delta' + \frac{k^2}{3(1+R)} \Delta = -4\phi_N'' - \frac{4\mathcal{H}R}{(1+R)} \phi_N' - \frac{4}{3} k^2 \psi_N. \quad (4.3.73)$$

For adiabatic initial conditions, the cosine solution is excited and *all* modes with $k \int^{t_*} c_s d\eta = n\pi$ are at extrema of their oscillation at last scattering. This gives a series of acoustic oscillations in the temperature power spectrum [301], of which the first five have now been observed by a combination of sub-orbital experiments and the WMAP satellite [19]. An example of the CMB power spectra in a Λ CDM model is shown in Fig. 4, computed with the CAMB code [262]. The acoustic peaks are a rich source of cosmological information. Their relative heights depends on the baryon density (i.e. R) and matter density since these affect the midpoint of the acoustic oscillation and the efficacy of the gravitational driving in Eq. (4.3.73) [300]. The angular position of the peaks depends on the type of initial condition and the angular diameter distance to last scattering. Moreover, the general morphology of the spectra is related to the distribution of primordial power with scale, i.e. the power spectrum $\mathcal{P}_\phi(k)$. For current constraints from the CMB see Ref. [302].

On smaller scales photon diffusion becomes important. The breakdown of tight-coupling has two important effects on the CMB. First, the acoustic oscillations are exponentially damped and this gives the rapid decline of C_l^T at high l apparent in Fig. 4. Second, anisotropies can start to grow in the CMB intensity and this produces linear polarization on Thomson

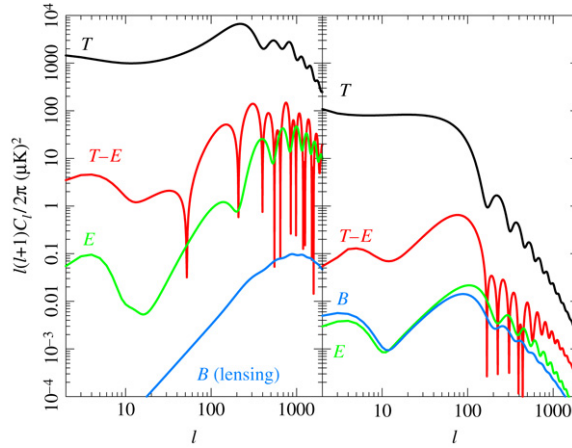


Fig. 4. Power spectra produced by adiabatic scalar perturbations (left) and tensor perturbations (right) for a tensor-to-scalar ratio $r = 0.20$ and optical depth to reionization of 0.08. The power spectrum of the B -modes produced by gravitational lensing of the scalar E -mode polarization is also shown on the left.

scattering. To first-order in the ratio of the mean-free time to the expansion time or the wavelength of the perturbation, the polarization is an E -mode quadrupole:

$$\mathcal{E}_{ab} \approx \frac{8}{45} \frac{I}{n_e \sigma_T} (\sigma_{ab} + D_{(a} v_{b)}), \quad (4.3.74)$$

which follows quite generally from Eqs. (4.3.45) and (4.3.48). For scalar perturbations, the polarization thus traces the projected derivative of the baryon velocity relative to the Newtonian frame. The peaks in the C_l^E spectrum thus occur at the minima of C_l^I as the baryon velocity oscillates $\pi/2$ out of phase with Δ . This behaviour can be seen in Fig. 4. The large-angle polarization from recombination is necessarily small by causality, but a large-angle signal is generated by re-scattering at reionization [303,304].

4.3.5. Vector perturbations

Vector modes describe vortical motions of the cosmic fluids. They have received considerably less attention than scalar and tensor modes mainly because they are not excited during inflation. Furthermore, due to conservation of angular momentum, the vorticity of radiation decays as $1/a$ and matter as $1/a^2$ so that vector modes are generally singular.³⁷ Vector modes are important in models with sources, such as magnetic fields (see Ref. [66] for a recent review) or topological defects [305].

The CMB anisotropies from vector modes were first studied comprehensively in Ref. [221]; the full kinetic theory treatment was developed in the total-angular-momentum method in Refs. [228,229]. Here we shall extend the $1 + 3$ -covariant treatment of Ref. [226] to general spatial curvature.

We expand the radiation anisotropies and polarization in the PSTF derivatives of the vector harmonics as

$$I_{A_l} = I \sum_k \left(\prod_{n=1}^l \kappa_n^{(1)} \right)^{-1} I_l^{(\pm 1)} \mathcal{Q}_{A_l}^{(\pm 1)}, \quad l \geq 1, \quad (4.3.75)$$

$$\mathcal{E}_{A_l} = I \sum_k \left(\prod_{n=1}^l \kappa_n^{(1)} \right)^{-1} \mathcal{E}_l^{(\pm 1)} \mathcal{Q}_{A_l}^{(\pm 1)}, \quad l \geq 1, \quad (4.3.76)$$

$$\mathcal{B}_{A_l} = I \sum_k \left(\prod_{n=1}^l \kappa_n^{(1)} \right)^{-1} \mathcal{B}_l^{(\pm 1)} \mathcal{Q}_{A_l}^{(\pm 1)}, \quad l \geq 1, \quad (4.3.77)$$

where the sum over harmonics, \sum_k , includes a sum over the two parity states of vector harmonics labelled ± 1 . For non-scalar perturbations, B -mode polarization is produced from E -modes by advection so we must now include \mathcal{B}_{A_l} . Substituting in Eq. (4.3.48), we find

$$\begin{aligned} j_l^{(\pm 1)} + \frac{k}{a} \left[\frac{l(l+2)}{(l+1)(2l+1)} \kappa_{l+1}^{(1)} j_{l+1}^{(\pm 1)} - \frac{l}{(2l+1)} \kappa_l^{(1)} j_{l-1}^{(\pm 1)} \right] + \frac{4}{3} \frac{k}{a} \kappa_1^{(1)} A^{(\pm 1)} \delta_{l1} + \frac{8}{15} \frac{k}{a} \kappa_1^{(1)} \kappa_2^{(1)} \sigma^{(\pm 1)} \delta_{l2} \\ = -n_e \sigma_T \left[j_l^{(\pm 1)} - \frac{4}{3} \kappa_1^{(1)} v^{(\pm 1)} \delta_{l1} - \left(\frac{1}{10} j_2^{(\pm 1)} - \frac{3}{5} \mathcal{E}_2^{(\pm 1)} \right) \delta_{l2} \right], \end{aligned} \quad (4.3.78)$$

³⁷ It is possible to construct solutions with compensating singular vortices in the neutrinos and radiation that leave the perturbations to the spacetime geometry non-singular (e.g. Ref. [226]), but these are rather contrived.

for $l \geq 1$. Here, $v^{(\pm 1)}$ and $\sigma^{(\pm 1)}$ are the harmonic expansions of the baryon peculiar velocity and shear respectively:

$$v_a = \sum_k v^{(\pm 1)} \mathcal{Q}_a^{(\pm 1)}, \quad \sigma_{ab} = \sum_k \frac{k}{a} \sigma^{(\pm 1)} \mathcal{Q}_{ab}^{(\pm 1)}. \quad (4.3.79)$$

The $l = 1$ moment I_a is generated from the projected gradient of the photon energy density by advection. For vector perturbations, Δ_a is non-vanishing in a general gauge. Indeed, the linearised identity

$$\text{curl } D_a \rho = -2\dot{\rho} \omega_a \quad (4.3.80)$$

relates the curl of Δ_a for any species to the vorticity and the time evolution of that species. Of course, if the frame is chosen to be hypersurface orthogonal, $\omega_a = 0$ and $\Delta_a = 0$ for vector perturbations. However, in general we have a non-zero Δ_a with

$$\sqrt{1 + \frac{2K}{k^2}} \Delta^{(\pm 1)} = -\frac{8}{3} \frac{a}{k} \Theta \omega^{(\mp 1)}, \quad (4.3.81)$$

where $\Delta_a = -\sum_k k \Delta^{(\pm 1)} \mathcal{Q}_a^{(\pm 1)}$ and $\omega_a = \sum_k (k/a) \omega^{(\pm 1)} \mathcal{Q}_a^{(\pm 1)}$. Note that since Δ_a and ω_a are related by a curl, Eq. (4.3.81) links the coefficients of vector harmonics of opposite parity. In Eq. (4.3.78), it should be understood that the quantity $I_0^{(\pm 1)} \equiv \Delta^{(\pm 1)}$; it can be considered a source term since it is algebraic in the vorticity.

The integral solution for the vector-mode anisotropies in a general almost-FLRW model is [229]

$$I_l^{(\pm 1)} = \frac{4l}{\sqrt{v^2 + 2}} \int^{tr} dt e^{-\tau} \left\{ \left[\frac{k}{a} \left(\frac{1}{4} \Delta^{(\pm 1)} - A^{(\pm 1)} \right) + n_e \sigma_T v^{(\pm 1)} \right] \frac{\Phi_l^v(x)}{\sinh x} - \left[\frac{k}{a} \sigma^{(\pm 1)} + \frac{3}{16} n_e \sigma_T \left(\kappa_1^{(1)} \kappa_2^{(1)} \right)^{-1} \left(I_2^{(\pm 1)} - 6\mathcal{E}_2^{(\pm 1)} \right) \right] \frac{1}{\sqrt{v^2 + 2}} \frac{d}{dx} \left(\frac{\Phi_l^v(x)}{\sinh x} \right) \right\}. \quad (4.3.82)$$

This is valid for $l \geq 1$ in an open universe; the equivalent result for $K > 0$ follows from the usual replacements of $v^2 + n$ with $v^2 - n$ and hyperbolic functions by their trigonometric counterparts. The geometric terms $\Phi_l^v(x)/\sinh x$ arise from the projection of sources going like $e^a \mathcal{Q}_a^{(\pm 1)}$ along the line of sight (see Eq. (4.1.47)). The terms involving the derivative of this function arise from shear and scattering sources going like $e^a e^b \mathcal{Q}_{ab}^{(\pm 1)}$, as follows from Eq. (4.1.46). If we integrate the shear term by parts in Eq. (4.3.82) we find

$$\frac{1}{4} I_l^{(\pm 1)} - \frac{1}{3} \kappa_1^{(1)} \sigma^{(\pm 1)} \delta_{l1} = \frac{l}{\sqrt{v^2 + 2}} \int^{tr} dt e^{-\tau} \left\{ \left[\frac{k}{a} \left(\frac{1}{4} \Delta^{(\pm 1)} - A^{(\pm 1)} \right) - \dot{\sigma}^{(\pm 1)} \right] \frac{\Phi_l^v(x)}{\sinh x} + n_e \sigma_T \left(v^{(\pm 1)} - \sigma^{(\pm 1)} \right) \frac{\Phi_l^v(x)}{\sinh x} + \frac{3}{16} n_e \sigma_T \left(\kappa_1^{(1)} \kappa_2^{(1)} \right)^{-1} \left(I_2^{(\pm 1)} - 6\mathcal{E}_2^{(\pm 1)} \right) \frac{1}{\sqrt{v^2 + 2}} \frac{d}{dx} \left(\frac{\Phi_l^v(x)}{\sinh x} \right) \right\}. \quad (4.3.83)$$

Each term in this expression can be shown to frame-invariant. We see that the anisotropy is sourced at last scattering by the baryon velocity in a zero-shear frame (i.e. by the Doppler effect) and by anisotropic Thomson scattering. There is also an integrated effect which involves

$$\frac{k}{a} \left(\frac{1}{4} \Delta^{(\pm 1)} - A^{(\pm 1)} \right) - \dot{\sigma}^{(\pm 1)} = -\frac{2}{\sqrt{1 + 2K/k^2}} \dot{H}^{(\mp 1)}, \quad (4.3.84)$$

where the vector-mode contribution to the magnetic part of the Weyl tensor is

$$H_{ab} = \sum_k \left(\frac{k}{a} \right)^2 H^{(\pm 1)} \mathcal{Q}_{ab}^{(\pm 1)}, \quad (4.3.85)$$

and we have used Eq. (4.3.81), the vorticity propagation equation and the magnetic Weyl equation. The integrated contribution to the vector-mode temperature anisotropies is similar to the scalar-mode ISW effect, but involves the gravito-magnetic part of the Weyl curvature rather than gravito-electric. Eq. (4.3.83) extends the analysis of Ref. [226] to non-flat models; the velocity source term and the integrated contribution are consistent with the gauge-invariant analysis of Ref. [221] though the (small) anisotropic scattering term is not included there.

The polarization multipole equations for vector modes are

$$\begin{aligned} \dot{\xi}_l^{(\pm 1)} + \frac{k}{a} \left[\frac{(l-1)l(l+2)(l+3)}{(l+1)^3(2l+1)} \kappa_{l+1}^{(1)} \mathcal{E}_{l+1}^{(\pm 1)} - \frac{l}{(2l+1)} \kappa_l^{(1)} \mathcal{E}_{l-1}^{(\pm 1)} \right] - \frac{2}{l(l+1)} \frac{k}{a} \sqrt{1 + \frac{2K}{k^2}} \mathcal{B}_l^{(\mp 1)} \\ = -n_e \sigma_T \left[\mathcal{E}_l^{(\pm 1)} + \left(\frac{1}{10} I_2^{(\pm 1)} - \frac{3}{5} \mathcal{E}_2^{(\pm 1)} \right) \delta_{l2} \right], \end{aligned} \quad (4.3.86)$$

$$\begin{aligned} \dot{\mathcal{B}}_l^{(\pm 1)} + \frac{k}{a} \left[\frac{(l-1)l(l+2)(l+3)}{(l+1)^3(2l+1)} \kappa_{l+1}^{(1)} \mathcal{B}_{l+1}^{(\pm 1)} - \frac{l}{(2l+1)} \kappa_l^{(1)} \mathcal{B}_{l-1}^{(\pm 1)} \right] + \frac{2}{l(l+1)} \frac{k}{a} \sqrt{1 + \frac{2K}{k^2}} \mathcal{E}_l^{(\mp 1)} \\ = -n_e \mathcal{B}_l^{(\pm 1)}. \end{aligned} \quad (4.3.87)$$

Note here how the curl couplings between E and B -modes in Eqs. (4.3.45) and (4.3.46) lead to a coupling between the $\mathcal{E}_l^{(\pm 1)}$ and $\mathcal{B}_l^{(\mp 1)}$. If we choose the primordial perturbation variable $\phi_{(k)}^{(\pm 1)}$ so that the radiation anisotropies $I_l^{(\pm 1)}$ are linear in them (for example, $\phi_{(k)}^{(\pm 1)} = \sigma_{(k)}^{(\pm 1)} + v_{(k)}^{(\pm 1)}$), $\mathcal{B}_l^{(\pm 1)}$ will be linear in the primordial perturbation of the opposite parity, $\phi_{(k)}^{(\mp 1)}$. It follows that, if the primordial fluctuations are parity-symmetric in the mean, there will be no correlations between \mathcal{B}_{A_l} and either of \mathcal{E}_{A_l} or I_{A_l} .

The integral solutions for the polarization are [229]

$$\mathcal{E}_l^{(\pm 1)} = -\frac{3l(l-1)}{4(l+1)} \frac{1}{(v^2+2)} \int^{tr} dt n_e \sigma_T e^{-\tau} (\kappa_1^{(1)} \kappa_2^{(1)})^{-1} \left[(I_2^{(\pm 1)} - 6\mathcal{E}_2^{(\pm 1)}) \times \left(\frac{1}{\sinh x} \frac{d}{dx} \Phi_l^v(x) + \frac{\cosh x}{\sinh^3 x} \Phi_l^v(x) \right) \right], \quad (4.3.88)$$

$$\mathcal{B}_l^{(\pm 1)} = \frac{3l(l-1)}{4(l+1)} \frac{v}{(v^2+2)} \int^{tr} dt n_e \sigma_T e^{-\tau} (\kappa_1^{(1)} \kappa_2^{(1)})^{-1} (I_2^{(\mp 1)} - 6\mathcal{E}_2^{(\mp 1)}) \frac{\Phi_l^v(x)}{\sinh x}, \quad (4.3.89)$$

in an open universe. Examples of CMB power spectra from vector modes are given in Figure 2 of Ref. [226].

4.3.6. Tensor perturbations

The imprint of tensor perturbations, or gravitational waves, is implicit in the original work of Sachs and Wolfe [102], although the first detailed calculations for temperature were reported in Ref. [253] and for polarization in Ref. [254]. Other important milestones include constraints on the gravitational wave background from the large-angle temperature anisotropy from COBE [306–308], the introduction of the E - B decomposition (which was already implicit in the early work of Dautcourt and Rose [284]) and the realisation that B -mode polarization is a particularly sensitive probe of tensor modes [275, 276], and constraints from the WMAP data [302]. The effect of tensor modes on the CMB from the 1+3-covariant perspective is discussed in Refs. [169,224].

For tensor perturbations we expand the temperature and polarization multipoles in the $\mathcal{Q}_{A_l}^{(\pm 2)}$:

$$I_{A_l} = I \sum_k \left(\prod_{n=2}^l \kappa_n^{(2)} \right)^{-1} I_l^{(\pm 2)} \mathcal{Q}_{A_l}^{(\pm 2)}, \quad l \geq 2, \quad (4.3.90)$$

$$\mathcal{E}_{A_l} = I \sum_k \left(\prod_{n=2}^l \kappa_n^{(2)} \right)^{-1} \mathcal{E}_l^{(\pm 2)} \mathcal{Q}_{A_l}^{(\pm 2)}, \quad l \geq 2, \quad (4.3.91)$$

$$\mathcal{B}_{A_l} = I \sum_k \left(\prod_{n=2}^l \kappa_n^{(2)} \right)^{-1} \mathcal{B}_l^{(\pm 2)} \mathcal{Q}_{A_l}^{(\pm 2)}, \quad l \geq 2, \quad (4.3.92)$$

where implicit in the sum over harmonics, \sum_k , is a sum over the parity states labelled ± 2 . As for vector modes, we must now include the B -mode polarization. The linearised multipole hierarchy for the temperature anisotropies, Eq. (4.3.48), becomes

$$\begin{aligned} i_l^{(\pm 2)} + \frac{k}{a} \left[\frac{(l+3)(l-1)}{(l+1)(2l+1)} \kappa_{l+1}^{(2)} i_{l+1}^{(\pm 2)} - \frac{l}{(2l+1)} \kappa_l^{(2)} i_{l-1}^{(\pm 2)} \right] + \frac{8}{15} \frac{k}{a} \kappa_2^{(2)} \sigma^{(\pm 2)} \delta_{l2} \\ = -n_e \sigma_T \left[I_l^{(\pm 2)} - \left(\frac{1}{10} I_2^{(\pm 2)} - \frac{3}{5} \mathcal{E}_2^{(\pm 2)} \right) \delta_{l2} \right], \end{aligned} \quad (4.3.93)$$

for $l \geq 2$, where $\sigma_{ab} = \sum_k (k/a) \sigma^{(\pm 2)} \mathcal{Q}_{ab}^{(\pm 2)}$. For the polarization, we have

$$\begin{aligned} \dot{\mathcal{E}}_l^{(\pm 2)} + \frac{k}{a} \left[\frac{(l+3)^2(l-1)^2}{(2l+1)(l+1)^3} \kappa_{l+1}^{(2)} \mathcal{E}_{l+1}^{(\pm 2)} - \frac{l}{(2l+1)} \kappa_l^{(2)} \mathcal{E}_{l-1}^{(\pm 2)} \right] - \frac{4}{l(l+1)} \frac{k}{a} \sqrt{1 + \frac{3K}{k^2}} \mathcal{B}_l^{(\mp 2)} \\ = -n_e \sigma_T \left[\mathcal{E}_l^{(\pm 2)} + \left(\frac{1}{10} I_2^{(\pm 2)} - \frac{3}{5} \mathcal{E}_2^{(\pm 2)} \right) \delta_{l2} \right], \end{aligned} \quad (4.3.94)$$

$$\dot{\mathcal{B}}_l^{(\pm 2)} + \frac{k}{a} \left[\frac{(l+3)^2(l-1)^2}{(2l+1)(l+1)^3} \kappa_{l+1}^{(2)} \mathcal{B}_{l+1}^{(\pm 2)} - \frac{l}{(2l+1)} \kappa_l^{(2)} \mathcal{B}_{l-1}^{(\pm 2)} \right] + \frac{4}{l(l+1)} \frac{k}{a} \sqrt{1 + \frac{3K}{k^2}} \mathcal{E}_l^{(\mp 2)} = -n_e \mathcal{B}_l^{(\pm 2)}. \quad (4.3.95)$$

Note, again, how the curl couplings between E and B -modes in Eqs. (4.3.45) and (4.3.46) lead to a coupling between the $\mathcal{E}_l^{(\pm 2)}$ and $\mathcal{B}_l^{(\mp 2)}$ so there will be no correlations between \mathcal{B}_{A_l} and either of \mathcal{E}_{A_l} or I_{A_l} for fluctuations that are parity-invariant in the mean.

The integral solution for the tensor contribution to the intensity anisotropy in a general almost-FLRW model is [224,229]

$$I_l^{(\pm 2)} = \frac{4l(l-1)}{[(v^2+1)(v^2+3)]^{1/2}} \int^{tr} dt e^{-\tau} \left(-\frac{k}{a} \sigma^{(\pm 2)} + \frac{3}{16} n_e \sigma_T (\kappa_2^{(2)})^{-1} (I_2^{(\pm 2)} - 6\mathcal{E}_2^{(\pm 2)}) \right) \frac{\Phi_l^v(x)}{\sinh^2 x}, \quad (4.3.96)$$

in an open universe. The sources for the tensor-mode Boltzmann equation are of the form $\mathcal{Q}_{ab}^{(\pm 2)} e^a e^b$ and the projection of these at x along the line of sight gives rise to the geometric factor $l(l-1)\phi_l^v(x)/\sinh^2 x$; see Eq. (4.1.47). The solution for the polarization is

$$\mathcal{E}_l^{(\pm 2)} = \frac{-3M_l^2}{8[(v^2+1)(v^2+3)]^{1/2}} \int^{tr} dt n_e \sigma_T e^{-\tau} (\kappa_2^{(2)})^{-1} (I_2^{(\pm 2)} - 6\mathcal{E}_2^{(\pm 2)}) \phi_l^v(x), \quad (4.3.97)$$

$$\mathcal{B}_l^{(\pm 2)} = \frac{-3M_l^2}{8[(v^2+1)(v^2+3)]^{1/2}} \int^{tr} dt n_e \sigma_T e^{-\tau} (\kappa_2^{(2)})^{-1} (I_2^{(\mp 2)} - 6\mathcal{E}_2^{(\mp 2)}) \psi_l^v(x), \quad (4.3.98)$$

where the geometric terms $\phi_l^v(x)$ and $\psi_l^v(x)$ are

$$\phi_l^v(x) = \frac{d^2}{dx^2} \Phi_l^v(x) + 4 \coth x \frac{d}{dx} \Phi_l^v(x) - (v^2 - 1 - 2 \coth^2 x) \Phi_l^v(x), \quad (4.3.99)$$

$$\psi_l^v(x) = -2v \left[\frac{d}{dx} \Phi_l^v(x) + 2 \coth x \Phi_l^v(x) \right] \quad (4.3.100)$$

in an open universe. The results for $K > 0$ follow from the usual replacements.

We relate the radiation multipoles to the gauge-invariant primordial metric perturbation \mathcal{H}_{ab} of Section 4.2.2 via transfer functions

$$I_l^{(\pm 2)} = T_l^T(k) \mathcal{H}_{(k)}^{(\pm 2)}, \quad \mathcal{E}_l^{(\pm 2)} = \frac{M_l}{\sqrt{2}} T_l^E(k) \mathcal{H}_{(k)}^{(\pm 2)}, \quad \mathcal{B}_l^{(\pm 2)} = \frac{M_l}{\sqrt{2}} T_l^B(k) \mathcal{H}_{(k)}^{(\mp 2)}, \quad (4.3.101)$$

where it should be noted that $\mathcal{B}_l^{(\pm 2)}$ are linear in $\mathcal{H}_{(k)}^{(\mp 2)}$. The primordial tensor power spectrum is defined so that

$$\langle \mathcal{H}_{ab} \mathcal{H}^{ab} \rangle = \int \frac{vdv}{(v^2+1)} \mathcal{P}_{\mathcal{H}}(k), \quad (4.3.102)$$

and the non-vanishing CMB power spectra from tensor modes are then given by Eq. (4.3.63) with $m = 2$.

Examples of CMB power spectra from tensor modes are shown in the right-hand panel of Fig. 4. CMB anisotropies are produced by the anisotropic expansion of gravitational waves along the line of sight after last-scattering. Since gravitational waves damp away once they become sub-Hubble, with amplitude falling as $1/a$, the CMB spectra fall away rapidly above $l \sim 100$ (corresponding to the horizon size at recombination). Gravitational waves produce roughly equal power in E - and B -mode polarization, but with a somewhat sharper projection between linear scales k and angular scales l for E -modes than B [228]. As for scalar perturbations, reionization produces additional large-angle polarization.

The ratio of primordial tensor modes to curvature perturbations is set to $r = 0.20$ in Fig. 4, corresponding to the current upper limit on tensor modes from a combination of WMAP temperature and E -mode polarization data, and distance measures from the baryon acoustic oscillations in large-scale structure data and supernovae [302]. The constraint comes mainly from the (large-angle) temperature anisotropies. Simple inflation models naturally produce a background of gravitational waves with an almost scale-invariant power spectrum [309]

$$\mathcal{P}_{\mathcal{H}}(k) = \frac{16}{\pi} \left(\frac{H}{m_{\text{Pl}}} \right)^2, \quad (4.3.103)$$

where H is the Hubble parameter during inflation and $m_{\text{Pl}} = 1.22 \times 10^{19} \text{ GeV}/c^2$ is the Planck mass. The current upper limit on tensor modes implies $H < 1.2 \times 10^{14} \text{ GeV}$ and so an energy scale of inflation $E_{\text{inf}} < 2.2 \times 10^{16} \text{ GeV}$. The B -mode of polarization is yet to be observed and the constraints B -modes place on gravitational waves are not yet competitive with the temperature anisotropies. However, because of cosmic variance from the dominant scalar perturbations, the constraint on tensor modes from the temperature anisotropies will not improve much as instruments get more sensitive but that from B -modes will. Ratios $r \sim 10^{-2}$ are targets for a new generation of sensitive B -mode surveys from ground and balloon-borne platforms. Ultimately, the detection of tensor modes via the B -mode polarization may only be limited by our ability to control instrumental systematic effects (e.g. [310,311]) and clean out foreground contaminants [312] and secondary processes such as gravitational lensing [313,314].

4.3.7. Cosmic microwave background in other cosmological models

Anisotropies in the CMB in other cosmological models have been considered by a number of authors. A very general approach to constrain departures from FLRW symmetries with the CMB was pioneered by Ref. [255], subsequently improved in Ref. [256], and applied to COBE data in Ref. [257] (with an erratum in Ref. [315]). The idea was to use observational constraints on the F_{A_i} plus the 1 + 3-covariant dynamical equations to constrain the geometry in a model-independent way. Given that we can only measure the anisotropies here and now, assumptions about the size of temporal and spatial derivatives are required to extract useful constraints. For the large-angle anisotropies, the above papers assumed the expansion-normalised derivatives of the multipoles were bounded by the multipoles themselves. This is indeed the case

for almost FLRW models, but examples are known where the assumption is violated and an isotropic CMB observed now can still be accompanied by large spacetime anisotropy [316].

The CMB has also been investigated extensively in models with a subset of FLRW symmetry. Spatially-homogeneous but (globally-)anisotropic models were considered in the pioneering work of Collins & Hawking [317] and its subsequent extensions (e.g. Ref. [318]). Bianchi models that can be considered small perturbations of FLRW are particularly well studied and the anisotropies have a rich phenomenology due to geodesic focusing and spiralling. Constraints on global rotation and shear in these early papers were superseded by analysis based on the COBE data [319,320] for Bianchi VII_h models. More recently, a curious correlation has been found between the anisotropy template in this model and the WMAP data [321]. The result is statistically-significant, and removal of the correlated pattern can explain a number of anomalous features in the WMAP data. However, the cosmological parameters required are at odds with those needed to explain the CMB anisotropy on smaller scales and, moreover, the predicted polarization anisotropy has now been computed and it exceeds the WMAP polarization observations on large angular scales [322].

Spherically-symmetric models have also received considerable attention. They provide a useful analytic model of the secondary anisotropies due to non-linear gravitational effects (such as the imprint of a forming cluster of galaxies) [274, 323] and a simple way of relaxing the Copernican assumption.

5. Beyond the linear regime

Linear perturbation theory is a good approximation only at the initial stages of gravitational collapse, when the density contrast is well below unity. Most of the observed structures in the universe, however, have density contrasts well in excess of unity. The density within a cluster of galaxies, for example, is between 10^2 and 10^3 times greater than the average density of the universe, while that of a galaxy is about 10^5 times larger. To understand the evolution of these objects we need to go beyond the limits of the linear regime.

5.1. Nonlinear peculiar kinematics

In an unperturbed, idealised Friedmann universe, comoving particles have velocities that follow Hubble's law. When perturbations are present, however, the Hubble flow is distorted and matter acquires 'peculiar' velocities. The dipole anisotropy of the CMB seems to suggest that our Local Group is moving with respect the smooth Hubble flow, which defines the frame where the CMB dipole vanishes, at a speed of approximately 600 km/s (e.g. see [139]). Such, rather large, velocity perturbations can have important implications for any nonlinear structure formation scenario. After recombination and on scales well inside the Hubble length, one can use the Newtonian theory to study peculiar velocities. As we are successively probing significant fractions of the Hubble radius, however, the need for a relativistic treatment increases.

5.1.1. 1 + 3 peculiar-velocity decomposition

When studying peculiar motions one needs to define the associated velocities relative to a preferred reference frame. The latter is not comoving with the fluid, since there are no peculiar velocities relative to the matter frame by construction. Following [324–326], we choose our reference velocity field (u_a) to be both irrotational and shear-free. Then, we assume the presence of matter moving with 4-velocity

$$\tilde{u}_a = \gamma(u_a + v_a), \quad (5.1.1)$$

where v_a is the peculiar velocity of the fluid (with $u_a v^a = 0$) and $\gamma = (1 - v^2)^{-1/2}$ is the Lorentz-boost factor (see also Section 2.4.1). The $u_a v^a = 0$ condition guarantees that \tilde{u}_a is also timelike (i.e. that $\tilde{u}_a \tilde{u}^a = -1$), irrespective of the value of the γ -factor. Here, however, we will be dealing with non-relativistic peculiar motions, which means that $v^2 \ll 1$ and consequently that $\gamma \simeq 1$. The instantaneous rest space of the 'tilded' observer is defined by means of the tensor

$$\tilde{h}_{ab} = g_{ab} + \tilde{u}_a \tilde{u}_b, \quad (5.1.2)$$

which projects orthogonal to \tilde{u}_a .³⁸ When the u_a -frame has $\sigma_{ab} = 0 = \omega_{ab}$, it corresponds to Bardeen's quasi-Newtonian gauge (see [103]) and is related to the comoving (Lagrangian) reference system via the transformation laws given in Appendix A.2.

The peculiar kinematics are covariantly determined by the irreducible variables of the motion, obtained by decomposing the gradient of the v_a -field. To simplify the equations we choose as our time direction the one along \tilde{u}_a . Then, $\tilde{D}_a = \tilde{h}_a^b \nabla_b$ defines the associated orthogonally projected covariant derivative operator and the projected gradient of the peculiar velocity splits as

$$\tilde{D}_b v_a = \frac{1}{3} \hat{\Theta} \tilde{h}_{ab} + \hat{\sigma}_{ab} + \hat{\omega}_{ab}, \quad (5.1.3)$$

³⁸ The peculiar velocity field is orthogonal to u_a by construction but does not lie in the rest frame of the 'tilded' observers, even for non-relativistic peculiar velocities. Indeed, following (5.1.1) and (5.1.2), we find that $\tilde{u}_a v^a = \gamma v^2 \neq 0$ and that $\tilde{h}_a^b v_b = v_a + \gamma v^2 \tilde{u}_a \neq v_a$.

with $\hat{\Theta} = \tilde{D}_a v^a$, $\hat{\sigma}_{ab} = \tilde{D}_{[b} v_{a]}$ and $\hat{\omega}_{ab} = \tilde{D}_{[b} v_{a]}$. In analogy with $D_b u_a$ (see Section 1.3.1), the tensor $\tilde{D}_b v_a$ describes the relative motion of neighbouring peculiar flow lines, while $\hat{\Theta}$, $\hat{\sigma}_{ab}$ and $\hat{\omega}_{ab}$ represent the volume expansion (or contraction), the shear and the vorticity of the peculiar motion respectively. Thus, a region that has ‘decoupled’ from the background expansion and is collapsing has $\hat{\Theta} < 0$.

5.1.2. Nonlinear peculiar motions

In cosmological studies, it helps to identify the preferred 4-velocity field, that is the reference u_a -frame of the previous section, with the one setting the CMB dipole to zero. Then, an observer moving with the matter (i.e. with the \tilde{u}_a -frame of Section 5.1.1) monitors its motion using the nonlinear evolution equation of the peculiar velocity field

$$\dot{v}_{(a)} = \tilde{A}_a - A_{(a)} - \frac{1}{3} (\tilde{\Theta} - \hat{\Theta}) v_{(a)} - (\tilde{\sigma}_{ab} - \hat{\sigma}_{ab}) v^b - (\tilde{\omega}_{ab} - \hat{\omega}_{ab}) v^b, \quad (5.1.4)$$

with $\dot{v}_{(a)} = \tilde{h}_a^b \dot{v}_b$, $A_{(a)} = \tilde{h}_a^b A_b$ and $v_{(a)} = \tilde{h}_a^b v_b$, and that of its projected gradients. The latter, which is obtained by applying the Ricci identity to v_a (see [326] for the technical details), reads

$$\begin{aligned} \tilde{h}_b^d \tilde{h}_a^c (\tilde{D}_d v_c) &= \tilde{D}_b \dot{v}_a - \frac{1}{9} \tilde{\Theta} \tilde{\Theta} \tilde{h}_{ab} - \frac{1}{3} [\tilde{\Theta} (\tilde{\sigma}_{ab} + \hat{\omega}_{ab}) + \hat{\Theta} (\tilde{\sigma}_{ab} + \tilde{\omega}_{ab})] - \hat{\sigma}_{ca} \tilde{\sigma}^c_b - \hat{\sigma}_{ca} \tilde{\omega}^c_b + \hat{\omega}_{ca} \tilde{\sigma}^c_b + \hat{\omega}_{ca} \tilde{\omega}^c_b \\ &+ \dot{v}_{(a)} \tilde{A}_b - \frac{1}{3} \tilde{\Theta} \tilde{A}_a v_{(b)} - \tilde{A}_a (\tilde{\sigma}^c_b + \tilde{\omega}^c_b) v_c + \tilde{A}_a \tilde{D}_b v^2 - \tilde{h}_a^c \tilde{h}_b^d R_{csdq} v^s \tilde{u}^a, \end{aligned} \quad (5.1.5)$$

where $\tilde{A}_a = \tilde{\dot{u}}_a = \tilde{u}^b \nabla_b \tilde{u}_a$ is the 4-acceleration of the ‘tilded’ observers. Note the Riemann-curvature term in the right-hand side of the above, which couples the geometry of the spacetime with both \tilde{u}_a and v_a . As we will see next, the latter coupling has profound implications for the kinematics of peculiar motions.

The trace, the symmetric trace-free component and the skew part of the above lead to the propagation equations of $\hat{\Theta}$, $\hat{\sigma}_{ab}$ and $\hat{\omega}_{ab}$ respectively [326]. Thus, the time evolution of the volume scalar is monitored by

$$\begin{aligned} \dot{\hat{\Theta}} &= -\frac{1}{3} \tilde{\Theta} \hat{\Theta} - \tilde{\sigma}_{ab} \hat{\sigma}^{ab} + \tilde{\omega}_{ab} \hat{\omega}^{ab} + \tilde{D}_a \dot{v}^a + \tilde{A}_a \dot{v}^a + \tilde{q}_a v^a + \left[\frac{1}{2} (\tilde{\rho} + 3\tilde{p}) - \Lambda \right] v^2 \\ &- \left(\tilde{\sigma}_{ab} - \tilde{\omega}_{ab} + \frac{1}{3} \tilde{\Theta} \tilde{h}_{ab} \right) \tilde{A}^a v^b + \tilde{A}^a \tilde{D}_a v^2, \end{aligned} \quad (5.1.6)$$

which is the Raychaudhuri analogue of the peculiar flow and shares several close analogies with its standard counterpart (compare to Eq. (1.3.3) in Section 1.3.1).³⁹ There are differences, however, and the main one is in the role played by the matter fields. According to the seventh term in the right-hand side of (5.1.6), the total gravitational mass of the system tends to increase the average separation between two neighbouring peculiar flow lines. A positive cosmological constant, on the other hand, brings the aforementioned flow lines closer together.

Taking the symmetric trace-free and the antisymmetric parts of (5.1.5) leads to the respective propagation formulae of the shear and the vorticity. Following [326], we have

$$\begin{aligned} \tilde{h}_{(b}^d \tilde{h}_{a)}^c \dot{\tilde{\sigma}}_{cd} &= -\frac{1}{3} (\tilde{\Theta} \hat{\sigma}_{ab} + \hat{\Theta} \tilde{\sigma}_{ab}) - \hat{\sigma}_{c[a} \tilde{\sigma}^c_{b]} - \hat{\sigma}_{c[a} \tilde{\omega}^c_{b]} + \hat{\omega}_{c[a} \tilde{\sigma}^c_{b]} + \hat{\omega}_{c[a} \tilde{\omega}^c_{b]} + \tilde{D}_{(b} \dot{v}_{a)} + \tilde{A}_{(a} \tilde{h}_{b)}^c \dot{v}_c - \frac{1}{3} \tilde{\Theta} \tilde{A}_{(a} \tilde{h}_{b)}^c v_c \\ &- \tilde{A}_{(a} (\tilde{\sigma}^c_{b]} + \tilde{\omega}^c_{b]}) v_c + \tilde{A}_{(a} \tilde{D}_{b)} v^2 + v^2 \tilde{E}_{ab} - \tilde{\varepsilon}_{cd(a} v^c \tilde{H}^d_{b)} - \frac{1}{2} v^2 \tilde{\pi}_{ab} - \frac{1}{2} \tilde{q}_{(a} \tilde{h}_{b)}^c v_c, \end{aligned} \quad (5.1.7)$$

for the ‘peculiar shear’, and

$$\begin{aligned} \tilde{h}_{[b}^d \tilde{h}_{a]}^c \dot{\tilde{\omega}}_{cd} &= -\frac{1}{3} (\tilde{\Theta} \hat{\omega}_{ab} + \hat{\Theta} \tilde{\omega}_{ab}) - \hat{\sigma}_{c[a} \tilde{\sigma}^c_{b]} - \hat{\sigma}_{c[a} \tilde{\omega}^c_{b]} + \hat{\omega}_{c[a} \tilde{\sigma}^c_{b]} + \hat{\omega}_{c[a} \tilde{\omega}^c_{b]} + \tilde{D}_{[b} \dot{v}_{a]} - \tilde{A}_{[a} \tilde{h}_{b]}^c \dot{v}_c - \frac{1}{3} \tilde{\Theta} \tilde{A}_{[a} \tilde{h}_{b]}^c v_c \\ &- \tilde{A}_{[a} (\tilde{\sigma}^c_{b]} + \tilde{\omega}^c_{b]}) v_c + \tilde{A}_{[a} \tilde{D}_{b]} v^2 - \tilde{\varepsilon}_{cd[a} v^c \tilde{H}^d_{b]} - \frac{1}{2} \tilde{q}_{[a} \tilde{h}_{b]}^c v_c, \end{aligned} \quad (5.1.8)$$

for the peculiar vorticity. When compared to their conventional analogues (see Eqs. (1.3.4) and (1.3.5) in Section 1.3.1), these expressions show a high degree of complexity. Among others, the coupling between spacetime geometry and peculiar velocity seen in (5.1.5), has led to a extra (magnetic) Weyl effects on both the shear and the vorticity.

³⁹ We remind the reader that in this section overdots denote time-derivatives along \tilde{u}_a and angled brackets projection orthogonal to \tilde{u}_a . Also, ‘tilded’ quantities are measured in the matter (\tilde{u}_a) frame.

5.1.3. The case of dust

To this point, the peculiar velocity field is arbitrary and v_a has not been associated with any particular matter source. The system (5.1.4)–(5.1.8) is completely general and $\gamma \simeq 1$ is the only restriction imposed so far.

Consider now an almost-FLRW spacetime with a family of observers having a non-relativistic peculiar velocity relative to the comoving frame, which is irrotational and shear-free. This defines the covariant analogues of the quasi-Newtonian cosmologies (see also Section 5.1.1). In what follows we will identify the quasi-Newtonian frame with the reference 4-velocity field u_a and align \hat{u}_a with a non-relativistic (dust) component; a picture that reflects the post-recombination universe. In that case $\tilde{p} = 0 = \tilde{\pi}_{ab} = \tilde{A}_a$ and, given that $\tilde{q}_a = 0$ in the matter frame, expressions (5.1.4) and (5.1.6)–(5.1.8) reduce to

$$\dot{v}_a = -A_{(a)} - \frac{1}{3} (\tilde{\Theta} - \hat{\Theta})v_{(a)} - (\tilde{\sigma}_{ab} - \hat{\sigma}_{ab})v^b - (\tilde{\omega}_{ab} - \hat{\omega}_{ab})v^b, \quad (5.1.9)$$

$$\dot{\hat{\Theta}} = -\frac{1}{3} \tilde{\Theta} \hat{\Theta} - \tilde{\sigma}_{ab} \hat{\sigma}^{ab} + \tilde{\omega}_{ab} \hat{\omega}^{ab} + \tilde{D}_a \dot{v}^a + \left(\frac{1}{2} \tilde{\rho} - \Lambda \right) v^2, \quad (5.1.10)$$

$$\dot{\hat{\sigma}}_{ab} = -\frac{1}{3} \left(\tilde{\Theta} \hat{\sigma}_{ab} + \hat{\Theta} \tilde{\sigma}_{ab} \right) - \hat{\sigma}_{c[a} \tilde{\sigma}^c_{b]} - \hat{\sigma}_{c[a} \tilde{\omega}^c_{b]} + \hat{\omega}_{c[a} \tilde{\sigma}^c_{b]} + \hat{\omega}_{c[a} \tilde{\omega}^c_{b]} + \tilde{D}_{[b} \dot{v}_{a]} + v^2 \tilde{E}_{ab} - \tilde{\varepsilon}_{cd[a} v^c \tilde{H}^d_{b]}, \quad (5.1.11)$$

and

$$\dot{\hat{\omega}}_{ab} = -\frac{1}{3} \left(\tilde{\Theta} \hat{\omega}_{ab} + \hat{\Theta} \tilde{\omega}_{ab} \right) - \hat{\sigma}_{c[a} \tilde{\sigma}^c_{b]} - \hat{\sigma}_{c[a} \tilde{\omega}^c_{b]} + \hat{\omega}_{c[a} \tilde{\sigma}^c_{b]} + \hat{\omega}_{c[a} \tilde{\omega}^c_{b]} + \tilde{D}_{[b} \dot{v}_{a]} - \tilde{\varepsilon}_{cd[a} v^c \tilde{H}^d_{b]}, \quad (5.1.12)$$

respectively. Following (5.1.11), the effect of electric (the tidal) part of the Weyl field on the shear anisotropy of the peculiar motion is of higher perturbative order, compared to that of its magnetic counterpart. This will make a difference when taking the second-order limit of the above named expressions around a quasi-Newtonian background.

5.2. The mildly nonlinear regime

Linear perturbation theory provides an adequate description of the early stages of galactic collapse when the distortions from homogeneity and isotropy are still relatively small. As the perturbations grow stronger, however, the linear approximation brakes down and one needs to incorporate nonlinear effects. During the transition from the linear to the fully nonlinear regime, a period which one might call the “mildly nonlinear era”, one can monitor the perturbed quantities by means of the second-order equations instead of the fully nonlinear ones.

5.2.1. Second order evolution

Assuming an Einstein–de Sitter background, the linear relations between the quasi-Newtonian reference system and the matter frame are $\hat{\Theta} = \Theta + \hat{\Theta}$, $\hat{\sigma}_{ab} = \hat{\sigma}_{ab}$ and $\hat{\omega}_{ab} = \hat{\omega}_{ab}$. Also to first order, $\tilde{\rho} = \rho$, $\tilde{E}_{ab} = E_{ab}$ and $\tilde{H}_{ab} = 0$, while A_a can be shown to derive from a potential [325]. Then, Eq. (5.1.9) reduces to the linear expression

$$\dot{v}_a = -\frac{1}{3} \Theta v_a - \tilde{D}_a \phi, \quad (5.2.1)$$

where ϕ is the peculiar gravitational potential. Following [325], the above accepts a single growing mode with $v_a \propto a^{1/2}$ and has the same linear evolution in both the quasi-Newtonian and the matter frames. Given that, it helps to describe the peculiar motions by means of the rescaled velocity field $V_a = a^{-1/2} v_a$, which remains constant to leading order both in the reference and the matter frame (see [325,326] and also Section 5.2.2).

The mildly nonlinear regime monitors the early proto-galactic collapse of a dust cloud that has decoupled from the background expansion. On these grounds we may set $\Theta \ll \hat{\Theta}$. Then, on a quasi-Newtonian Einstein–de Sitter background, the nonlinear system (5.1.9)–(5.1.12) leads to the second-order set [326]

$$\dot{v}_a = -A_{(a)} - \frac{1}{3} \Theta v_{(a)}, \quad (5.2.2)$$

$$\dot{\hat{\Theta}} = -\frac{1}{3} \hat{\Theta}^2 - 2\hat{\sigma}^2 + 2\hat{\omega}^2 + \tilde{D}_a \dot{v}^a + \frac{1}{2} \rho v^2, \quad (5.2.3)$$

$$\dot{\hat{\sigma}}_{ab} = -\frac{2}{3} \hat{\Theta} \hat{\sigma}_{ab} - \hat{\sigma}_{c[a} \hat{\sigma}^c_{b]} + \hat{\omega}_{c[a} \hat{\omega}^c_{b]} + \tilde{D}_{[b} \dot{v}_{a]}, \quad (5.2.4)$$

$$\dot{\hat{\omega}}_{ab} = -\frac{2}{3} \hat{\Theta} \hat{\omega}_{ab} - 2\hat{\sigma}_{c[a} \hat{\omega}^c_{b]} + \tilde{D}_{[b} \dot{v}_{a]}. \quad (5.2.5)$$

Hence, at this perturbative level, both the electric and the magnetic Weyl tensors have decoupled and the peculiar kinematics evolve unaffected by the long-range gravitational field. The absence of the electric Weyl component in (5.2.4), in particular, is a consequence of the two-frame approach. Next, we will see the implications of this absence for the asymptotic final shape of a collapsing overdensity.

5.2.2. Rescaling the variables

For our purposes, the key linear result is that the peculiar velocity evolves as $v_a \propto a^{1/2}$. This then means that, to leading order, the rescaled velocity field $V_a = a^{-1/2}v_a$ remains constant (i.e. $\dot{V}_a = 0$) and that V_a comes from a potential and it is therefore non-rotating, namely that $\tilde{D}_{[b}V_{b]} = 0$ (see [325,326] for details). Hence, one can always assume that the linear peculiar velocity of the dust component is both acceleration-free and irrotational. To fully exploit this linear result we rescale our kinematical variables according to $\dot{\Theta} = a^{1/2}\bar{\Theta}$, $\dot{\sigma}_{ab} = a^{1/2}\bar{\sigma}_{ab}$ and $\dot{\omega}_{ab} = a^{1/2}\bar{\omega}_{ab}$, with $\bar{\Theta} = \tilde{D}_a V^a$, $\bar{\sigma}_{ab} = \tilde{D}_{(b}V_{a)}$ and $\bar{\omega}_{ab} = \tilde{D}_{[b}V_{a]}$. Then, the system (5.2.2)–(5.2.5) also rescales to [326]

$$\dot{V}_a = -a^{-1/2}A_{(a)} - \frac{1}{2}\Theta V_{(a)}, \quad (5.2.6)$$

$$\dot{\Theta} = -\frac{1}{3}a^{1/2}\bar{\Theta}^2 - 2a^{1/2}\bar{\sigma}^2 + 2a^{1/2}\bar{\omega}^2 + \tilde{D}_a \dot{V}^a + \frac{1}{2}a^{1/2}\rho V^2, \quad (5.2.7)$$

$$\dot{\sigma}_{ab} = -\frac{2}{3}a^{1/2}\bar{\Theta}\bar{\sigma}_{ab} - a^{1/2}\bar{\sigma}_{c(a}\bar{\sigma}^c_{b)} + a^{1/2}\bar{\omega}_{c(a}\bar{\omega}^c_{b)} + \tilde{D}_{(b}\dot{V}_{a)}, \quad (5.2.8)$$

$$\dot{\omega}_{ab} = -\frac{2}{3}a^{1/2}\bar{\Theta}\bar{\omega}_{ab} - 2a^{1/2}\bar{\sigma}_{c[a}\bar{\omega}^c_{b]} + \tilde{D}_{[b}\dot{V}_{a]}. \quad (5.2.9)$$

5.3. The relativistic Zeldovich approximation

Most analytical models of nonlinear structure formation are based on the assumption of spherical symmetry [327]—see also [139]. The spherical collapse scenario became popular because of its simplicity, but in reality it stops short from explaining key features of the observed universe. Galactic collapse does not seem to proceed isotropically. Galaxy surveys show complicated triaxial structures, which require non-spherical analysis if they were to be explained.

5.3.1. The Zeldovich ansatz

The Zeldovich approximation is not restricted to spherical symmetry and addresses the mildly nonlinear collapse of protogalactic clouds. It applies to scales well within the Hubble radius, as they decouple from the background expansion and ‘turn around’. The approximation works by extrapolating into the nonlinear regime the exact linear result of the acceleration-free and irrotational motion of the dust component. This considerably simplifies the equations and allows for analytical solutions.

When the Zeldovich ansatz $\dot{V}_a = 0 = \bar{\omega}_{ab}$ is applied to the second-order set (5.2.6)–(5.2.5), the motion of the collapsing pressure-free matter is determined by the reduced pair

$$\dot{\Theta} = -\frac{1}{3}a^{1/2}\bar{\Theta}^2 - 2a^{1/2}\bar{\sigma}^2 + \frac{1}{2}a^{1/2}\rho V^2, \quad (5.3.1)$$

$$\dot{\sigma}_{ab} = -\frac{2}{3}a^{1/2}\bar{\Theta}\bar{\sigma}_{ab} - a^{1/2}\bar{\sigma}_{c(a}\bar{\sigma}^c_{b)}. \quad (5.3.2)$$

Note the matter term in the right-hand side of (5.3.1). Given that V_a is constant and that $\rho \propto a^{-3}$ for dust, the impact of the background matter upon $\bar{\Theta}$ decays away. This means that the collapse is increasingly dominated by the kinematics and that gravity becomes progressively less important. The situation is closely analogous to that seen in studies of silent universes (e.g. see [328]), or during the Kasner regime of the Bianchi I cosmologies (see Section 1.5.2).

Referring the reader to [326] for the details of the relativistic analysis and also to [329] for the original (Newtonian) covariant treatment of the Zeldovich approximation, we introduce the new ‘time’ variable τ (constructed so that $\dot{\tau} = -a^{1/2}\bar{\Theta}$). The minus sign compensates for the fact that we are dealing with a collapsing region (i.e. $\bar{\Theta} < 0$) and guarantees that $\dot{\tau} > 0$ always. Note that $\tau \rightarrow \infty$ as we approach the singularity, where $\bar{\Theta} \rightarrow -\infty$. Then, the system (5.3.1) and (5.3.2), the former without the matter term, transforms into the set

$$\bar{\Theta}' = \frac{1}{3}\bar{\Theta} + 2\bar{\Theta}^{-1}\bar{\sigma}^2, \quad (5.3.3)$$

$$\bar{\sigma}'_{ab} = \frac{2}{3}\bar{\sigma}_{ab} + \bar{\Theta}^{-1}\bar{\sigma}_{c(a}\bar{\sigma}^c_{b)}, \quad (5.3.4)$$

where primes indicating differentiation with respect to τ . Our last step is to assume the shear eigenframe, where $\bar{\sigma}_{ab} = \text{diag}(\bar{\sigma}_1, \bar{\sigma}_2, \bar{\sigma}_3)$ with $\bar{\sigma}_3 = -(\bar{\sigma}_1 + \bar{\sigma}_2)$. Relative to this frame equations (5.3.3) and (5.3.4) lead to

$$\bar{\Theta}' = \frac{1}{3}\bar{\Theta} + 2\bar{\Theta}^{-1}(\bar{\sigma}_1^2 + \bar{\sigma}_2^2 + \bar{\sigma}_1\bar{\sigma}_2), \quad (5.3.5)$$

$$\bar{\sigma}'_1 = \frac{2}{3}\bar{\sigma}_1 + \frac{1}{3}\bar{\Theta}^{-1}\bar{\sigma}_1^2 - \frac{2}{3}\bar{\Theta}^{-1}(\bar{\sigma}_1 + \bar{\sigma}_2)\bar{\sigma}_2, \quad (5.3.6)$$

$$\bar{\sigma}'_2 = \frac{2}{3}\bar{\sigma}_2 + \frac{1}{3}\bar{\Theta}^{-1}\bar{\sigma}_2^2 - \frac{2}{3}\bar{\Theta}^{-1}(\bar{\sigma}_1 + \bar{\sigma}_2)\bar{\sigma}_1. \quad (5.3.7)$$

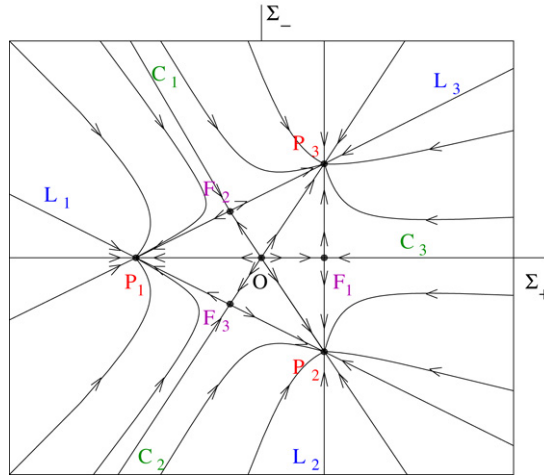


Fig. 5. Phase plane with $\Sigma_+ \equiv X$ and $\Sigma_- \equiv Y$. The lines L_1 , L_2 and L_3 that form the central triangle correspond to $\sigma_i = -\bar{\theta}/3$ ($i = 1, 2, 3$), with the three pancakes located at P_1 , P_2 and P_3 where these lines intersect. The points F_1 , F_2 , F_3 represent filamentary solutions and spindle-like singularities, while O corresponds to spherically symmetric, isotropic collapse (see [326] and also [329]).

with the behaviour of $\bar{\sigma}_3$ determined from that of $\bar{\sigma}_1$ and $\bar{\sigma}_2$ [326]. The above second-order system provides a fully covariant formulation of the Zeldovich approximation, which governs the small-scale evolution of pressure-free matter, as the latter decouples from the background expansion and starts to turn around and collapse.

5.3.2. The Zeldovich pancakes

The question is whether or not the relativistic analysis also predicts that one-dimensional pancakes are the final configurations of any generic overdensity. Given the qualitative nature of the question, one can employ a dynamical-systems approach to look for the answer. To begin with, consider the dimensionless variables (see [326] and also [329])

$$\Sigma_+ = \frac{3}{2} \bar{\theta}^{-1} (\bar{\sigma}_1 + \bar{\sigma}_2) \quad \text{and} \quad \Sigma_- = \frac{\sqrt{3}}{2} \bar{\theta}^{-1} (\bar{\sigma}_1 - \bar{\sigma}_2), \quad (5.3.8)$$

which measure the anisotropy of the collapse. Clearly, when both Σ_+ and Σ_- vanish, we are dealing with a spherically symmetric collapse. On introducing Σ_{\pm} , Eq. (5.3.5) transforms into

$$\bar{\theta}' = \frac{1}{3} \bar{\theta} + \frac{2}{3} \bar{\theta} (\Sigma_+^2 + \Sigma_-^2), \quad (5.3.9)$$

while expressions (5.3.6) and (5.3.7) become

$$\Sigma_+' = \frac{1}{3} [1 - \Sigma_+ - 2(\Sigma_+^2 + \Sigma_-^2)] \Sigma_+ + \frac{1}{3} \Sigma_-^2 \quad (5.3.10)$$

and

$$\Sigma_-' = \frac{1}{3} [1 + 2\Sigma_+ - 2(\Sigma_+^2 + \Sigma_-^2)] \Sigma_-, \quad (5.3.11)$$

respectively. Thus, the evolution of Σ_+ and Σ_- has decoupled from that of $\bar{\theta}$ and the shape of the collapsing overdensity is monitored by the subsystem (5.3.10) and (5.3.11). Technically speaking, the problem has been reduced to the study of the planar dynamical system depicted in Fig. 5. Physically, this dimensional reduction means that the shape of the collapsing dust cloud does not depend on the collapse timescale.

Referring to [326] for the technical details, we point out that the vertices P_1 , P_2 and P_3 of the triangle seen in Fig. 5 are stationary points of the system (5.3.10) and (5.3.11) and also act as attractors. Each one of the three vertices corresponds to an one-dimensional pancake solution, which is stationary along two directions and collapses along the third shear eigendirection. Generic solutions are asymptotic to one of these pancakes (one for each eigendirection). The bisecting lines C_1 , C_2 and C_3 intersect to the stationary point O that represents shear-free spherically symmetric collapse. Finally, where the bisecting lines intersect the triangle, we have the stationary points F_1 , F_2 and F_3 that correspond to exact filamentary solutions. The pancakes are stable nodes, the filaments are saddle points and the spherically symmetric collapse is an unstable node. In other words, once the collapse sets in, the pancakes are the natural attractors for a generic overdensity. This result is in disagreement with dynamical studies of silent universes, which argue for spindle-like rather than pancake singularities (see [328] and also [330]). The reason for this difference appears to lie in the role of the tidal field. Silent universes allow for a nonzero electric Weyl component, but set its magnetic counterpart to zero. In the presence of this “truncated” tidal

field, the collapsing overdensity appears to evolve towards a Kasner-type singularity where pancakes are a set of measure zero. The analysis of peculiar velocities in Einstein–de Sitter models, however, has shown that the long-range gravitational field has negligible input at second order. As a result, the relativistic equations have been reduced to the Newtonian ones and pancakes have been reinstated as the natural attractors of generic protogalactic collapse, this time within the realm of general relativity. This also appears in agreement with numerical simulations which also seem to favor pancake formation over all other types of singularities [331].

5.4. Averaging and backreaction

The increasing complexity of the equations, as one moves beyond the linear regime, means that analytical studies of nonlinear structure formation are always bound within the limits of certain approximations. So, a common feature in every (local) nonlinear analytical work is that some effects are sidestepped in favour of others. The spherical collapse scenario and the Zeldovich ansatz are probably the two best known examples. It therefore seems likely that only via a consistent averaging method one can achieve a tractable, all inclusive (and nonlocal) nonlinear study.

5.4.1. Spatial averaging

The question of whether an inhomogeneous spacetime behaves on average like a homogeneous solution of Einstein's equations is a long-standing issue, directly related to the problem of averaging general-relativistic spacetimes [332,333]. The reason is the generic nonlinearity of general relativity and the difficulty of establishing a unique and unambiguous way of averaging the spacetime metric, without throwing away crucial information during the process. This ongoing ambiguity has led a number of cosmologists to argue that the averaging problem may be crucial in our understanding of the recent expansion of the visible universe. Thus, when it was recently suggested that structure formation may be responsible for the acceleration of the expansion, averaging techniques were employed to study the dark energy problem in terms of kinematic backreaction effects from spatial inhomogeneities [25–35]. Although the backreaction idea has been criticised, primarily on the basis of current observations [334–341], it is attractive in principle because it can solve the coincidence problem without appealing to a cosmological constant, speculative quintessence fields, or nonlinear corrections to Einstein's gravity. For recent reviews on the backreaction question the reader is referred to [342,343].

General relativity has been applied to a range of averaging scales, depending on the scale of the physical system under study. Cosmology applies to the largest astrophysical scales, which typically extend over a significant fraction of the Hubble length (see [344] for a discussion and further references). Nevertheless, the major issue of defining a suitable averaging process remains open. The literature contains more than one averaging methods that have been applied to cosmology [345–349]. Here we will follow the Buchert approach, which builds on the Newtonian theory (where spatial averaging is relatively simple) before extending to general relativity by confining to scalar variables [347,348]. These averaged scalars become the effective dynamical sources that an observer will expect to measure. In the averaged equations one isolates an explicit source term, which is commonly referred to as the *backreaction term*. The latter quantifies the deviation from a given 'fitting model' that usually coincides with the Friedmann spacetime. From our point of view, the Buchert scheme has an additional advantage because it applies naturally to the 1 + 3 covariant formulae used here.

Following [347,348], the spatial averaging of an arbitrary scalar field $\phi = \phi(x^a)$, over a simply connected domain D is a covariant operation, defined by the averaging operator

$$\langle \phi \rangle_D = V_D^{-1} \int_D \phi \mathcal{H} d^3x^a, \quad (5.4.1)$$

where the angled brackets indicate spatial averaging, V_D is the volume of the domain in question and $\mathcal{H} = \sqrt{\det(h_{ab})}$. Since the volume is time-dependent in principle, the average of the time derivative of a locally defined scalar generally differs from the time derivative of the average. This non-commutativity between spatial averaging and temporal evolution can be formulated in a simple mathematical rule. For a rest-mass preserving domain the commutation between the volume-averaging and the time-averaging operators leads to

$$\langle \dot{\phi} \rangle_D - \dot{\langle \phi \rangle}_D = \langle \Theta \phi \rangle_D - \langle \Theta \rangle_D \langle \phi \rangle_D, \quad (5.4.2)$$

with Θ representing the local expansion rate. The right-hand side of the above generally introduces source terms in the evolution equations which can be interpreted as the backreaction effects of the averaging process. Such backreaction terms have recently been considered as a possible solution to the dark energy question.

5.4.2. The averaged equations

The first step is to write the average equations in the form of a chosen family of local equations, with any deviation between the two sets treated as a backreaction effect. Here, we will write the averaged formulae in a Friedmann-type form. To be precise, assuming irrotational dust moving along worldlines tangent to the 4-velocity u_a -field and applying rule (5.4.2)

to the continuity, the Friedmann and the Raychaudhuri equations we arrive at the system⁴⁰

$$\langle \rho \rangle + \langle \Theta \rangle \langle \rho \rangle = 0, \quad (5.4.3)$$

$$\langle \Theta \rangle + \frac{1}{3} \langle \Theta \rangle^2 + \frac{1}{2} \langle \rho \rangle - \Lambda = \frac{2}{3} \langle (\Theta - \langle \Theta \rangle)^2 \rangle - 2 \langle \sigma^2 \rangle, \quad (5.4.4)$$

$$\frac{1}{2} \langle \mathcal{R} \rangle - \langle \rho \rangle + \frac{1}{3} \langle \Theta \rangle^2 - \Lambda = -\frac{1}{3} \langle (\Theta - \langle \Theta \rangle)^2 \rangle + \langle \sigma^2 \rangle, \quad (5.4.5)$$

where the right-hand sides of (5.4.4) and (5.4.5) involve the identity $\langle (\phi - \langle \phi \rangle)^2 \rangle = \langle \phi^2 \rangle - \langle \phi \rangle^2$. The above provides a (non-closed) set of three effective Einstein equations for the spatially averaged scalars of non-rotating inhomogeneous universes containing pressure-free matter [25,350]. Within this environment, Eqs. (5.4.4) and (5.4.5) are exact and there is no need to assume that the inhomogeneity and the anisotropy are small perturbations. Note that, despite the non-commutativity between spatial averaging and temporal evolution, the averaged continuity equation has retained the form of its local counterpart. This means that $\langle \rho \rangle \propto a_D^{-3}$, where a_D is the average scale factor smoothed over the domain D and defined so that $\dot{a}_D/a_D = \langle \Theta \rangle/3$. However, the averaging process has led to extra terms in the right-hand sides of both (5.4.4) and (5.4.5), collectively given by the domain-dependent scalar

$$\mathcal{Q}_D = \frac{2}{3} \langle (\Theta - \langle \Theta \rangle)^2 \rangle - 2 \langle \sigma^2 \rangle = -\frac{2}{3} \left(\langle \Theta \rangle^2 - \langle \Theta^2 \rangle \right) - 2 \langle \sigma^2 \rangle. \quad (5.4.6)$$

We interpret this quantity, which in principle can be either positive or negative, as the kinematic backreaction of spatial averaging upon pressure-free FRW models, since in its absence the averaged formulae recover the form of their local counterparts.⁴¹ We also note that Eqs. (5.4.3)–(5.4.5) may be recast into the standard form of a spatially flat LFRW universe, provided the (effective) density and pressure are given by $\rho_{\text{eff}} = \langle \rho \rangle - (\mathcal{Q}_D + \langle \mathcal{R} \rangle)/2$ and $p_{\text{eff}} = -\mathcal{Q}_D/2 + \langle \mathcal{R} \rangle/6$ respectively [349]. Thus, given an effective equation of state, the set (5.4.3)–(5.4.5) is solvable and provides the scalar characteristics of the inhomogeneous universe on a given spatial scale.

5.5. Backreaction and accelerated expansion

The most striking feature of the Λ CDM model is that 95% of the matter in the universe today is in some unknown form. In fact, the bulk of the matter (approximately 70% of it) is in the form of dark energy, triggering the accelerated expansion. Nevertheless, we have not actually measured the cosmic acceleration nor detected the dark-energy component of the universe. What we have observed, are cosmological parameters sensitive to the expansion history of the universe. The results show that the Einstein–de Sitter model does not fit the current data (see [352] for an up-to-date discussion).

5.5.1. Conditions for acceleration

Observations also seem to indicate that the expansion of the universe started accelerating relatively recently; roughly at the time galaxy formation was moving from its linear phase into the nonlinear regime. This coincidence has led a number of cosmologists to suggest that cosmic acceleration could be the direct consequence of structure formation and of the fact that, at least on certain scales, the universe does not obey the FLRW symmetries. These ideas have led to what are now known as ‘kinematic backreaction’ scenarios.

Following (5.4.4) and in the absence of a cosmological constant, the backreaction effects will cause a given domain of the averaged universe to expand at an accelerated pace (i.e. with $\ddot{a}_D > 0$) provided that

$$\mathcal{Q}_D > \frac{1}{2} \langle \rho \rangle. \quad (5.5.1)$$

This means that \mathcal{Q}_D needs to be positive, which implies that shear fluctuations must be superseded by those in the volume expansion (see (5.4.6)). In the opposite case, \mathcal{Q}_D will be negative and the domain will decelerate further instead of accelerating. Condition (5.5.1) also implies that, if backreaction is to work, \mathcal{Q}_D should decay slower than the average density.

To examine the above condition further we introduce the dimensionless, domain-dependent parameters $\Omega_\rho = 3 \langle \rho \rangle / \langle \Theta \rangle^2$, $\Omega_{\mathcal{R}} = -3 \langle \mathcal{R} \rangle / 2 \langle \Theta \rangle^2$ and $\Omega_{\mathcal{Q}} = -3 \mathcal{Q}_D / 2 \langle \Theta \rangle^2$ and then recast expression (5.4.5) into

$$\Omega_\rho + \Omega_{\mathcal{R}} + \Omega_{\mathcal{Q}} = 1. \quad (5.5.2)$$

At the same time, condition (5.5.1) takes the form

$$\Omega_{\mathcal{Q}} < -\frac{1}{4} \Omega_\rho, \quad (5.5.3)$$

⁴⁰ The absence of rotation is necessary to guarantee the existence of flow-orthogonal hypersurfaces.

⁴¹ The fitting model does not always need to be the FRW spacetime. For example, in [351] the Buchert averaged formulae were written in Bianchi type- I form to study analogous backreaction effects on anisotropic cosmological models.

ensuring that \mathcal{Q}_D is positive as required. Substituting this result into Eq. (5.5.2), the latter leads to the constraint

$$\frac{3}{4} \Omega_\rho + \Omega_{\mathcal{R}} > 1, \quad (5.5.4)$$

between the matter and the 3-curvature contribution to the effective Ω -parameter of the averaged universe. Note that Ω_ρ is always positive, by construction, though $\Omega_{\mathcal{Q}}$ can take negative values as well.

It is conceivable that one can arrive to the same qualitative result, namely that the aforementioned backreaction effects can lead to accelerated expansion, through a perturbative approach as well. In [27], contributions from super-Hubble perturbations to the scale-factor evolution of a spatially flat FLRW background were claimed capable of accelerating the expansion rate of the inhomogeneous model, provided the initial perturbation had the right sign (see also [335] for the relevant counter-arguments).

5.5.2. Attractive aspects and caveats

The observed universe is full of nonlinear structures that, at least on certain scales, can seriously distort its perceived homogeneity and isotropy. As yet, the implications of these nonlinearities have not been fully appreciated, since the mathematical complexity of the problem has discouraged a systematic and in depth study. In this respect, spatial averaging may provide an alternative approach. An advantage of the Buchert averaging scheme is that it encodes the nonlinear effects mentioned above within scalars that are relatively simple to analyse (see [342] for recent review). It is not surprising therefore that the method has been applied to study the backreaction idea by several authors. Different averaging schemes, primarily that of Zalaletdinov [346], have also been employed with similar qualitative results [29,353,354]. The common factor is that backreaction can substantially change the behaviour of the exact model and therefore potentially explain the accelerated expansion of our universe without appealing to exotic matter fields, or introducing corrections to standard gravity. Moreover, since the effects are triggered by the onset of structure formation, they can offer a natural answer to the coincidence problem. All these make the backreaction idea attractive in principle.

However, current observations support a low density universe with Ω_ρ considerably less than unity. If accelerated expansion is to be triggered solely by backreaction effects, condition (5.5.4) is only satisfied in models with a considerable amount of negative (recall that $\langle \mathcal{R} \rangle = -2\langle \Theta \rangle^2 \Omega_{\mathcal{R}}/3$) spatial curvature [350,355]. In other words, condition (5.5.4) seems to contradict the widespread belief that the 3-curvature of our universe is very small.

Even if we assume that the aforementioned backreaction effects are strong enough to supersede the ‘background’ kinematics, nothing guarantees that they will lead to accelerated expansion. The sign of the backreaction term, which will decide which way the effects will go to, has yet to be decided. In fact, because of this sign-ambiguity, the same kinematic backreaction term, which is now proposed as a conventional solution to the dark-energy problem, was earlier suggested as an effective dark-matter source [356].

We may not have clear information on the sign of \mathcal{Q}_D , but we can monitor its dynamical evolution by means of a consistency/integrability condition. In particular, taking the time derivative of (5.4.5) and then using (5.4.3) and (5.4.4) one arrives at

$$\dot{\mathcal{Q}}_D + 2\langle \Theta \rangle \mathcal{Q}_D = -\langle \mathcal{R} \rangle' - \frac{2}{3} \langle \Theta \rangle \langle \mathcal{R} \rangle. \quad (5.5.5)$$

The above means that if $\langle \mathcal{R} \rangle \propto a_D^{-2}$, namely if the averaged spatial curvature behaves like its FRW counterpart, the backreaction term will scale as $\mathcal{Q}_D \propto a_D^{-6}$ mimicking a ‘stiff’ dilatonic fluid. Clearly, the same evolution law holds when $\langle \mathcal{R} \rangle = 0$ as well.⁴² In either case, the density term in (5.4.4) will quickly dominate the backreaction effects, even when condition (5.5.4) is initially satisfied. Recall that $\langle \rho \rangle \propto a_D^{-3}$, just like in the standard Einstein–de Sitter model. Therefore, a change in the scale-factor dependence of the average 3-Ricci scalar seems necessary if the backreaction idea is to work [355]. Overall, if condition (5.5.1) holds, a typical Hubble volume would not correspond to a perturbative state near the LFRW model [355].

6. Summary and outlook

Cosmology is now firmly a data-driven science. The principal drivers over the past fifteen years have been large surveys of galaxy redshifts and hence their three-dimensional clustering, precision measurements of the CMB temperature anisotropies and, recently, polarization, and measurements of the magnitude-redshift relation of distant supernovae. Together, these have revolutionised our understanding of the constituents, geometry and initial conditions for the observable universe.

General relativity, or its Newtonian approximation on sub-Hubble scales, appears to provide a satisfactory description of the dynamics of the universe and its fluctuations. This is a considerable triumph for the theory since it extends the range of

⁴² As expected, when \mathcal{Q}_D vanishes, $\langle \mathcal{R} \rangle \propto a_D^{-2}$ and the averaged model behaves like its standard LFRW counterpart (see Eqs. (5.4.3)–(5.4.5)). Generally speaking, however, the reduced system still describes inhomogeneous non-equilibrium states.

scales over which it has passed observational tests by many orders of magnitude. Bold theoretical predictions of the simplest models, in which initially small, adiabatic and Gaussian fluctuations evolve passively under gravity and hydrodynamics on an FRW background, have been impressively verified. The observed acoustic peaks in the CMB and galaxy clustering power spectrum are a particularly noteworthy example. Our aim here has been to provide a comprehensive review of the dynamics and perturbations of cosmological models that are based in general relativity, and to describe how a simple class of these models provides an excellent and remarkably efficient description of current data. By uniformly working with 1+3-covariant methods throughout, we hope to have presented a unified treatment of a range of topics in contemporary cosmology.

However, general relativity can only explain the observed structure and evolution of the universe by supplementing the known particles of the standard model with additional components that have, as yet, not yielded to direct detection. Cold dark matter is invoked to reconcile the rotation curves of galaxies with their distribution of luminous matter and the observed clustering of galaxies with the small amplitude of CMB fluctuations. One or more potential-dominated scalar fields are invoked in the early universe to drive inflation and hence remove some of the fine-tuning issues that beset the standard hot big-bang scenario. A further violation of the strong energy condition is required at late times to explain the observed accelerated expansion. This requires that a further component dubbed dark energy – either a woefully small contribution from the vacuum, or an additional, unclustered dynamical component – be introduced. Alternatives to CDM, a fundamental inflaton field and dark energy are all being pursued actively. For example, in attempts to realise inflation in string/M theory, the role of the inflaton can be played by scalar fields describing the geometry of the compactified dimensions (see Ref. [8] for a recent review). Infra-red modifications to general relativity can lead to late-time acceleration without dark energy, but do not address the problem of why the vacuum does not gravitate [357]. Introducing additional scalar, vector and tensor degrees of freedom (not all of which are dynamical) yields a relativistic theory of gravity [358] which reduces to the phenomenology of Milgrom’s modified Newtonian dynamics [359] on galactic scales and may therefore offer an alternative to CDM. Future astronomical observations should be able to discriminate between some of these alternatives. For example, a detection of CMB B -mode polarization may signal trouble for many string-theory scenarios of inflation [360]; gravitational lensing observations may provide further support for a dominant, non-interacting, cold component (i.e. CDM) in galaxy clusters and large-scale structures [361,362]; and accurately mapping out the expansion rate of the universe to high redshift with future supernovae surveys and baryon acoustic oscillations, and the growth rate of structure with tomographic lensing surveys, may distinguish between modifications of gravity or a physical dark energy component [363].

Acknowledgements

CT wishes to thank John Barrow and George Ellis for many helpful discussions. AC acknowledges a Royal Society University Research Fellowship and gratitude to his long-term collaborators in this area, particularly Antony Lewis. The work of RM is supported by STFC and he thanks George Ellis and Malcolm MacCallum for useful discussions.

Appendix A

A.1. 1 + 3 Covariant decomposition

The skew part of a projected rank-2 tensor is spatially dual to the projected vector $S_a = \varepsilon_{abc}S^{bc}/2$, and any projected second-rank tensor has the irreducible covariant decomposition

$$S_{ab} = \frac{1}{3}Sh_{ab} + \varepsilon_{abc}S^c + S_{(ab)}, \quad (\text{A.1})$$

where $S = S_{cd}h^{cd}$ is the spatial trace and $S_{(ab)} = S_{(ab)} - Sh_{ab}/3$ is the projected symmetric and trace-free (PSTF) part of S_{ab} . In the 1 + 3 covariant formalism, all quantities are either scalars, projected vectors or PSTF tensors.

The projected derivative operator, $D_a = h_a{}^b\nabla_b$, further splits irreducibly into a 1 + 3 covariant spatial divergence

$$D^aV_a, \quad D^bS_{ab}, \quad (\text{A.2})$$

a spatial curl

$$\text{curl } V_a = \varepsilon_{abc}D^bV^c, \quad \text{curl } S_{ab} = \varepsilon_{cd(a}D^cS_{b)}{}^d, \quad (\text{A.3})$$

and a 1 + 3 covariant spatial distortion

$$D_{(a}V_{b)} = D_{(a}V_{b)} - \frac{1}{3}D_cV^c h_{ab}, \quad D_{(a}S_{bc)} = D_{(a}S_{bc)} - \frac{2}{5}h_{(ab}D^dS_{c)d}. \quad (\text{A.4})$$

Note that, as a result of the commutation laws between projected covariant derivatives (see next section), the div curl operation is *not* in general zero.

The covariant irreducible decompositions of the derivatives of scalars, vectors and rank-2 tensors are given in exact (nonlinear) form by

$$\nabla_a \psi = -\dot{\psi} u_a + D_a \psi, \quad (\text{A.5})$$

$$\nabla_b V_a = -u_b (\dot{V}_{(a)} + A_c V^c u_a) + u_a \left[\frac{1}{3} \Theta V_b + \sigma_{bc} V^c + \varepsilon_{bde} \omega^d V^e \right] + \frac{1}{3} D_c V^c h_{ab} - \frac{1}{2} \varepsilon_{abc} \text{curl } V^c + D_{(a} V_{b)}, \quad (\text{A.6})$$

and

$$\nabla_c S_{ab} = -u_c (\dot{S}_{(ab)} + 2u_{(a} S_{b)d} A^d) + 2u_{(a} \left[\frac{1}{3} \Theta S_{b)c} + S_{b)}^d (\sigma_{cd} - \varepsilon_{cde} \omega^e) \right] + \frac{3}{5} D^d S_{(a|d|} h_{b)c} - \frac{2}{3} \varepsilon_{dc(a} S_{b)}^d + D_{(a} S_{bc)}, \quad (\text{A.7})$$

respectively. The algebraic correction terms in Eqs. (A.6) and (A.7) arise from the relative motion of comoving observers, as encoded in the kinematic quantities.

A.2. Transformations under a 4-velocity boost

Consider an observer moving with 4-velocity \tilde{u}_a relative to the u_a -frame. The two motions are related by the transformation

$$\tilde{u}_a = \gamma(u_a + v_a), \quad (\text{A.8})$$

where $\gamma = (1 - v^2)^{-1/2}$ is the Lorentz-boost factor and v_a is the ‘peculiar’ velocity (with $v_a u^a = 0$). Note that for non-relativistic peculiar motions $v \ll 1$ and $\gamma \simeq 1$. Also,

$$\tilde{h}_{ab} = h_{ab} + \gamma^2 (v^2 u_a u_b + 2u_{(a} v_{b)} + v_a v_b), \quad (\text{A.9})$$

$$\tilde{\varepsilon}_{abc} = \gamma \varepsilon_{abc} + \gamma (2u_{[a} \varepsilon_{b]cd} + u_c \varepsilon_{abd}) v^d, \quad (\text{A.10})$$

are the relations between the fundamental algebraic tensors of the two frames.

The irreducible kinematic quantities, as measured in the \tilde{u}_a -frame, are defined by means of the decomposition (see Eq. (1.3.1) in Section 1.3.1)

$$\nabla_b \tilde{u}_a = \frac{1}{3} \tilde{\Theta} \tilde{h}_{ab} + \tilde{\sigma}_{ab} + \tilde{\varepsilon}_{abc} \tilde{\omega}^c - \tilde{A}_a \tilde{u}_b. \quad (\text{A.11})$$

Using the above, the relation $\nabla_a \gamma = \gamma^3 v^b \nabla_a v_b$ and expression (A.6) from Appendix A.1, we arrive at the following kinematic transformation laws:

$$\tilde{\Theta} = \gamma \Theta + \gamma (D_a v^a + A^a v_a) + \gamma^3 W, \quad (\text{A.12})$$

$$\begin{aligned} \tilde{A}_a = & \gamma^2 A_a + \gamma^2 \left[\dot{v}_{(a)} + \frac{1}{3} \Theta v_a + \sigma_{ab} v^b - \varepsilon_{abc} \omega^b v^c + \left(\frac{1}{3} \Theta v^2 + A^b v_b + \sigma_{bc} v^b v^c \right) u_a \right. \\ & \left. + \frac{1}{3} (D_b v_a) v_a + \frac{1}{2} \varepsilon_{abc} v^b \text{curl } v^c + v^b D_{(b} v_{a)} \right] + \gamma^4 W (u_a + v_a), \end{aligned} \quad (\text{A.13})$$

$$\begin{aligned} \tilde{\omega}_a = & \gamma^2 \left[\left(1 - \frac{1}{2} v^2 \right) \omega_a - \frac{1}{2} \text{curl } v_a + \frac{1}{2} v_b (2\omega^b - \text{curl } v^b) u_a + \frac{1}{2} v_b \omega^b v_a \right. \\ & \left. + \frac{1}{2} \varepsilon_{abc} A^b v^c + \frac{1}{2} \varepsilon_{abc} \dot{v}^b v^c u_a + \frac{1}{2} \varepsilon_{abc} \sigma^b{}_d v^c v^d \right], \end{aligned} \quad (\text{A.14})$$

$$\begin{aligned} \tilde{\sigma}_{ab} = & \gamma \sigma_{ab} + \gamma (1 + \gamma^2) u_{(a} \sigma_{b)c} v^c + \gamma^2 A_{(a} [v_{b)} + v^2 u_{b)}] + \gamma D_{(a} v_{b)} - \frac{1}{3} h_{ab} [A_c v^c + \gamma^2 (W - \dot{v}_c v^c)] \\ & + \gamma^3 u_a u_b \left[\sigma_{cd} v^c v^d + \frac{2}{3} v^2 A_c v^c - v^c v^d D_{(c} v_{d)} + \left(\gamma^4 - \frac{1}{3} v^2 \gamma^2 - 1 \right) W \right] \\ & + \gamma^3 u_{(a} v_{b)} \left[A_c v^c + \sigma_{cd} v^c v^d - \dot{v}_c v^c + 2\gamma^2 \left(\gamma^2 - \frac{1}{3} \right) W \right] + \frac{1}{3} \gamma^3 v_a v_b [D_c v^c - A_c v^c + \gamma^2 (3\gamma^2 - 1) W] \\ & + \gamma^3 v_{(a} \dot{v}_{b)} + v^2 \gamma^3 u_{(a} \dot{v}_{b)} + \gamma^3 v_{(a} \sigma_{b)c} v^c - \gamma^3 \omega^b v^c \varepsilon_{bc(a} (v_{b)} + v^2 u_{b)} + 2\gamma^3 v^c D_{(c} v_{a)} (v_{b)} + u_{b)}), \end{aligned} \quad (\text{A.15})$$

with

$$W \equiv \dot{v}_c v^c + \frac{1}{3} v^2 D_c v^c + v^c v^d D_{(c} v_{d)}. \quad (\text{A.16})$$

Similarly, one can decompose the energy–momentum tensor of the matter with respect to the \tilde{u}_a -frame (see expression (1.2.11) in Section 1.2.3). Then,

$$T_{ab} = \tilde{\rho} \tilde{u}_a \tilde{u}_b + \tilde{p} \tilde{h}_{ab} + 2\tilde{q}_{(a} \tilde{u}_{b)} + \tilde{\pi}_{ab}, \quad (\text{A.17})$$

and the transformed dynamic quantities are given by

$$\tilde{\rho} = \rho + \gamma^2 \left[v^2(\rho + p) - 2q_a v^a + \pi_{ab} v^a v^b \right], \quad (\text{A.18})$$

$$\tilde{p} = p + \frac{1}{3} \gamma^2 \left[v^2(\rho + p) - 2q_a v^a + \pi_{ab} v^a v^b \right], \quad (\text{A.19})$$

$$\tilde{q}_a = \gamma q_a - \gamma \pi_{ab} v^b - \gamma^3 \left[(\rho + p) - 2q_b v^b + \pi_{bc} v^b v^c \right] v_a - \gamma^3 \left[v^2(\rho + p) - (1 + v^2)q_b v^b + \pi_{bc} v^b v^c \right] u_a, \quad (\text{A.20})$$

$$\begin{aligned} \tilde{\pi}_{ab} = & \pi_{ab} + 2\gamma^2 v^c \pi_{c(a} \{u_{b)} + v_{b)}\} - 2v^2 \gamma^2 q_{(a} u_{b)} - 2\gamma^2 q_{(a} v_{b)} - \frac{1}{3} \gamma^2 \left[v^2(\rho + p) + \pi_{cd} v^c v^d \right] h_{ab} \\ & + \frac{1}{3} \gamma^4 \left[2v^4(\rho + p) - 4v^2 q_c v^c + (3 - v^2) \pi_{cd} v^c v^d \right] u_a u_b + \frac{2}{3} \gamma^4 \left[2v^2(\rho + p) - (1 + 3v^2)q_c v^c + 2\pi_{cd} v^c v^d \right] u_{(a} v_{b)} \\ & + \frac{1}{3} \gamma^4 \left[(3 - v^2)(\rho + p) - 4q_c v^c + 2\pi_{cd} v^c v^d \right] v_a v_b. \end{aligned} \quad (\text{A.21})$$

Finally, relative to the \tilde{u}_a -frame, the Gravito-electric/magnetic field decomposes according to (see Eq. (1.2.9) in Section 1.3.6),

$$C_{ab}{}^{cd} = 4 \left(\tilde{u}_{[a} \tilde{u}^{[c} + \tilde{h}_{[a}{}^{[c} \right) \tilde{E}_{b]}{}^{d]} + 2\tilde{\epsilon}_{abe} \tilde{u}^{[c} \tilde{H}^{d]e} + 2\tilde{u}_{[a} \tilde{H}_{b]e} \tilde{\epsilon}^{cde}, \quad (\text{A.22})$$

where \tilde{E}_{ab} and \tilde{H}_{ab} are respectively the Lorentz-boosted electric and magnetic components of the free gravitational field. Then,

$$\tilde{E}_{ab} = \gamma^2 \left\{ (1 + v^2)E_{ab} + v^c \left[2\epsilon_{cd(a} H_{b)}{}^d + 2E_{c(a} u_{b)} + (u_a u_b + h_{ab})E_{cd} v^d - 2E_{c(a} v_{b)} + 2u_{(a} \epsilon_{b)cd} H^{de} v_e \right] \right\}, \quad (\text{A.23})$$

$$\tilde{H}_{ab} = \gamma^2 \left\{ (1 + v^2)H_{ab} + v^c \left[-2\epsilon_{cd(a} E_{b)}{}^d + 2H_{c(a} u_{b)} + (u_a u_b + h_{ab})H_{cd} v^d - 2H_{c(a} v_{b)} - 2u_{(a} \epsilon_{b)cd} E^{de} v_e \right] \right\}. \quad (\text{A.24})$$

The transformation laws of the electric and magnetic Weyl components, may be compared to those of their Maxwell counterparts, namely to

$$\tilde{E}_a = \gamma \left(E_a + \epsilon_{abc} v^b B^c + v^b E_b u_a \right), \quad (\text{A.25})$$

$$\tilde{B}_a = \gamma \left(B_a - \epsilon_{abc} v^b E^c + v^b B_b u_a \right), \quad (\text{A.26})$$

where (see decomposition (1.3.9) in Section 1.3.2)

$$F_{ab} = 2u_{[a} E_{b]} + \epsilon_{abc} B^c = 2\tilde{u}_{[a} \tilde{E}_{b]} + \tilde{\epsilon}_{abc} \tilde{B}^c. \quad (\text{A.27})$$

Note that all the transformations are given explicitly in terms of irreducible quantities (i.e. irreducible in the original u_a -frame).

A.3. Covariant commutation laws

According to definition (1.2.2)(a), the orthogonally projected covariant derivative operator satisfies the condition $D_a h_{bc} = 0$. This means that we can use h_{ab} to raise and lower indices in equations acted upon by this operator. Following Frobenius' theorem, however, rotating spaces do not possess integrable 3-D submanifolds (e.g. see [49,65]). Therefore, the D_a -operator cannot be used as a standard 3-D derivative in such spaces and it does not always satisfy the usual commutation laws (see below and also [138]).

When acting on a scalar quantity the orthogonally projected covariant derivative operators commute according to

$$D_{[a} D_{b]} f = -\omega_{ab} \dot{f}. \quad (\text{A.28})$$

The above is a purely relativistic result and underlines the different behaviour of rotating spacetimes within Einstein's theory. Similarly, the commutation law for the orthogonally projected derivatives of spacelike vectors reads

$$D_{[a} D_{b]} v_c = -\omega_{ab} \dot{v}_{(c)} + \frac{1}{2} \mathcal{R}_{dcba} v^d, \quad (\text{A.29})$$

where $v_a u^a = 0$ and \mathcal{R}_{abcd} represents the Riemann tensor of the observer's local rest-space. Finally, when dealing with orthogonally projected tensors, we have

$$D_{[a} D_{b]} S_{cd} = -\omega_{ab} h_c{}^e h_d{}^f \dot{S}_{ef} + \frac{1}{2} (\mathcal{R}_{ecba} S^e{}_d + \mathcal{R}_{edba} S_c{}^e), \quad (\text{A.30})$$

with $S_{ab} u^a = 0 = S_{ab} u^b$. Note that in the absence of rotation, \mathcal{R}_{abcd} is the Riemann tensor of the (integrable) 3-D hypersurfaces orthogonal to the u_a -congruence. For details on the definition, the symmetries and the key equations involving \mathcal{R}_{abcd} , the

reader is referred to Section 1.3.5. We also note that the above equations are fully nonlinear and hold at all perturbative levels.

In general relativity, time derivatives do not generally commute with their spacelike counterparts. For scalars, in particular, we have

$$D_a \dot{f} - h_a^b (D_b \dot{f}) = -\dot{f} A_a + \frac{1}{3} \Theta D_a f + D_b f (\sigma^b_a + \omega^b_a), \quad (\text{A.31})$$

at all perturbative levels. Assuming an FLRW background, we find that the orthogonally projected gradient and the time derivative of the first-order vector v_a commute as

$$a D_a \dot{v}_b = (a D_a v_b) \dot{}, \quad (\text{A.32})$$

to linear order. Similarly, when dealing with first-order spacelike tensors, we have the following linear commutation law

$$a D_a \dot{S}_{bc} = (a D_a S_{bc}) \dot{}. \quad (\text{A.33})$$

A.4. Scalar, vector and tensor modes

In the coordinate-based approach, perturbations are decomposed from the start into scalar, vector and tensor modes, using appropriate harmonics. The covariant approach does not depend on a priori splitting into harmonic modes and it is independent of any Fourier-type decomposition. Instead, all the perturbative quantities are described as spatial vectors $V_a = V_{(a)}$ or as spatial, symmetric and trace-free (PSTF) rank-2 tensors $S_{ab} = S_{(ab)}$ (higher-rank PSTF tensors are needed in kinetic theory—see Section 4).

The scalar modes are characterised by the fact that all vectors and tensors are generated by scalar potentials. For instance,

$$V_a = D_a V \quad \text{and} \quad S_{ab} = D_{(a} D_{b)} S, \quad (\text{A.34})$$

for some V, S . This implies that $\text{curl } V_a = 0 = \text{curl } S_{ab}$.

For vector modes, all vectors are transverse (solenoidal) and proportional to ω_a . Also, all tensors are generated by transverse vector potentials. Thus,

$$D^a V_a = 0 \quad \text{and} \quad S_{ab} = D_{(a} S_{b)}, \quad (\text{A.35})$$

where $D^a S_a = 0$. Vector modes are nonzero if and only if the vorticity is nonzero.

Tensor modes are characterised by the vanishing of all vectors and by the transverse traceless nature of all tensors. In other words,

$$V_a = 0 \quad \text{and} \quad D^b S_{ab} = 0. \quad (\text{A.36})$$

This way no perturbative scalars or vectors can be formed.

We can expand these modes in harmonic basis functions (Fourier modes in the case $K = 0$). For example, for scalar modes, the harmonics are time-independent eigenfunctions that satisfy the scalar Laplace–Beltrami equation. In other words, $\hat{\mathcal{Q}}^{(k)} = 0$ and

$$D^2 \mathcal{Q}^{(k)} = -\left(\frac{k}{a}\right)^2 \mathcal{Q}^{(k)}, \quad (\text{A.37})$$

where k is the eigenvalue of the associated harmonic mode and $D^2 = D^a D_a$. The latter takes continuous values when $K = 0, -1$ and discrete ones for $K = +1$. In particular, $k = \nu \geq 0$ when the 3-space has Euclidean geometry and $k^2 = \nu^2 + 1 \geq 0$ for hyperbolic spatial sections, with ν representing the comoving wavenumber of the mode in all cases. Supercurvature modes have $\lambda = a/k > a$ and in open FLRW models correspond to $0 \leq k^2 < 1$. Those with $k^2 > 1$, on the other hand, span scales smaller than the curvature radius and are therefore termed subcurvature. Clearly, the $k^2 = 1$ threshold indicates the curvature scale, with $\lambda = \lambda_K = a$ (see also Section 1.4.2). Note that, although they are often ignored (e.g. see [180]), supercurvature modes are necessary if we want perturbations with correlations lengths bigger than the curvature radius [71]. Finally, when the 3-curvature is positive, $k^2 = \nu(\nu + 2)$ and $\nu = 1, 2, \dots$

Appendix B. Notation

• Spacetime Geometry

Line element: $ds^2 = g_{ab} dx^a dx^b = -d\tau^2$, with $c = 1$.

4-velocity: $u^a = dx^a/d\tau$, 3-D projection tensor: $h_{ab} = g_{ab} + u_a u_b$.

4-D permutation tensor: η_{abcd} , 3-D permutation tensor: $\varepsilon_{abc} = \eta_{abcd} u^d$.

Covariant derivative: $\nabla_b T_a = \partial T_a / \partial x^b - \Gamma_{ab}^c T_c$.

Time derivative: $\dot{T}_a = u^b \nabla_b T_a$, 3-D covariant derivative: $D_b T_a = h_b^d h_a^c \nabla_d T_c$.

Riemann tensor: R_{abcd} , Ricci tensor: $R_{ab} = R^c_{acb}$, Ricci scalar: $R = R^a_a$.

3-Riemann tensor: \mathcal{R}_{abcd} , 3-Ricci tensor: $\mathcal{R}_{ab} = \mathcal{R}^c_{acb}$, 3-Ricci scalar: $\mathcal{R} = \mathcal{R}^a_a$.

3-curvature index: $K = 0, \pm 1$, with $\mathcal{R} = 6K/a^2$ (in FLRW models).

Weyl Tensor: C_{abcd} , electric Weyl: $E_{ab} = C_{acbd} u^c u^d$, magnetic Weyl: $H_{ab} = \varepsilon_a^{cd} C_{cdbe} u^e / 2$.

• Kinematics

Expansion scalar: $\Theta = \nabla^a u_a = D^a u_a$, scale factor: a , with $\dot{a}/a = \Theta/3$.

Conformal time: η , with $\dot{\eta} = 1/a$.

Vorticity tensor: $\omega_{ab} = D_{[b} u_{a]}$, vorticity vector: $\omega_a = \varepsilon_{abc} \omega^{bc}/2$.

Shear tensor: $\sigma_{ab} = D_{(b} u_{a)} = D_{(b} u_{b)} - (D^c u_c) h_{ab}/3$, 4-acceleration: $A_a = u^b \nabla_b u_a$.

Hubble parameter: $H = \dot{a}/a$, deceleration parameter: $q = -\ddot{a}/\dot{a}^2$.

• Matter Fields

Field equations: $R_{ab} - (R/2)g_{ab} = T_{ab}$, with $\kappa = 8\pi G = 1$.

Matter energy-momentum tensor: $T_{ab} = \rho u_a u_b + p h_{ab} + 2u_{(a} q_{b)} + \pi_{ab}$.

Matter density: $\rho = T_{ab} u^a u^b$, isotropic pressure: $p = T_{ab} h^{ab}/3$.

Barotropic index: $w = p/\rho$, adiabatic sound speed: $c_s^2 = \dot{p}/\dot{\rho}$.

Energy flux: $q_a = h_a^b T_{bc} u^c$, anisotropic pressure: $\pi_{ab} = T_{(ab)} = T_{(ab)} - (T/3)h_{ab}$.

Particle flux vector: $N_a = n u_a + \mathcal{N}_a$.

Particle number density: $n = -N_a u^a$, particle drift: $\mathcal{N}_a = h_a^b N_b$.

Entropy flux: $S_a = S u_a$ (in equilibrium).

Entropy density: $S = -S_a u^a$, specific entropy: $s = S/n$.

• Electromagnetism

Electromagnetic tensor: F_{ab} , magnetic field: $B_a = \varepsilon_{abc} F^{bc}/2$, electric field: $E_a = F_{ab} u^b$.

Energy density: $(B^2 + E^2)/2$, isotropic pressure: $(B^2 + E^2)/6$.

Poynting vector: $\mathcal{P}_a = \varepsilon_{abc} E^b B^c$, anisotropic pressure: $\Pi_{ab} = -E_{(a} E_{b)} - B_{(a} B_{b)}$.

Alfvén speed: $c_a^2 = B^2/(\rho + p + B^2)$.

Electric 4-current: J_a , electric 3-current: $\mathcal{J}_a = J_{(a)} = h_a^b J_b$.

Charge density: $\mu = -J_a u^a$, electrical conductivity: ζ .

• Minimally coupled scalar fields

Scalar field: φ , with $\dot{\varphi} = -\nabla_a \nabla^a \varphi > 0$ and $D_a \varphi = 0$.

Potential: $V(\varphi)$, with $\nabla_a \nabla^a \varphi - V'(\varphi) = 0$.

Energy density: $\rho^{(\varphi)} = \dot{\varphi}^2/2 + V(\varphi)$, pressure: $p^{(\varphi)} = \dot{\varphi}^2/2 - V(\varphi)$.

• Perturbations

Matter density gradients: $\Delta_a = (a/\rho) D_a \rho$, with $\Delta_{ab} = a D_b \Delta_a$ and $\Delta = \Delta^a_a$.

Matter vortices: $\mathcal{W}_{ab} = \Delta_{[ab]}$, with $\mathcal{W}_a = \varepsilon_{abc} \mathcal{W}^{bc}/2$.

Volume expansion gradients: $\mathcal{Z}_a = a D_a \Theta$, with $\mathcal{Z}_{ab} = a D_b \mathcal{Z}_a$ and $\mathcal{Z} = \mathcal{Z}^a_a$.

Magnetic density gradients: $\mathcal{B}_a = (a/B^2) D_b B^2$ and $\mathcal{B}_{ab} = a D_b \mathcal{B}_a$, with $\mathcal{B} = \mathcal{B}^a_a$.

Effective entropy perturbations: $\mathcal{E}_a, \mathcal{E}_a^{(ij)}$, with $\mathcal{E}_a^{(ij)} = -\mathcal{E}_a^{(ji)}$.

Peculiar velocity: v_a , with $v_a u^a = 0$.

References

- [1] D.N. Spergel, et al., *Astrophys. J. Suppl.* 170 (2007) 377. arXiv:astro-ph/0603449.
- [2] W.J. Percival, et al., *Astrophys. J.* 657 (2007) 645. arXiv:astro-ph/0608636.
- [3] M. Bojowald, *Living Rev. Rel.* 8 (2005) 11. arXiv:gr-qc/0601085.
- [4] J.K. Erickson, et al., *Phys. Rev. D* 75 (2007) 123507. arXiv:hep-th/0607164.
- [5] S.-H. Henry Tye, *Lect. Notes Phys.* 737 (2008) 949. arXiv:hep-th/0610221.
- [6] A. Ashtekar, T. Pawłowski, P. Singh, *Phys. Rev. D* 74 (2006) 084003. arXiv:gr-qc/0607039.
- [7] R. Brandenberger, *Mod. Phys. Lett. A* 22 (2007) 1885. arXiv:hep-th/0702001.
- [8] R. Kallosh, *Lect. Notes Phys.* 738 (2008) 199. arXiv:hep-th/0702059.
- [9] E.J. Copeland, M. Sami, S. Tsujikawa, *Internat. J. Modern Phys. D* (2006) arXiv:hep-th/0603057.
- [10] L. Perivolaropoulos, in: N. Solomos (Ed.), *Recent Advances in Astronomy and Astrophysics*, in: AIPCS, vol. 848, 2006, pp. 698–712.
- [11] S. Nojiri, S.D. Odintsov, *Int. J. Geom. Meth. Modern Phys.* 4 (2007) 115. arXiv:hep-th/0601213.
- [12] T. Padmanabhan, *Albert Einstein Century International Conference*, in: AIPCS, vol. 861, 2006, pp. 179–196.
- [13] D. Polarski, *Albert Einstein Century International Conference*, in: AIPCS, vol. 861, 2006, pp. 1013–1018.
- [14] N. Straumann, *Modern Phys. Lett. A* 21 (2006) 1083. arXiv:hep-ph/0604231.
- [15] S. Bludman, *VI Reunion Anual Sociedad Chilena de Astronomia (SOCHIAS)*, p. 9, 2007.
- [16] J.P. Uzan, *Gen. Relativity Gravitation* 4 (2007) arXiv:astro-ph/0605313.
- [17] R.A. Knop, et al., *Astrophys. J.* 598 (2003) 102. arXiv:astro-ph/0309368.
- [18] W.M. Wood-Vasey, et al., *Astrophys. J.* 666 (2007) 694. arXiv:astro-ph/0701041.
- [19] M.R. Nolta, et al. *ArXiv e-prints* 803, 2008. arXiv:0803.0593.
- [20] W.C. Jones, et al., *Astrophys. J.* 647 (2006) 823. arXiv:astro-ph/0507494.
- [21] A.C.S. Readhead, et al., *Astrophys. J.* 609 (2004) 498. arXiv:astro-ph/0402359.
- [22] C.L. Reichardt, et al. *ArXiv e-prints* 801, 2008. arXiv:0801.1491.
- [23] J. Polchinski, in: D. Gross, M. Henneaux, A. Servin (Eds.), *Quantum Structure of Space and Time*, World Scientific, Hackensack, 2007, arXiv:hep-th/0603249.
- [24] R. Bousso, *ArXiv high energy physics-theory e-prints*, 2006. arXiv:hep-th/0610211.
- [25] S. Räsänen, *JCAP* 2 (2004) 3. arXiv:astro-ph/0311257.
- [26] E. Barausse, S. Matarrese, A. Riotto, *Phys. Rev. D* 71 (2005) 063537. arXiv:astro-ph/0501152.
- [27] E.W. Kolb, et al. *ArXiv high energy physics-theory e-prints*, 2005. arXiv:hep-th/0503117.
- [28] G. Geshnizjani, D.J. Chung, N. Afshordi, *Phys. Rev. D* 72 (2005) 023517. arXiv:astro-ph/0503553.
- [29] A.A. Coley, N. Pelavas, R.M. Zalaletdinov, *Phys. Rev. Lett.* 95 (2005) 151102. arXiv:gr-qc/0504115.
- [30] Y. Nambu, M. Tanimoto, *ArXiv General Relativity and Quantum Cosmology e-prints*, 2005. arXiv:gr-qc/0507057.
- [31] E.W. Kolb, S. Matarrese, A. Riotto, *New J. Phys.* 8 (2006) 322. arXiv:astro-ph/0506534.

- [32] P. Martineau, R. Brandenberger, ArXiv astrophysics e-prints, 2005. arXiv:astro-ph/0510523.
- [33] R. Mansouri, ArXiv astrophysics e-prints 2006. arXiv:astro-ph/0601699.
- [34] J.W. Moffat, Albert Einstein Century International Conference, in: AIP Conference Series, vol 861, 2006, pp. 987–994.
- [35] A. Paranjape, T.P. Singh, Classical Quantum Gravity 23 (2006) 6955. arXiv:astro-ph/0605195.
- [36] C. Clarkson, B.A. Bassett, T. Hui-Ching Lu, ArXiv e-prints, 2007. arXiv:0712.3457.
- [37] R. Sanders, in: L. Papantonopoulos (Ed.), Lecture Notes in Physics, vol. 720, Springer Verlag, Berlin, 2007, 375–+.
- [38] F. Bourliot, et al., Phys. Rev. D 75 (2007) 063508. arXiv:astro-ph/0611255.
- [39] O. Heckmann, E. Schücking, Zeits. f. Astroph. 38 (1955) 95.
- [40] A. Raychaudhuri, Zeits. f. Astroph. 43 (1957) 161.
- [41] G.F.R. Ellis, in: R.K. Sachs (Ed.), General Relativity and Cosmology, 1971, pp. 104–180.
- [42] G.F.R. Ellis, in: E. Schatzman (Ed.), Cargèse Lectures in Physics, 1973, pp. 1–60.
- [43] G.F.R. Ellis, in: S. Dhurandhar, T. Padmanabhan (Eds.), Gravitation and Cosmology, in: Astrophysics and Space Science Library, vol. 211, 1997, pp. 53–74.
- [44] G.F.R. Ellis, H. van Elst, in: M. Lachièze-Rey (Ed.), Theoretical and Observational Cosmology, 1999, pp. 1–116. arXiv:gr-qc/9812046.
- [45] S.W. Hawking, G.F.R. Ellis, The Large Scale Structure of Space-Time, Cambridge University Press, Cambridge, 1973.
- [46] R. Maartens, Phys. Rev. D 55 (1997) 463. arXiv:astro-ph/9609198.
- [47] N. Dadhich, 2005. arXiv:gr-qc/0511123.
- [48] S. Kar, S. SenGupta, Pramana 69 (2007) 49. arXiv:gr-qc/0611123.
- [49] R.M. Wald, General Relativity, University of Chicago Press, Chicago, 1984.
- [50] C.G. Tsagas, Classical Quantum Gravity 22 (2005) 393. arXiv:gr-qc/0407080.
- [51] R.R. Caldwell, Phys. Lett. B 545 (2002) 23. arXiv:astro-ph/9908168.
- [52] R.R. Caldwell, M. Kamionkowski, N.N. Weinberg, Phys. Rev. Lett. 91 (2003) 071301. arXiv:astro-ph/0302506.
- [53] A. Starobinsky, in: V. Burduzha, G. Khozin (Eds.), The Future of the Universe and the Future of our Civilization, 2000, p. 71.
- [54] S. Nojiri, S.D. Odintsov, Phys. Lett. B 562 (2003) 147. arXiv:hep-th/0303117.
- [55] M.P. Dabrowski, T. Stachowiak, M. Szydlowski, Phys. Rev. D 68 (2003) 103519. arXiv:hep-th/0307128.
- [56] P. Singh, M. Sami, N. Dadhich, Phys. Rev. D 68 (2003) 023522. arXiv:hep-th/0305110.
- [57] L.P. Chimento, R. Lazkoz, Phys. Rev. Lett. 91 (2003) 211301. arXiv:gr-qc/0307111.
- [58] H. Štefančić, Phys. Rev. D 71 (2005) 124036. arXiv:astro-ph/0504518.
- [59] J.D. Barrow, Classical Quantum Gravity 21 (2004) L79. arXiv:gr-qc/0403084.
- [60] J.D. Barrow, C.G. Tsagas, Classical Quantum Gravity 22 (2005) 1563. arXiv:gr-qc/0411045.
- [61] W. Israel, Ann. Phys. 100 (1976) 310.
- [62] R. Maartens, ArXiv astrophysics e-prints (1996). arXiv:astro-ph/9609119.
- [63] W. Israel, J.M. Stewart, Ann. Phys. 118 (1979) 341.
- [64] E. Bertschinger, in: R. Schaeffer, et al. (Eds.), Cosmology and Large Scale Structure, 1996, p. 273. arXiv:astro-ph/9503125.
- [65] E. Poisson, A Relativist's Toolkit, Cambridge University Press, Cambridge, 2004.
- [66] J.D. Barrow, R. Maartens, C.G. Tsagas, Phys. Rep. 449 (2007) 131. arXiv:astro-ph/0611537.
- [67] L. Bel, Acad. Sci. Paris 247 (1958) 1094.
- [68] R. Penrose, Ann. Phys. 138 (1969) 59.
- [69] R. Maartens, B.A. Bassett, Classical Quantum Gravity 15 (1998) 705. arXiv:gr-qc/9704059.
- [70] N. Dadhich, Gen. Relativity Gravitation 32 (2000) 1009. arXiv:gr-qc/9909067.
- [71] D.H. Lyth, A. Woszczyna, Phys. Rev. D 52 (1995) 3338. arXiv:astro-ph/9501044.
- [72] J.V. Narlikar, Introduction to Cosmology, Jones and Bartlett, Boston, 1983.
- [73] M.S. Turner, A.G. Riess, Astrophys. J. 569 (2002) 18. arXiv:astro-ph/0106051.
- [74] M. Visser, Classical Quantum Gravity 21 (2004) 2603. arXiv:gr-qc/0309109.
- [75] A.G. Riess, et al., Astrophys. J. 607 (2004) 665. arXiv:astro-ph/0402512.
- [76] J.D. Barrow, The Observatory 113 (1993) 210.
- [77] G.F.R. Ellis, M.A.H. MacCallum, Comm. Math. Phys. 12 (1969) 108.
- [78] M.A.H. MacCallum, in: E. Schatzman (Ed.), Cargèse Lectures in Physics, 1973, pp. 61–174.
- [79] C.B. Collins, G.F.R. Ellis, Phys. Rep. 56 (1979) 65.
- [80] J.D. Barrow, D.H. Sonoda, Phys. Rep. 139 (1986) 1.
- [81] A. Krasinski, Inhomogeneous Cosmological Models, Cambridge University Press, Cambridge, 1997.
- [82] J. Wainwright, G.F.R. Ellis, Dynamical Systems in Cosmology, Cambridge University Press, Cambridge, 1997.
- [83] A.A. Coley, Dynamical Systems and Cosmology, Springer, Berlin, 2003.
- [84] J. Plebanski, A. Krasinski, An Introduction to General Relativity and Cosmology, Cambridge University Press, Cambridge, 2006.
- [85] G.F.R. Ellis, Gen. Relativity Gravitation 38 (2006) 1003.
- [86] L.P. Hughston, K.C. Jacobs, Astrophys. J. 160 (1970) 147.
- [87] V.A. Belinskij, I.M. Khalatnikov, E.M. Lifshits, Adv. Phys. 19 (1970) 525.
- [88] L.D. Landau, E.M. Lifshitz, The Classical Theory of Fields, Pergamon Press, Oxford, 1975.
- [89] H. Stephani, et al., Exact Solutions of Einstein's Field Equations, Cambridge University Press, Cambridge, 2003.
- [90] L. Hsu, J. Wainwright, Classical Quantum Gravity 3 (1986) 1105.
- [91] J.D. Barrow, C.G. Tsagas, Classical Quantum Gravity 22 (2005) 825. arXiv:gr-qc/0411070.
- [92] C.W. Misner, Phys. Rev. Lett. 22 (1969) 1071.
- [93] J.D. Barrow, Phys. Rep. 85 (1982) 1.
- [94] J.D. Barrow, Nuclear Physics B 296 (1988) 697.
- [95] N.J. Cornish, J.J. Levin, Phys. Rev. Lett. 78 (1997) 998. arXiv:gr-qc/9605029.
- [96] C.B. Collins, S.W. Hawking, Astrophys. J. 180 (1973) 317.
- [97] A.S. Kompaneets, A.S. Chernov, Sov. Phys. JETP 20 (1965) 1303.
- [98] R. Kantowski, R.K. Sachs, J. Math. Phys. 7 (1966) 443.
- [99] J.D. Barrow, M.P. Dabrowski, Phys. Rev. D 55 (1997) 630. arXiv:hep-th/9608136.
- [100] E.M. Lifshitz, J. Phys. (Moscow) 10 (1946) 116.
- [101] E.M. Lifshitz, I.M. Khalatnikov, Adv. Phys. 12 (1963) 185.
- [102] R.K. Sachs, A.M. Wolfe, Astrophys. J. 147 (1967) 73.
- [103] J.M. Bardeen, Phys. Rev. D 22 (1980) 1882.
- [104] G.F.R. Ellis, M. Bruni, Phys. Rev. D 40 (1989) 1804.
- [105] G.F.R. Ellis, W. Stoeger, Classical Quantum Gravity 4 (1987) 1697.
- [106] J.M. Stewart, M. Walker, Proc. R. Soc. A 341 (1974) 49.
- [107] J.M. Stewart, Classical Quantum Gravity 7 (1990) 1169.
- [108] M. Bruni, et al., Classical Quantum Gravity 14 (1997) 2585. arXiv:gr-qc/9609040.
- [109] C.G. Tsagas, J.D. Barrow, Classical Quantum Gravity 14 (1997) 2539. arXiv:gr-qc/9704015.
- [110] A.A. Starobinskii, JETP Lett. 42 (1985) 152.
- [111] D. Langlois, F. Vernizzi, JCAP 2 (2007) 17. arXiv:astro-ph/0610064.

- [112] P.S. Letelier, *Phys. Rev. D* 22 (1980) 807.
- [113] H. Kodama, M. Sasaki, *Progr. Theoret. Phys. Sup.* 78 (1984) 1.
- [114] A. Challinor, A. Lasenby, *Astrophys. J.* 513 (1999) 1. [arXiv:astro-ph/9804301](#).
- [115] A.R. King, G.F.R. Ellis, *Comm. Math. Phys.* 31 (1973) 209.
- [116] C. Eckart, *Phys. Rev.* 58 (1940) 919.
- [117] P.K.S. Dunsby, M. Bruni, G.F.R. Ellis, *Astrophys. J.* 395 (1992) 54.
- [118] P.P. Kronberg, *Rep. Progr. Phys.* 57 (1994) 325.
- [119] R.M. Kulsrud, et al., *Astrophys. J.* 480 (1997) 481. [arXiv:astro-ph/9607141](#).
- [120] J.L. Han, R. Wielebinski, *Chin. J. Astron. Astrophys.* 2 (2002) 293. [arXiv:astro-ph/0209090](#).
- [121] M. Giovannini, *Internat. J. Modern Phys. D* 13 (2004) 391. [arXiv:astro-ph/0312614](#).
- [122] P.J. Greenberg, *Astrophys. J.* 164 (1971) 589.
- [123] J.D. Jackson, *Classical Electrodynamics*, Wiley, New York, 1975.
- [124] A. Kandus, C.G. Tsagas, *Mon. Not. R. Astron. Soc.* 385 (2008) 883. [arXiv:0711.3573](#).
- [125] Y.B. Zeldovich, A.A. Ruzmaikin, D.D. Sokoloff, *Magnetic Fields in Astrophysics*, Gordon and Breach, New York, 1983.
- [126] Y.B. Zel'Dovich, *Soviet Astron.* 13 (1970) 608.
- [127] J.D. Barrow, *Phys. Rev. D* 55 (1997) 7451. [arXiv:gr-qc/9701038](#).
- [128] C.G. Tsagas, R. Maartens, *Classical Quantum Gravity* 17 (2000) 2215. [arXiv:gr-qc/9912044](#).
- [129] K.S. Thorne, D. MacDonald, *Mon. Not. R. Astron. Soc.* 198 (1982) 339.
- [130] N.K. Spyrou, C.G. Tsagas, *Mon. Not. R. Astron. Soc.* 388 (2008) 187. [arXiv:0804.1702](#).
- [131] M.S. Madsen, *Astrophys. Space Sci.* 113 (1985) 205.
- [132] M.S. Madsen, *Classical Quantum Gravity* 5 (1988) 627.
- [133] G.F.R. Ellis, M.S. Madsen, *Classical Quantum Gravity* 8 (1991) 667.
- [134] F. Vernizzi, *Phys. Rev. D* 71 (2005) 061301. [arXiv:astro-ph/0411463](#).
- [135] M. Bruni, G.F.R. Ellis, P.K.S. Dunsby, *Classical Quantum Gravity* 9 (1992) 921.
- [136] G.F.R. Ellis, J. Hwang, M. Bruni, *Phys. Rev. D* 40 (1989) 1819.
- [137] J.D. Barrow, *Mon. Not. R. Astron. Soc.* 179 (1977) 47P.
- [138] G.F.R. Ellis, M. Bruni, J. Hwang, *Phys. Rev. D* 42 (1990) 1035.
- [139] T. Padmanabhan, *Structure Formation in the Universe*, Cambridge University Press, Cambridge, 1993.
- [140] P.J.E. Peebles, *The Large-Scale Structure of the Universe*, Princeton University Press, Princeton, 1980.
- [141] P. Coles, F. Lucchin, *Cosmology: The Origin and Evolution of Cosmic Structure*, Wiley, Chichester, 1995.
- [142] C.G. Tsagas, in: S. Cotsakis, E. Papantonopoulos (Eds.), *Cosmological Crossroads*, 2002, pp. 223–261.
- [143] V.F. Mukhanov, H.A. Feldman, R.H. Brandenberger, *Phys. Rep.* 215 (1992) 203.
- [144] A.R. Liddle, D.H. Lyth, *Cosmological Inflation and Large-Scale Structure*, Cambridge University Press, Cambridge, 2000.
- [145] J. Silk, *Nature* 215 (1967) 1155.
- [146] J. Silk, *Astrophys. J.* 151 (1968) 459.
- [147] P. Meszaros, *Astron. Astrophys.* 37 (1974) 225.
- [148] E.J. Groth, P.J.E. Peebles, *Astron. Astrophys.* 41 (1975) 143.
- [149] D.J. Eisenstein, et al., *Astrophys. J.* 633 (2005) 560. [arXiv:astro-ph/0501171](#).
- [150] S. Cole, et al., *Mon. Not. R. Astron. Soc.* 362 (2005) 505. [arXiv:astro-ph/0501174](#).
- [151] P.J.E. Peebles, *Astrophys. J. Lett.* 263 (1982) L1.
- [152] G.R. Blumenthal, et al., *Nature* 311 (1984) 517.
- [153] M. Davis, et al., *Nature* 356 (1992) 489.
- [154] A.R. Liddle, D.H. Lyth, *Phys. Rep.* 231 (1993) 1. [arXiv:astro-ph/9303019](#).
- [155] M. White, et al., *Mon. Not. R. Astron. Soc.* 276 (1995) L69. [arXiv:astro-ph/9508009](#).
- [156] D.N. Spergel, et al., *Astrophys. J. Suppl.* 148 (2003) 175. [arXiv:astro-ph/0302209](#).
- [157] C.G. Tsagas, J.D. Barrow, *Classical Quantum Gravity* 15 (1998) 3523. [arXiv:gr-qc/9803032](#).
- [158] C.G. Tsagas, R. Maartens, *Phys. Rev. D* 61 (2000) 083519. [arXiv:astro-ph/9904390](#).
- [159] J. Adams, et al., *Phys. Lett. B* 388 (1996) 253. [arXiv:astro-ph/9607043](#).
- [160] T.V. Ruzmaikina, A.A. Ruzmaikin, *Soviet Astron.* 14 (1971) 963.
- [161] C. Pitrou, J.P. Uzan, *Phys. Rev. D* 75 (2007) 087302. [arXiv:gr-qc/0701121](#).
- [162] S.W. Hawking, *Astrophys. J.* 145 (1966) 544.
- [163] P.K.S. Dunsby, B.A.C.C. Bassett, G.F.R. Ellis, *Classical Quantum Gravity* 14 (1997) 1215. [arXiv:gr-qc/9811092](#).
- [164] R. Maartens, C.G. Tsagas, C. Ungarelli, *Phys. Rev. D* 63 (2001) 123507. [arXiv:astro-ph/0101151](#).
- [165] C.G. Tsagas, *Classical Quantum Gravity* 19 (2002) 3709. [arXiv:gr-qc/0202095](#).
- [166] C. Caprini, R. Durrer, *Phys. Rev. D* 65 (2002) 023517. [arXiv:astro-ph/0106244](#).
- [167] M. Bruni, P.K.S. Dunsby, G.F.R. Ellis, *Astrophys. J.* 395 (1992) 34.
- [168] S.W. Goode, *Phys. Rev. D* 39 (1989) 2882.
- [169] A. Challinor, *Classical Quantum Gravity* 17 (2000) 871. [arXiv:astro-ph/9906474](#).
- [170] P.K.S. Dunsby, *Phys. Rev. D* 48 (1993) 3562.
- [171] T.E. Perko, R.A. Matzner, L.C. Shepley, *Phys. Rev. D* 6 (1972) 969.
- [172] G.F.R. Ellis, R. Maartens, *Classical Quantum Gravity* 21 (2004) 223. [arXiv:gr-qc/0211082](#).
- [173] G.F.R. Ellis, J. Murugan, C.G. Tsagas, *Classical Quantum Gravity* 21 (2004) 233. [arXiv:gr-qc/0307112](#).
- [174] T. Clifton, J.D. Barrow, *Phys. Rev. D* 72 (2005) 123003. [arXiv:gr-qc/0511076](#).
- [175] D.J. Mulryne, et al., *Phys. Rev. D* 71 (2005) 123512. [arXiv:astro-ph/0502589](#).
- [176] S. Mukherjee, et al., *Classical Quantum Gravity* 23 (2006) 6927. [arXiv:gr-qc/0605134](#).
- [177] A.S. Eddington, *Mon. Not. R. Astron. Soc.* 90 (1930) 668.
- [178] G.W. Gibbons, *Nuc. Phys. B* 292 (1987) 784.
- [179] J.D. Barrow, et al., *Classical Quantum Gravity* 20 (2003) L155. [arXiv:gr-qc/0302094](#).
- [180] E.R. Harrison, *Rev. Modern Phys.* 39 (1967) 862.
- [181] B. Lotic, W.G. Unruh, *Phys. Rev. D* 71 (2005) 044011. [arXiv:gr-qc/0408011](#).
- [182] A. Borde, A. Vilenkin, *Phys. Rev. Lett.* 72 (1994) 3305. [arXiv:gr-qc/9312022](#).
- [183] A. Borde, A. Vilenkin, *Phys. Rev. D* 56 (1997) 717. [arXiv:gr-qc/9702019](#).
- [184] A. Borde, A.H. Guth, A. Vilenkin, *Phys. Rev. Lett.* 90 (2003) 151301. [arXiv:gr-qc/0110012](#).
- [185] C. Molina-Paris, M. Visser, *Phys. Lett. B* 455 (1999) 90. [arXiv:gr-qc/9810023](#).
- [186] K. Gödel, *Rev. Modern Phys.* 21 (1949) 447.
- [187] B.R. Steadman, *Gen. Relativity Gravitation* 35 (2003) 1721.
- [188] O. Gron, S. Johannesen, 2007. [arXiv:gr-qc/0703139](#).
- [189] A. Krasinski, *J. Math. Phys.* 39 (1998) 2148. [arXiv:gr-qc/9707021](#).
- [190] I. Ozsváth, E. Schucking, *Amer. J. Phys.* 71 (2003) 801.
- [191] J.D. Barrow, 2006. [arXiv:physics/0612253](#).
- [192] J.D. Barrow, C.G. Tsagas, *Classical Quantum Gravity* 21 (2004) 1773. [arXiv:gr-qc/0308067](#).

- [193] F. de Felice, *Mon. Not. R. Astron. Soc.* 252 (1991) 197.
- [194] C. Barrabes, B. Boisseau, W. Israel, *Mon. Not. R. Astron. Soc.* 276 (1995) 432. arXiv:gr-qc/9505025.
- [195] M.A. Abramowicz, J.P. Lasota, *Acta Phys. Polon. B* 5 (1974) 327.
- [196] V.M. Rosa, P.S. Letelier, *Gen. Relativity Gravitation* 39 (2007) 1419. arXiv:gr-qc/0703100.
- [197] M.J. Rebouças, J. Tiomno, *Phys. Rev. D* 28 (1983) 1251.
- [198] J.D. Barrow, M.P. Dabrowski, *Phys. Rev. D* 58 (1998) 103502. arXiv:gr-qc/9803048.
- [199] P. Kanti, C.E. Vayonakis, *Phys. Rev. D* 60 (1999) 103519. arXiv:gr-qc/9905032.
- [200] J.D. Barrow, C.G. Tsagas, *Phys. Rev. D* 69 (2004) 064007. arXiv:gr-qc/0309030.
- [201] A.G. Doroshkevich, V.N. Lukash, I.D. Novikov, *Sov. Phys. JETP* 37 (1973) 739.
- [202] V.N. Lukash, *Nuovo Cimento B* 35 (1976) 268.
- [203] S. Hervik, A. Coley, *Classical Quantum Gravity* 22 (2005) 3391. arXiv:gr-qc/0505108.
- [204] J.M. Stewart, in: J.M. Stewart (Ed.), *Non-equilibrium Relativistic Kinetic Theory*, in: LNP, vol. 10, 1971, pp. 1–113.
- [205] S.L. de Groot, W.A. van Leeuwen, C.G. van Weert, *Relativistic Kinetic Theory*, North-Holland, Amsterdam, 1980.
- [206] J. Bernstein, *Kinetic Theory in the Expanding Universe*, Cambridge University Press, Cambridge and New York, 1988, 157p.
- [207] M. Bartelmann, P. Schneider, *Phys. Rep.* 340 (2001) 291. arXiv:astro-ph/9912508.
- [208] K.S. Thorne, *Mon. Not. R. Astron. Soc.* 194 (1981) 439.
- [209] G.F.R. Ellis, D.R. Matravers, R. Treciokas, *Ann. Phys.* 150 (1983) 455.
- [210] T. Gebbie, G.F.R. Ellis, *Ann. Phys.* 282 (2000) 285. arXiv:astro-ph/9804316.
- [211] A. Lewis, A. Challinor, *Phys. Rev. D* 66 (2002) 023531. arXiv:astro-ph/0203507.
- [212] R. Maartens, J. Triginer, D.R. Matravers, *Phys. Rev. D* 60 (1999) 103503. arXiv:astro-ph/9901213.
- [213] R.C. Tolman, *Relativity, Thermodynamics, and Cosmology*, Clarendon Press, Oxford, 1934.
- [214] J. Ehlers, P. Geren, R.K. Sachs, *J. Math. Phys.* 9 (1968) 1344.
- [215] W.R. Stoeger, R. Maartens, G.F.R. Ellis, *Astrophys. J.* 443 (1995) 1.
- [216] C.A. Clarkson, et al., *Gen. Relativity Gravitation* 35 (2003) 969. arXiv:gr-qc/0302068.
- [217] C.P. Ma, E. Bertschinger, *Astrophys. J.* 455 (1995) 7. arXiv:astro-ph/9401007.
- [218] U. Seljak, M. Zaldarriaga, *Astrophys. J.* 469 (1996) 437. arXiv:astro-ph/9603033.
- [219] S. Dodelson, E. Gates, A. Stebbins, *Astrophys. J.* 467 (1996) 10. arXiv:astro-ph/9509147.
- [220] K. Tomita, *Progr. Theoret. Phys.* 68 (1982) 310.
- [221] L.F. Abbott, R.K. Schaefer, *Astrophys. J.* 308 (1986) 546.
- [222] M. Bucher, A.S. Goldhaber, N. Turok, *Phys. Rev. D* 52 (1995) 3314. arXiv:hep-ph/9411206.
- [223] A. Challinor, A. Lasenby, *Phys. Rev. D* 58 (1998) 023001. arXiv:astro-ph/9804150.
- [224] A. Challinor, *Phys. Rev. D* 62 (2000) 043004. arXiv:astro-ph/9911481.
- [225] M.L. Wilson, *Astrophys. J.* 273 (1983) 2.
- [226] A. Lewis, *Phys. Rev. D* 70 (2004) 043518. arXiv:astro-ph/0403583.
- [227] K.S. Thorne, *Rev. Modern Phys.* 52 (1980) 299.
- [228] W. Hu, M. White, *Phys. Rev. D* 56 (1997) 596. arXiv:astro-ph/9702170.
- [229] W. Hu, et al., *Phys. Rev. D* 57 (1998) 3290. arXiv:astro-ph/9709066.
- [230] J. Lesgourgues, S. Pastor, *Phys. Rep.* 429 (2006) 307. arXiv:astro-ph/0603494.
- [231] A.G. Doroshkevich, Y.B. Zeldovich, R.A. Sunyaev, M.Y. Khlopov, *Sov. Astron. Lett.* 6 (1980) 252.
- [232] J.R. Bond, A.S. Szalay, *Astrophys. J.* 274 (1983) 443.
- [233] U. Seljak, A. Slosar, P. McDonald, *JCAP* 10 (2006) 14. arXiv:astro-ph/0604335.
- [234] M. Viel, M.G. Haehnelt, A. Lewis, *Mon. Not. R. Astron. Soc.* 370 (2006) L51. arXiv:astro-ph/0604310.
- [235] J.R. Bond, G. Efstathiou, J. Silk, *Phys. Rev. Lett.* 45 (1980) 1980.
- [236] S. Weinberg, *Phys. Rev. D* 69 (2004) 023503. arXiv:astro-ph/0306304.
- [237] W. Hu, D.J. Eisenstein, M. Tegmark, *Phys. Rev. Lett.* 80 (1998) 5255. arXiv:astro-ph/9712057.
- [238] R.J. Michney, R.R. Caldwell, *JCAP* 1 (2007) 14. arXiv:astro-ph/0608303.
- [239] J.M. Stewart, *Astrophys. J.* 176 (1972) 323.
- [240] G.F. Smoot, et al., *Astrophys. J. Lett.* 396 (1992) L1.
- [241] W. Hu, S. Dodelson, *Ann. Rev. Astron. Astrophys.* 40 (2002) 171. arXiv:astro-ph/0110414.
- [242] W. Hu, *Ann. Phys.* 303 (2003) 203. arXiv:astro-ph/0210696.
- [243] A. Challinor, in: K. Tamvakis (Ed.), *The Physics of the Early Universe*, in: LNP, vol. 653, 2005, p. 71.
- [244] A. Challinor, A. Lasenby, *Anisotropy and polarization of the cosmic microwave background*, *Living Reviews in Relativity*, 2008 (submitted for publication).
- [245] A.A. Penzias, R.W. Wilson, *Astrophys. J.* 142 (1965) 419.
- [246] M.J. Rees, *Astrophys. J. Lett.* 153 (1968) L1.
- [247] P.J.E. Peebles, *Astrophys. J.* 153 (1968) 1.
- [248] Y.B. Zeldovich, V.G. Kurt, R.A. Sunyaev, *Zhurnal Eksperimental noi i Teoreticheskoi Fiziki* 55 (1968) 278.
- [249] P.J.E. Peebles, J.T. Yu, *Astrophys. J.* 162 (1970) 815.
- [250] R.A. Sunyaev, Y.B. Zeldovich, *Astrophys. Space Sci.* 7 (1970) 3.
- [251] J.R. Bond, G. Efstathiou, *Astrophys. J. Lett.* 285 (1984) L45.
- [252] N. Kaiser, *Mon. Not. R. Astron. Soc.* 202 (1983) 1169.
- [253] G. Dautcourt, *Mon. Not. R. Astron. Soc.* 144 (1969) 255.
- [254] A.G. Polnarev, *Soviet Astron.* 29 (1985) 607.
- [255] R. Maartens, G.F.R. Ellis, W.R. Stoeger, *Phys. Rev. D* 51 (1995) 1525. arXiv:astro-ph/9501016.
- [256] R. Maartens, G.F.R. Ellis, W.R. Stoeger, *Phys. Rev. D* 51 (1995) 5942.
- [257] W.R. Stoeger, M.E. Araujo, T. Gebbie, *Astrophys. J.* 476 (1997) 435. arXiv:astro-ph/9904346.
- [258] P.K.S. Dunsby, *Classical Quantum Gravity* 14 (1997) 3391. arXiv:gr-qc/9707022.
- [259] R. Maartens, T. Gebbie, G.F.R. Ellis, *Phys. Rev. D* 59 (1999) 083506. arXiv:astro-ph/9808163.
- [260] T. Gebbie, P.K.S. Dunsby, G.F.R. Ellis, *Ann. Phys.* 282 (2000) 321. arXiv:astro-ph/9904408.
- [261] A. Challinor, *Gen. Relativity Gravitation* 32 (2000) 1059. arXiv:astro-ph/9903283.
- [262] A. Lewis, A. Challinor, A. Lasenby, *Astrophys. J.* 538 (2000) 473. arXiv:astro-ph/9911177.
- [263] A. Lewis, *Phys. Rev. D* 70 (2004) 043011. arXiv:astro-ph/0406096.
- [264] A. Blanchard, J. Schneider, *Astron. Astrophys.* 184 (1987) 1.
- [265] S. Cole, G. Efstathiou, *Mon. Not. R. Astron. Soc.* 239 (1989) 195.
- [266] E.V. Linder, *Mon. Not. R. Astron. Soc.* 243 (1990) 353.
- [267] U. Seljak, *Astrophys. J.* 463 (1996) 1. arXiv:astro-ph/9505109.
- [268] A. Lewis, A. Challinor, *Phys. Rep.* 429 (2006) 1.
- [269] A. Gruzinov, W. Hu, *Astrophys. J.* 508 (1998) 435. arXiv:astro-ph/9803188.
- [270] R.A. Sunyaev, Y.B. Zeldovich, *Comment. Astrophys. Space Phys.* 4 (1972) 173.
- [271] J.P. Ostriker, E.T. Vishniac, *Astrophys. J. Lett.* 306 (1986) L51.
- [272] E.T. Vishniac, *Astrophys. J.* 322 (1987) 597.

- [273] M. Birkinshaw, *Phys. Rep.* 310 (1999) 97. [arXiv:astro-ph/9808050](#).
- [274] M.J. Rees, D.W. Sciama, *Nature* 217 (1968) 511.
- [275] M. Kamionkowski, A. Kosowsky, A. Stebbins, *Phys. Rev. Lett.* 78 (1997) 2058. [arXiv:astro-ph/9609132](#).
- [276] U. Seljak, M. Zaldarriaga, *Phys. Rev. Lett.* 78 (1997) 2054. [arXiv:astro-ph/9609169](#).
- [277] M. Zaldarriaga, U. Seljak, *Phys. Rev. D* 58 (1998) 023003. [arXiv:astro-ph/9803150](#).
- [278] N. Bartolo, S. Matarrese, A. Riotto, *JCAP* 6 (2006) 24. [arXiv:astro-ph/0604416](#).
- [279] N. Bartolo, S. Matarrese, A. Riotto, *JCAP* 1 (2007) 19. [arXiv:astro-ph/0610110](#).
- [280] A. Challinor, A. Lewis, *Phys. Rev. D* 71 (2005) 103010. [arXiv:astro-ph/0502425](#).
- [281] V.B. Berestetskii, E.M. Lifshitz, L.P. Pitaevskii, *Quantum Electrodynamics* (Course of Theoretical Physics—Pergamon International Library of Science, Technology, Engineering and Social Studies, 2nd Engl. ed., Pergamon Press, Oxford, 1982).
- [282] S. Chandrasekhar, *Radiative Transfer*, Dover, New York, 1960.
- [283] J.N. Goldberg, et al., *J. Math. Phys.* 8 (1967) 2155.
- [284] G. Dautcourt, K. Rose, *Astronom. Nachr.* 299 (1978) 13.
- [285] D.J. Fixsen, J.C. Mather, *Astrophys. J.* 581 (2002) 817.
- [286] N. Bartolo, et al., *Phys. Rep.* 402 (2004) 103.
- [287] C. Burigana, R. Salvaterra, *Mon. Not. R. Astron. Soc.* 342 (2003) 543. [arXiv:astro-ph/0301133](#).
- [288] M. Kamionkowski, A. Kosowsky, A. Stebbins, *Phys. Rev. D* 55 (1997) 7368. [arXiv:astro-ph/9611125](#).
- [289] K.W. Ng, G.C. Liu, *Internat. J. Modern Phys. D* 8 (1999) 61. [arXiv:astro-ph/9710012](#).
- [290] G. Chon, et al., *Mon. Not. R. Astron. Soc.* 350 (2004) 914. [arXiv:astro-ph/0303414](#).
- [291] A. Challinor, F. van Leeuwen, *Phys. Rev. D* 65 (2002) 103001. [arXiv:astro-ph/0112457](#).
- [292] S. Bildhauer, *Classical Quantum Gravity* 6 (1989) 1171.
- [293] A.D. Challinor, M.T. Ford, A.N. Lasenby, *Mon. Not. R. Astron. Soc.* 312 (2000) 159. [arXiv:astro-ph/9905227](#).
- [294] N. Itoh, S. Nozawa, Y. Kohyama, *Astrophys. J.* 533 (2000) 588. [arXiv:astro-ph/0005390](#).
- [295] J. Portsmouth, E. Bertschinger, *ArXiv astrophysics e-prints*, 2004. [arXiv:astro-ph/0412095](#).
- [296] J. Portsmouth, E. Bertschinger, *ArXiv astrophysics e-prints*, 2004. [arXiv:astro-ph/0412094](#).
- [297] M. Zaldarriaga, U. Seljak, E. Bertschinger, *Astrophys. J.* 494 (1998) 491. [arXiv:astro-ph/9704265](#).
- [298] J. García-Bellido, et al., *Phys. Rev. D* 55 (1997) 4596. [arXiv:astro-ph/9608106](#).
- [299] M. Bucher, J.D. Cohn, *Phys. Rev. D* 55 (1997) 7461. [arXiv:astro-ph/9701117](#).
- [300] W. Hu, N. Sugiyama, *Astrophys. J.* 444 (1995) 489. [arXiv:astro-ph/9407093](#).
- [301] Y.B. Zeldovich, R.A. Sunyaev, *Astrophys. Space Sci.* 4 (1969) 301.
- [302] E. Komatsu, et al. *ArXiv e-prints* 803, 2008. [arXiv:0803.0547](#).
- [303] M. Zaldarriaga, *Phys. Rev. D* 55 (1997) 1822.
- [304] L. Page, et al., *Astrophys. J. Suppl.* 170 (2007) 335. [arXiv:astro-ph/0603450](#).
- [305] N. Turok, U. Pen, U. Seljak, *Phys. Rev. D* 58 (1998) 023506. [arXiv:astro-ph/9706250](#).
- [306] L.M. Krauss, M. White, *Phys. Rev. Lett.* 69 (1992) 869. [arXiv:hep-ph/9205212](#).
- [307] J.E. Lidsey, P. Coles, *Mon. Not. R. Astron. Soc.* 258 (1992) 57P.
- [308] A.R. Liddle, D.H. Lyth, *Phys. Lett. B* 291 (1992) 391. [arXiv:astro-ph/9208007](#).
- [309] A.A. Starobinskiĭ, *J. Exp. Theoret. Phys. Lett.* 30 (1979) 682.
- [310] W. Hu, M.M. Hedman, M. Zaldarriaga, *Phys. Rev. D* 67 (2003) 043004. [arXiv:astro-ph/0210096](#).
- [311] D. O’Dea, A. Challinor, B.R. Johnson, *Mon. Not. R. Astron. Soc.* 292 (2007) [arXiv:astro-ph/0610361](#).
- [312] M. Tucci, et al., *Mon. Not. R. Astron. Soc.* 360 (2005) 935. [arXiv:astro-ph/0411567](#).
- [313] W. Hu, *Phys. Rev. D* 65 (2002) 023003. [arXiv:astro-ph/0108090](#).
- [314] A. Lewis, A. Challinor, N. Turok, *Phys. Rev. D* 65 (2002) 023505. [arXiv:astro-ph/0106536](#).
- [315] W.R. Stoeger, M.E. Araujo, T. Gebbie, *Astrophys. J.* 522 (1999) 559.
- [316] U.S. Nilsson, et al., *Astrophys. J. Lett.* 522 (1999) L1. [arXiv:astro-ph/9904252](#).
- [317] C.B. Collins, S.W. Hawking, *Mon. Not. R. Astron. Soc.* 162 (1973) 307.
- [318] J.D. Barrow, R. Juszkiewicz, D.H. Sonoda, *Mon. Not. R. Astron. Soc.* 213 (1985) 917.
- [319] E.F. Bunn, P.G. Ferreira, J. Silk, *Phys. Rev. Lett.* 77 (1996) 2883. [arXiv:astro-ph/9605123](#).
- [320] A. Kogut, G. Hinshaw, A.J. Banday, *Phys. Rev. D* 55 (1997) 1901. [arXiv:astro-ph/9701090](#).
- [321] T.R. Jaffe, et al., *Astrophys. J. Lett.* 629 (2005) L1. [arXiv:astro-ph/0503213](#).
- [322] A. Pontzen, A. Challinor, *Mon. Not. R. Astron. Soc.* 380 (2007) 1387. [arXiv:astro-ph/0706.2075](#).
- [323] M. Panek, *Astrophys. J.* 388 (1992) 225.
- [324] H. van Elst, G.F.R. Ellis, *Classical Quantum Gravity* 15 (1998) 3545. [arXiv:gr-qc/9805087](#).
- [325] G.F.R. Ellis, H. van Elst, R. Maartens, *Classical Quantum Gravity* 18 (2001) 5115. [arXiv:gr-qc/0105083](#).
- [326] G.F.R. Ellis, C.G. Tsagas, *Phys. Rev. D* 66 (2002) 124015. [arXiv:astro-ph/0209143](#).
- [327] J.E. Gunn, J.R.I. Gott, *Astrophys. J.* 176 (1972) 1.
- [328] M. Bruni, S. Matarrese, O. Pantano, *Astrophys. J.* 445 (1995) 958. [arXiv:astro-ph/9406068](#).
- [329] M. Bruni, in: P. Coles, V. Martinez, M.J. Pons-Borderia (Eds.), *Mapping, Measuring, and Modelling the Universe*, in: ASP Conference Series, vol. 94, 1996, pp. 31–37.
- [330] E. Bertschinger, B. Jain, *Astrophys. J.* 431 (1994) 486. [arXiv:astro-ph/9307033](#).
- [331] S.F. Shandarin, et al., *Phys. Rev. Lett.* 75 (1995) 7. [arXiv:astro-ph/9407050](#).
- [332] M.F. Shirokov, I.Z. Fisher, *Sov. Astron.* 6 (1963) 699.
- [333] G.F.R. Ellis, in: B. Bertotti, F. de Felice, A. Pascolini (Eds.), *General Relativity and Gravitation Conference*, 1984, pp. 215–288.
- [334] É.É. Flanagan, *Phys. Rev. D* 71 (2005) 103521. [arXiv:hep-th/0503202](#).
- [335] C.M. Hirata, U. Seljak, *Phys. Rev. D* 72 (2005) 083501. [arXiv:astro-ph/0503582](#).
- [336] M. Giovannini, *JCAP* 9 (2005) 9. [arXiv:astro-ph/0506715](#).
- [337] H. Alnes, M. Amarzguioui, Ø Grøn, *JCAP* 1 (2007) 7. [arXiv:astro-ph/0506449](#).
- [338] A. Ishibashi, R.M. Wald, *Classical Quantum Gravity* 23 (2006) 235. [arXiv:astro-ph/0509108](#).
- [339] T. Buchert, *Classical Quantum Gravity* 23 (2006) 817. [arXiv:gr-qc/0509124](#).
- [340] R.A. Vanderveld, É.É. Flanagan, I. Wasserman, *Phys. Rev. D* 74 (2006) 023506. [arXiv:astro-ph/0602476](#).
- [341] A. Paranjape, T.P. Singh, *JCAP* 03 (2008) 023. [arXiv:0801.1546](#).
- [342] T. Buchert, *Gen. Relativity Gravitation* 40 (2008) 467. [arXiv:0707.2153](#).
- [343] S. Räsänen, *ArXiv e-prints* 801, 2008. [arXiv:0801.2692](#).
- [344] G.F.R. Ellis, T. Buchert, *Phys. Lett. A* 347 (2005) 38. [arXiv:gr-qc/0506106](#).
- [345] R.A. Isaacson, *Phys. Rev.* 166 (1968) 1263.
- [346] R.M. Zalaletdinov, *Bull. Astron. Soc. India* 25 (1997) 401. [arXiv:gr-qc/9703016](#).
- [347] T. Buchert, J. Ehlers, *Astron. Astrophys.* 320 (1997) 1. [arXiv:astro-ph/9510056](#).
- [348] T. Buchert, *Gen. Relativity Gravitation* 32 (2000) 105. [arXiv:gr-qc/9906015](#).
- [349] T. Buchert, *Gen. Relativity Gravitation* 33 (2001) 1381. [arXiv:gr-qc/0102049](#).
- [350] S. Räsänen, *Classical Quantum Gravity* 23 (2006) 1823. [arXiv:astro-ph/0504005](#).

- [351] J.D. Barrow, C.G. Tsagas, *Classical Quantum Gravity* 24 (2007) 1023. [arXiv:gr-qc/0609078](#).
- [352] E.W. Kolb, ArXiv e-prints, 709, 2007. [arXiv:0709.3102](#).
- [353] A.A. Coley, ArXiv e-prints 704, 2007. [arXiv:0704.1734](#).
- [354] A. Paranjape, T.P. Singh, *Phys. Rev. D* 76 (2007) 044006. [arXiv:gr-qc/0703106](#).
- [355] T. Buchert, *Classical Quantum Gravity* 22 (2005) L113. [arXiv:gr-qc/0507028](#).
- [356] T. Buchert, in: P. Coles, V. Martinez, M.J. Pons-Borderia (Eds.), *Mapping, Measuring, and Modelling the Universe*, in: ASP Conference Series, vol. 94, 1996, pp. 349–355.
- [357] S.M. Carroll, et al., *Phys. Rev. D* 70 (2004) 043528. [arXiv:astro-ph/0306438](#).
- [358] J.D. Bekenstein, *Phys. Rev. D* 70 (2004) 083509. [arXiv:astro-ph/0403694](#).
- [359] M. Milgrom, *Astrophys. J.* 270 (1983) 365.
- [360] D. Baumann, L. McAllister, *Phys. Rev. D* 75 (2007) 123508. [arXiv:hep-th/0610285](#).
- [361] D. Clowe, et al., *Astrophys. J. Lett.* 648 (2006) L109. [arXiv:astro-ph/0608407](#).
- [362] M.J. Jee, et al., *Astrophys. J.* 661 (2007) 728. [arXiv:0705.2171](#).
- [363] E. Bertschinger, *Astrophys. J.* 648 (2006) 797. [arXiv:astro-ph/0604485](#).