35

Cosmological Perturbation Theory and the Variational Principle in Gravitation

C D Laatz

A thesis submitted to
The University of Cape Town
In fulfillment of the requirements for the degree
Master of Science

November 30, 2000

Abstract

In this thesis firstly the theory of Relativistic Cosmological Perturbations is studied, in the process being reviewed over the period 1960 – 1993. Secondly the Variational Principle, apropos of gravitation, is formulated and discussed. These two fields are then synthesised via a variational formulation of General Relativity and Cosmological Perturbation Theory. In the process new light is shed on Covariant Perturbation Theory via the development of generalised alternative variables, culminating in a unique variational formulation.

Acknowledgements

I would like to thank both Prof. GFR Ellis and Dr PKS Dunsby for their guidance and patience in both the practical and theoretical aspects involved in composing this thesis. I would also like to express my gratitude to all the members of the UCT Cosmology Group for their stimulating conversation and advice, to whom I owe much of my present understanding of Cosmology and research in general.

Contents

Chapter 1

Introduction

Since the inception thereof eighty-five years ago, Einstein's Theory of General Relativity has humbled intellectually, and challenged technically, successive generations of physicists. The former response has generated a literature characterised by a demeanour of profound awe and aesthetic appreciation, while the latter has seen the emergence of a plethora of academic treatises comprising progressively more complex mathematical formalisms.

When confronted and captivated by a physical theory as intuitively appealing as General Relativity, it is not difficult to digress into philosophical and aesthetic reverie, perhaps consequently losing sight of the more practical motivation and vindication thereof. This practicality lies somewhat more modestly displaced from the elevated ontological debates around the theory, in that it attempts more soberly to explain existing physical observations, and predict new ones. Indeed, one may argue that this is the foremost objective of physics in general, and that the philosophical issues are of peripheral or secondary nature due to their inherent subjectivity, although intellectually rewarding.

This empirical attribute of physics, by its practical nature, manifests itself in a variety of mathematical formalisms of varying complexity. Apropos of this, it has been repeatedly noted by numerous scientists outside of the regime of General Relativity and Cosmology that its inherent mathematical complexity has detracted from, and obfuscated the initial simplicity and philosophical appeal of the theory. Although this opinion does enjoy considerable support, it is the partial aim of this thesis to enhance the intuitive appreciation and understanding of General Relativity through physical application thereof to Cosmological Perturbation Theory, as well as formulating the theory via the Calculus of Variations. In short, it is the belief of the author that any physical theory, no matter how philosophical and grandiose, is motivated and vindicated primarily through the utility and empiricism of its practical application. In Cosmology, the most expansive application of General Relativity, Cosmological Perturbation Theory can be seen as such a vindication. The latter is a scientific endeavour which demonstrates the validity of Einstein's Theory of General Relativity through the tangible physical results which it delivers; results which comprise observable and verifiable properties of the physical universe.

Consequently, in elucidating Cosmological Perturbation Theory in the subsequent chapters, emphasis is placed on the technical workings of the fundamental research that has gone into the field over the past half century by discussing the relevant landmark results and scientific papers. This is intended not only to provide insight into the functioning of General Relativity, but also to demonstrate how fundamental research is done in physics, a process which is best described through example.

The Variational Principle, in contrast to General Relativity, entered the scientific arena in a far more pragmatic vein and Baroque spirit via Fermat's Principle in the seventeenth century. Commencing with its use in the formulation of Classical Mechanics through the efforts of D'Alembert, Hamilton and Lagrange in the nineteenth century, through its parallel formulation of General Relativity in the early twentieth century, and culminating in the formulations of Quantum Field Theory and String Theory in the last fifty years, it has become increasingly evident that the Variational Principle underpins most of Physics in, at the very least, a mathematically unified sense. Due to this utility as a basic formulation, conceptual and mathematical, of so many branches of Physics, modern Physics in particular, the question of an inherent physicality in the Variational Principle has been earnestly debated by theoreticians and experimentalists alike, particularly in the latter half of the twentieth century. The two principal stances on the issue are the following: either the Variational Principle can be seen purely as a convenient and elegant mathematical tool used pragmatically in formulating Physics, or it can be regarded as a fundamental physical principle underlying nature. These two issues have by no means yet been resolved, nor is it the purpose of this thesis to resolve them; rather, one aim here is to illustrate both points of view via a mathematical formulation of the Variational Principle, and application of the principle to Cosmological Perturbation Theory.

Another aim, as alluded to previously, is to cover Cosmological Perturbation Theory in its own right for the reasons expounded above. However, one is consequently inclined to wonder as to the motivation behind selecting two such seemingly disparate topics as Cosmological Perturbation Theory and the Variational Principle for comparison and discussion. In motivating this choice it suffices to say, quite honestly, that it is a purely subjective decision by the author, one based on a number of personal reasons: firstly, an aesthetic appreciation of the philosophical development of General Relativity; secondly, the vindication of the theory through practical application via Cosmological Perturbation Theory; thirdly, a fascination with the utility and philosophy behind the Variational Principle; and fourthly, an enhanced understanding and appreciation of both Cosmological Perturbation Theory and the Variational Principle through studying their interrelationship. Consequently, this thesis clearly reflects a personal fascination with two particular branches of science. This is in apposition to what is all too often observed when perusing scientific treatises across the academic spectrum: the tenuous and strained attempted justification of a given scientist for his or her research into a particular field, a motivation which quite often comprises an inflated assessment of importance and inane promises of future ground-breaking results. In this author's opinion, an affinity to scientific endeavour, an unbridled personal curiosity, and the belief that unshackled academic freedom optimises scientific progress, is sufficient as a justification.

Firstly then, in expounding the above topics, an approach will be adopted whereby the relevant requisite mathematics is presented in full at the beginning, and in a compact form. This is done so as to minimise technical digressions when discussing the principal topics in turn, thus optimising conceptual clarity. This method also simplifies mathematical referencing, and results in the thesis as whole being complete in that it expounds unambiguously the underlying mathematics. This philosophy is also intended at conveying the aesthetic mathematical appeal of Differential Geometry and the Variational Calculus, a sound understanding of which is crucial in comprehending the Variational Principle and General Relativity. Consequently, Differential Geometry and the Calculus of Variations appear together as chapters one and two of the Part One of the thesis, entitled 'The Underlying Mathematics and Physics'. In the process, it is intended that many of the mathematical concepts of Differential Geometry and the Variational Calculus will be enhanced through later application. In completing Part One, a chapter comprising the basics of Gravitation is included; this provides the necessary background for Cosmological Perturbation Theory.

Secondly, having established the preliminaries, Part Two covers the numerous approaches to Cosmological Perturbation Theory. The method adopted herein, is to present each approach separately by starting with the basic theory and then progressing to the relevant literature, the emphasis being placed on the methodology, various practical techniques and theoretical paradigms. The philosophy behind discussing the various fundamental scientific papers which comprise this literature, is the demonstration of scientific research and progress. Quite often when one encounters a summary of given topic in Science, one is presented with a revisionist, sanitised slant which would seem to indicate that Science always progresses in a perfectly logical sequential, linear and neat fashion; this is also a slant which doesn't do justice to the labour and functioning of research. Hence, by looking at the seminal papers that have created Cosmological Perturbation Theory, one not only gains insight into how the field has developed, but also an appreciation of the technical demands of the subject by noting the success, failures and ambiguities. It is also believed here that an understanding of the technical details of the seminal developments greatly enhances an intuitive understanding of the subject; hence liberty is taken in including many of the details behind these results. However, as this is indeed a vast area of research, the literature is indeed extensive; consequently, it has been necessary not only to select the seminal research papers and results, but also to restrict these to the time span 1960 – 1993. The subsequent years have seen a veritable explosion in research into cosmological perturbation theory, especially in the light of the COBE large scale structure observational results of the early 1990's, and more recently the smaller scale BOOMERanG results. This has been accompanied theoretically by the inclusion of kinetic theory methods and results. These more recent developments have seen a renewed interest in especially observational Cosmology, as they have provided for the first time tangible evidence for or against the numerous Cosmological models and theories; this has resulted in bringing Cosmology back from a largely theoretical and mathematical discipline to one based in observational Physics. These recent developments cannot be stressed enough, especially in the light of Cosmology in general, and for the future. Indeed, for these reasons, much of the theory and results subsequent to the early 1990's can be separately treated review-wise.

This approach is thus different from standard reviews on Perturbation Theory in that it focuses on a time span as opposed to a particular subcategory of the theory. It is intended thus to provide a perspective of the numerous approaches to Perturbation

Theory, and a regime in which to compare and assess them, while at the same time maintaining a suitable air of generality.

In formulating what is here termed 'metric' Perturbation Theory, a more systematic approach has been attempted by the author; systematic in that the formal theory has been collated and synthesised from the various provenances and placed collectively at the beginning of Part Two, before reviewing the standard research results and literature. This has been done so as to clarify the conceptual basis of the theory, thus optimising an understanding of the differences between the various paradigms. In addition to reviewing the standard material, in Chapter 5 on Covariant Perturbation Theory a section on alternative perturbation variables has been included; this material is the author's independent contribution, and is intended to systemise the standard formulations in Covariant Perturbation Theory.

Finally, Part Three attempts to synthesise the above by formulating General Relativity and Cosmological Perturbation Theory in terms of the Variational Principle. In the course of the variational formulation of Cosmological Perturbation Theory, and in addition to the formulations apropos of Metric Perturbation Theory, there is the author's own contribution to Covariant Perturbation Theory via the unique formulation of an action and Lagrangian for this branch of Cosmological Perturbation Theory.

In treating the subject matter of this thesis using the approach outlined above, it has been deemed useful to include the references for Part I at the end of each chapter therein, as these pertain predominantly to book references; while for Part II and Part III the references, exclusively journal references, are incorporated in the bibliography at the very end.

Part I

The Underlying Mathematics and Physics

"The book of nature lies continuously open before our eyes (I speak of the universe) but it can't be understood without first learning to understand the language and characters in which it is written. It is written in mathematical language, and its characters are geometrical figures." Galileo Galilei (*Il Saggiatore*)

The aim of this section is to establish a sound mathematical and physical framework in which ultimately to study the theory of perturbation theory. The physical theory which will be used here to model cosmological perturbation theory will be that of Einstein's General Relativity. As this is inherently a geometrical theory requiring the notions of curved spaces with differential structures, it is formulated within the mathematical regime of Differential Geometry.

Consequently, a knowledge of the fundamentals of Differential Geometry is a prerequisite to understanding the calculations, mathematical manipulations and interpretations of subsequent results in the following chapters. The relevant mathematics pertaining thereto will be provided in chapter 2. This chapter will also focus on the formulation of integration over a manifold, as this is an area which is sometimes technically awkward and conceptually misleading; this will provide the structure for a variational approach to Gravitation, as it incorporates the notion of an 'action integral' and variation thereof apropos of the manifold. The tetrad formulation is included for similar reasons, but more for the sake of completion as, although it will not be reinforced as a *modus operandi* later, it is a significant formalism utilised in much of the current research done in the covariant approach to Cosmological Perturbation theory.

The second requisite branch of mathematics for the subsequent studies is the Calculus of Variations, pertaining to applications of the Variational Principle. This too, is a technical necessity and will be seen to be of particular use in simplifying and elucidating several variational calculations.

Following the mathematical preliminaries, an introduction into the foundations of General Relativity and Cosmology will be given, formulated within the afore-mentioned mathematical regime.

As this thesis is intended to provide a working as well theoretical understanding of the Variational Principle in Gravitation and Perturbation Theory, the above topics will be dealt with in some detail.

Chapter 2

Differential Geometry and Tensor Calculus

"The idea that physicists would in future have to study the theory of tensors created real panic amongst them following the first announcement that Einstein's predictions had been verified."
A. Whitehead

2.1 Manifolds and Tensors

The primary mathematical objects that underly the technical formulation of General Relativity are tensors; however, before one can define a tensor one needs to consider the underlying incorporating structure. Hence one is introduced to the notion of a manifold, which is used to formulate physical space-time, as will be seen later, and enables one to utilise the associated structure in formulating General Relativity. As will be seen in Chapter 3, the need for considering a generalisation of the notions of differentiation and geometry arises from the interdependence of matter and geometry, as postulated by the Einstein Field Equations.

2.1.1 Manifolds

Conceptually speaking, a manifold $\mathcal{M}$ consists of two parts: A topological space $\mathcal{T}$, and a set of bijections $\{\phi_\alpha\}$ from the open sets $\{U_\alpha\}$ of the topological space into Euclidean n-space $\Re^n$. This can be stated formally as follows:
Given a topological space $\mathcal{T}$,

- For each point P of $\mathcal{T}$ there exists an open set U_α such that $P \epsilon U_\alpha$ and the set of all such open sets $\{U_\alpha\}$ covers $\mathcal{T}$;
- For each α there exists a bijective *chart* ϕ_α such that $\phi_\alpha : U_\alpha \to S_\alpha$ where $S_\alpha \subset \Re^n$, where the set of all charts $\mathcal{A} = \{\phi_\alpha\}$, $\alpha \epsilon I$ is referred to as an *atlas*;
- For any two open sets such that $U_\alpha \cap U_\beta \neq \emptyset$, there exists a C^∞ bijection $\phi_{\beta\alpha} = \phi_\beta \circ {\phi_\alpha}^{-1}$ such that $\phi_{\beta\alpha} \ : \ \phi_\alpha[U_\alpha \cap U_\beta] \subset S_\alpha \ \to \ \phi_\beta[U_\alpha \cap U_\beta] \subset S_\beta$.

Such a topological space $\mathcal{T}$ together with the above defined atlas $\mathcal{A}$ is called a *manifold* $\mathcal{M}$.

In simplified terms one thus sees that 1) each separate chart of the atlas constitutes a co-ordinate system as it assigns a particular n-tuple to each point in the topological space; 2) the manifold is *locally* Euclidean (isomorphic to a region of Euclidean space) due to the definition of a chart, but globally the manifold may differ from Euclidean space; 3) One cannot necessarily find a *single* chart to cover the entire topological space, unless it is isomorphically Euclidean; an example of this is illustrated below. In the physical context, the topology will beassumed to be *Hausdorff* and *paracompact*,as defined in the appendix.

The simplest example of a manifold is $\Re^n$ itself. Naturally this requires only one chart which covers itself. However, note that a finite set of points does not constitute a manifold, as this is not locally isomorphic to a region of $\Re^n$.

The surface of a sphere is another example of a manifold, and can be shown by application of the famous *fixed point theorem* to require at least two separate charts to cover it; for example, if one uses polar co-ordinates one runs into a problem at the poles.

In physical contexts it is readily seen that, in the simplest manifestation, a manifold will be used to represent the physical space-time of General Relativity itself; however, as will be seen, the concept of a manifold being suitably abstract, can be generalised well beyond this level of simplicity. When formulating Cosmological Perturbation Theory later, the utility of the mathematical intricacies of the manifold definition will become more apparent.

Submanifolds

Having defined the notion of a manifold, one can go one step farther in defining the notion of a *sub*manifold; following Schutz, this is defined as follows: a p-dimensional submanifold $\mathcal{N}$ of an n-dimensional manifold $\mathcal{M}$ is a collection of points of $\mathcal{M}$ having the property whereby in some open neighbourhood $\mathcal{U}$ of an arbitrary point Q in $\mathcal{N}$ there exists a co-ordinate system for M in which the points of $\mathcal{N}$ in that neighbourhood are defined by $x^{p+1} = x^{p+2} = \ldots = x^n = 0$. It follows trivially that this definition is consistent with the definitive properties of a manifold, as expounded above, thus guaranteeing that a submanifold is a manifold in its own right.

Physically, the notion of a submanifold permits the application of the subsequent manifold and tensor theory to certain regions of space-time which are of particular relevance in the study of Cosmology, as will be seen later. Such examples of submanifolds in Minkowski space-time are the future and past null cone of an observer, and a constant-time hypersurface - this will be expounded in more detail later. In general, a hypersurface in an n-dimensional manifold is an $(n-1)$ submanifold of that manifold. As will be seen shortly, the way in which one chooses such a submanifold is usually linked to *Frobenius' Theorem.*

2.1.2 Tangent Spaces and Tensors

Tangent Spaces

By the above definition, one may assign an n-tuple $x^a \equiv x^1, \ldots x^n$ to every point P in the underlying topological space; this defines a 'point' in the manifold. Note that this definition of a point does not satisfy the criteria for a vector, as the notions of vector addition *etc.* are ill-defined, if not meaningless; for example, how does one 'add' two different points in a general topological space? This is often an area of misconception. In order to gain such a vectorial structure, one defines a vector as an element of the *tangent space* at a point P in the manifold. This is achieved through the formulation of a vector as *directional derivative operator* acting on a point $P = x^a$ as follows:

$$\mathbf{v} \equiv \sum_a v^a \frac{\partial}{\partial x^a} \quad . \tag{2.1}$$

This is defined such that the operator $\mathbf{v}$ maps the point P onto the underlying field of real numbers as *a linear functional* in the algebraic sense, and is referred to as a *contravariant* vector. By this definition one also notices that the tangent space can be defined as the collection of all such vectors formulated in terms of the *co-ordinate basis* $\{\mathrm{e}_a\} = \left\{\frac{\partial}{\partial x^a}\right\}$; namely, such an arbitrary vector would be $\mathbf{u} = u^a \frac{\partial}{\partial x^a}$ where the Einstein summation convention across like indices is henceforth used. This tangent space is now readily seen to constitute the desired vector space, which can in turn trivially be shown to be an n-dimensional manifold itself. It is important to note that a tangent space is defined *at a specific point.* If one were to transform from $\{\mathrm{e}_a\}$ to another co-ordinate basis $\{\mathrm{e}'_a\}$ at P, corresponding to a different chart ϕ', one would have the following:

$$\frac{\partial}{\partial x^{a'}} = \frac{\partial x^a}{\partial x^{a'}} \frac{\partial}{\partial x^a} \quad . \tag{2.2}$$

For an arbitrary vector $\mathbf{v}$ in this new co-ordinate system, the components would thus obey:

$$v^{a'} = \frac{\partial x^{a'}}{\partial x^a} v^a \tag{2.3}$$

$$\equiv J_a^{a'} v^a \quad . \tag{2.4}$$

This is the basic vector transformation law, such that the reverse transformation J_a^c is defined by:

$$J_a^c J_b^a = \delta_b^c \quad , \tag{2.5}$$

where δ_b^c is the well-known Krönecker delta. Equivalent to the definition 2.1, one may consider the tangent vector in parametric form. Given a parametrically defined C^1 curve $l(t)$ in $\mathcal{M}$, i.e. a map of an interval of the real line into $\mathcal{M}$, one may define a contravariant vector *tangent* to $l(t)$ at the point $l(t_0)$ as being the operator $\left(\frac{\partial}{\partial t}\right)_l\Big|_{t_0}$ which maps each C^1 function f at t_0 onto the number $\left(\frac{\partial f}{\partial t}\right)_l\Big|_{t_0}$; i.e. the derivative of f in the *direction* of $l(t)$. Hence, in terms of a local co-ordinate system x^a in the neighbourhood of $l(t_0)$:

$$\left(\frac{\partial f}{\partial t}\right)_l\Big|_{t_0} = \frac{dx^a}{dt}\frac{\partial f}{\partial x^a}\Big|_{l(t_0)} \quad , \tag{2.6}$$

and hence, equivalent to the previous definition 2.1, one has:

$$\mathbf{v} = v^a \frac{\partial}{\partial x^a} \equiv \frac{dx^a}{dt}\frac{\partial}{\partial x^a} \quad , \tag{2.7}$$

which yields $v^a = \frac{dx^a}{dt}$ as the components of the vector. Physically, if one treats the parameter as proper time, then $v^a = \frac{dx^a}{dt}$ can be interpreted as a velocity quantity. Conversely, one can state that, associated with every vector $\mathbf{v}$ there exists an integral curve $l(t)$ in the manifold, such that the definition 2.6 holds. Hence, for the above arbitrarily defined function f on $l(t)$:

$$\mathbf{v}(f) = \frac{dx^a}{dt}\frac{\partial f}{\partial x^a} = v^a \frac{\partial f}{\partial x^a} \quad . \tag{2.8}$$

Naturally, one can also now define the dual (tangent) space $T^{\dagger}(\mathcal{M})$ associated with $T(\mathcal{M})$: if one has $\mathbf{v} \in T(\mathcal{M})$ then there exists a *covariant vector* or *one-form* $\psi \in T^{\dagger}(\mathcal{M})$ such that $\psi : \mathbf{v} \rightarrow \Re$ where $\Re$ refers to the underlying field of real numbers; i.e. ψ is a linear functional. Bearing the basis representation in mind, one may define then a dual space basis vector $\mathbf{e}^a$ as mapping the contravariant vector $\mathbf{v}$ onto its component v^a. In terms of the inner product associated with the linear functional, one therefore has:

$$\langle \mathbf{e}^a, \mathbf{v} \rangle = v^a \tag{2.9}$$

Hence, considering the basis for the tangent space one must have:

$$\langle \mathbf{e}^a, \mathbf{e}_b \rangle = \delta^a_b \tag{2.10}$$

Using a co-ordinate basis, for which one thus has $\mathbf{e}_a = \frac{\partial}{\partial x^a}$, one then immediately deduces that the basis for the dual space must be the set of differentials dx^a:

$$\langle dx^a, \frac{\partial}{\partial x^b} \rangle = \frac{\partial x^a}{\partial x^b} = \delta^a_b \tag{2.11}$$

Given then an arbitrary scalar function $f(x^a)$ defined on the manifold, one can define the differential of $f(x^a)$ as $df = \frac{\partial f}{\partial x^a} dx^a$ in terms of the basis of differentials; hence, for any contravariant vector $\mathbf{v}$ one has:

$$\langle df, \mathbf{v} \rangle = v^a \frac{\partial f}{\partial x^a} \tag{2.12}$$

and hence the differential is itself a *covariant* vector, with the components thereof being a normal to the surface $f(x^a)$.

This formulation is conceptually convenient as it enables one to envisage contravariant vectors as being *tangents* to curves lying within the surface $f(x^a)$, covariant vectors as being the *normal* to that surface, and the inner product of the two as the *gradient* of the surface $f(x^a) = 0$. Notice in the above formulation the index notation whereby an upper index indicates the components of a contravariant vector, as in v^a, while the lower index indicates the components of a covariant vector, as in v_a. it will shortly be explained, using the notion of a *metric tensor*, how to calculate the contravariant *and* covariant components of a vector and tensor in general.

Tensors

Having defined the notion of a vector in the tangent space, as well as the dual space, one can define a general type (m, n) tensor as being a multilinear functional $\mathbf{U}$ acting on m copies of tangent spaces $T^1 \ldots T^m$ and n copies of dual spaces $T_1^\dagger \ldots T_n^\dagger$:

$$\mathbf{U} : T^1 \otimes \ldots T^m \otimes T_1^\dagger \otimes \ldots T_n^\dagger \rightarrow \Re \quad , \tag{2.13}$$

where the symbol $\otimes$ indicates the standard Cartesian product, and refers to the multilinear functional nature of $\mathbf{U}$. Naturally then, in terms of the bases such a tensor can be written as:

$$\mathbf{U} = U^{a_1 \ldots a_m}_{b_1 \ldots b_n} \frac{\partial}{\partial x^{a_1}} \otimes \ldots \frac{\partial}{\partial x^{a_m}} \otimes dx^{b_1} \otimes \ldots dx^{a_n} \quad ; \tag{2.14}$$

such that the upper indices represent the contravariant components, and the lower indices the covariant components, as denoted in the definition 2.13. Under a co-ordinate transformation $x^a \rightarrow x^{a'}$, where the 'primed' sign indicates a new co-ordinate system, the above tensor then transforms as:

$$T^{a'_1 \ldots a'_m}_{b'_1 \ldots b'_n} = J^{a'_1}_{c_1} \ldots J^{a'_m}_{c_m} J^{d_1}_{b'_1} \ldots J^{d_n}_{b'_n} T^{c_1 \ldots c_m}_{d_1 \ldots d_n} \quad . \tag{2.15}$$

This is the definitive tensor co-ordinate transformation law, definitive in that it is, in some contexts, used to *define* a tensor; that is, one can define a tensor as an algebraic object which obeys the transformation property 2.15. This is probably the best interpretation of tensors in general as it is difficult, indeed impossible in many cases, to visualise them; alternatively, it suffices to say that tensors are the natural algebraic extensions of covariant and contravariant vectors. Some simple examples of type-two tensors are the various forms of energy-momentum stress tensors in classical physics, such as those containing the mechanical stresses for a given material in Engineering applications, or the inertia tensor for rigid bodies as formulated in Classical Mechanics, or even the Faraday tensor of Electromagnetism. From these examples one might be more familiar with the component matrix representation; however, due to the contravariant and covariant components, standard matrix algebra doesn't always apply: only when working in flat Euclidean space is this possible as such tensors, which are then referred to as *Cartesian* tensors, have the property whereby the contravariant and covariant components of a given tensor are identical. In general though, when one has the manifold property of *curvature*, as will be defined later in this chapter, the contravariant and covariant components are not equivalent, as will be seen.

Symmetrisation

As symmetry plays such a pivotal role in much of Physics, it is necessary to provide a suitable, useful and unambiguous mathematical framework in which to formulate symmetrical quantities. As the primary quantity used here will be, as for reasons previously expounded, the tensor, the notion of a *symmetric* tensor will thus be required.

A tensor is said to be *symmetric* in a pair of indices if it remains unchanged when these two indices are interchanged; e.g. for a four-tensor: $T_{abcd} = T_{acbd}$ is symmetric in the second and third indices b and c. the notation to indicate such a symmetry includes a pair of round brackets around the indices in question: $T_{abcd} = T_{a(bc)d}$ if and only if T_{abcd} is symmetric in the second and third indices b and c. Likewise, a tensor T_{abcd} is said to be *skew-symmetric* in the indices b and c if $T_{abcd} = -T_{acbd}$ and is indicated with square brackets: $T_{a[bc]d}$. For symmetrisation or skew-symmetrisation for any pair of non-adjacent indices, the notation of the *Bach bracket* is used: $T_{abcd} = T_{(a|bc|d)}$ or $T_{[a|bc|d]}$ if T_{abcd} is symmetric or skew-symmetric in the first and fourth indices respectively, such that the vertical lines exclude the indices between them from the symmetrisation. Note that one can always extricate a symmetric and anti-symmetric part from any (m, n) tensor. For a two-tensor, it can be decomposed exactly into two such parts:

$$\begin{aligned} T_{ab} &= \frac{1}{2}(T_{ab} + T_{ba}) + \frac{1}{2}(T_{ab} - T_{ba}) \\ &\equiv T_{(ab)} + T_{[ab]} \quad . \end{aligned} \tag{2.16}$$

However, for a general type (m, n) tensor, it cannot be decomposed into symmetric and anti-symmetric parts only - there will, in general, remain a part which is neither symmetric nor anti-symmetric in such a decomposition. For such tensors though, one can define symmetric and anti-symmetric parts as follows:

$$\begin{aligned} T_{(a_1 \ldots a_n)} &= \tfrac{1}{n!} \{\text{sum over permutations of indices}\} \\ T_{[a_1 \ldots a_n]} &= \tfrac{1}{n!} \{\text{alternating sum over permutations of indices}\} \end{aligned} \tag{2.17}$$

The Metric Tensor

In the conventional treatment of General Relativity the metric (in the standard use of the word) is regarded as the principal quantity, as this describes the geometrical structure of the space-time. In Metric Perturbation Theory it provides a simple quantity to perturb initially, as will be seen.

In full generality, a metric tensor is any symmetric, non-degenerate type $(0, 2)$ tensor $\mathbf{g}$ defined on the manifold. Hence, one can associate a scalar quantity, namely the *magnitude*:

$$u = |\mathbf{u}| = (|\mathbf{g}(\mathbf{u}, \mathbf{u})|)^{\frac{1}{2}} \quad , \tag{2.18}$$

with any vector $\mathbf{u}$. Two vectors $\mathbf{u}, \mathbf{v}$ are said to be *orthogonal* if:

$$\mathbf{g}(\mathbf{u}, \mathbf{v}) = 0 \quad . \tag{2.19}$$

By symmetry, one has that $\mathbf{g}(\mathbf{u}, \mathbf{v}) = \mathbf{g}(\mathbf{v}, \mathbf{u})$. With respect to a co-ordinate basis $\mathbf{e}_a = \frac{\partial}{\partial x^a}$, one may express the metric as:

$$\mathbf{g} = g_{ab} dx^a \otimes dx^b \quad , \tag{2.20}$$

such that the components are:

$$g_{ab} = \mathbf{g}(\mathbf{e}_a, \mathbf{e}_b) \quad . \tag{2.21}$$

One can now define the concept of a *length* or *distance* between points *in* the manifold as follows: Let $l(t)$ be some parametrically defined curve in the manifold incorporating points $P = l(a)$ and $Q = l(b)$, and with tangent vector $\frac{\partial}{\partial t}$ such that $\mathbf{g}(\frac{\partial}{\partial t}, \frac{\partial}{\partial t})$ has the same sign along all points of $l(t)$; then the path length between the points P and Q is defined as:

$$\mathcal{L} = \int_a^b \left(\left| \mathbf{g}\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}\right) \right| \right)^{\frac{1}{2}} dt \quad . \tag{2.22}$$

The infinitesimal distance ds along such a curve is usually expressed as:

$$ds^2 = g_{ab} dx^a dx^b \quad . \tag{2.23}$$

Now, assuming that $\mathbf{g}$ is *non-degenerate*; that is, there does not exist a non-zero vector $\mathbf{u}$ such that $\mathbf{g}(\mathbf{u}, \mathbf{v}) = 0$ for all $\mathbf{v}$, then there exists a unique symmetric *dual* tensor with respect to the dual basis $\{\mathbf{e}^a\}$ with components g^{ab} such that:

$$g^{ab} g_{ac} = \delta^b_c \quad . \tag{2.24}$$

Hence, in matrix form, the matrix of components g^{ab} is the inverse of the matrix of components g_{ab}, where non-degeneracy is equivalent to the matrices being non-singular. Hence the tensors g^{ab}, g_{ab} can be used to define, as in standard linear algebra, an isomorphism between contravariant and covariant vectors. Thus, u_a defined by:

$$u_a = g_{ab} u^b \tag{2.25}$$

is the covariant vector associated with the contravariant vector u^a. This result can be extended to tensors of arbitrary type; for example $g_{bd} T^{abc} = T^{a\ c}_{\ d}$. Hence g^{ab}, g_{ab} are used to 'raise' and 'lower' indices. In matrix form for type two tensors, $g_{ab} T^{ab} = T^a_{\ a}$ defines the *trace* of T^{ab}. Similarly, the magnitude of a vector is now $|\mathbf{u}| = \sqrt{g_{ab} u^a u^b} = \sqrt{u^a u_a}$. In this sense one sees that the metric is, as in the standard geometrical case, the natural inner product defined on space-time, as expected.

Note that, at any arbitrarily chosen point P, one can find an orthogonal basis such that at that point the matrix of components of the metric reduces to a diagonal matrix of positive and negative values. Hence one can formally define the *signature* of the metric as being the difference between the number of positive values and negative values. It can be shown that if $\mathbf{g}$ is non-degenerate and continuous on the manifold, the signature will remain constant throughout the manifold.

Forms and The Exterior Product

A type $(0, p)$ tensor $\mathbf{F}$ with components $F_{a_1 \ldots a_p}$ which is skew in all p indices is called a *p-form*; that is, $F_{a_1 \ldots a_p} = F_{[a_1 \ldots a_p]}$, using the anti-symmetrisation notation as before. Given then a p-form $F_{a_1 \ldots a_p}$ and q-form $G_{b_1 \ldots b_q}$, one can define a $(p+q)$-form, denoted $(F \wedge G)_{a_1 \ldots a_p b_1 \ldots b_q}$, in terms of their fully skew-symmetrised Cartesian product:

$$(F \wedge G)_{a_1 \ldots a_p b_1 \ldots b_q} \equiv F_{[a_1 \ldots a_p} G_{b_1 \ldots b_q]} \tag{2.26}$$

in component form. The above is also referred to as the *wedge product*. With this product, the space of all forms (defining scalars as zero forms) constitutes what is referred to as the Grassmann algebra of forms. As a Direct consequence of the above, one also has the following:

$$\mathbf{F} \wedge \mathbf{G} = (-1)^{pq} \mathbf{G} \wedge \mathbf{F} \quad . \tag{2.27}$$

A simple example of a form is the volume element which is used in formulating integrals and defining the determinants of type-two tensors; this will be developed in full shortly.

Relative Tensors

In ultimately formulating the notion of integration over a manifold which is required for the Variational Calculus approach to Gravitation later, one will need to formulate the following generalisation of a tensor. Consider the Jacobian of a co-ordinate transformation $x^a \to x^{a'}$:

$$J = \left| \frac{\partial x_a}{\partial x'^b} \right| \quad . \tag{2.28}$$

Then a *relative tensor* $\mathfrak{R}^{a_1 \ldots}_{b_1 \ldots}$ *of weight* W is defined by:

$$\mathfrak{R}'^{a_1 \ldots a_m}_{b_1 \ldots b_n} = J^W \frac{\partial x'^{a_1}}{\partial x^{c_1}} \cdots \frac{\partial x'^{a_m}}{\partial x^{c_m}} \frac{\partial x^{d_1}}{\partial x'^{b_1}} \cdots \frac{\partial x^{d_n}}{\partial x'^{b_n}} \mathfrak{R}^{c_1 \ldots c_m}_{d_1 \ldots d_m} \quad . \tag{2.29}$$

In the above definition the standard convention of using a Gothic letter to denote a relative tensor has been adopted. Hence it follows that a standard tensor is merely a relative tensor of weight 0. It also follows that one can combine relative tensors in the same manner as for standard tensors, with one exception: the product of a relative tensor of weight W_1 with one of weight W_2 is in turn a relative tensor of weight $W_1 + W_2$; a direct consequence of the definition above. Relative tensors of weight $+1$ are usually referred to as *tensor densities*; similarly, a relative scalar of unit weight is called a scalar density. These will be of specific importance later, simple examples of which are the metric determinant and the Lagrangian density, quantities which will be developed later.

2.2 Differential Structure

Naturally, in order to provide a suitable regime in which to formulate physics, it is necessary to impose a differential structure upon the above formulated manifold. The subsequent definitions will be seen to be broad generalisations of the differential concepts of Euclidean space, and will induce naturally the notion of *curvature*, the central geometric tenet of General Relativity.

In leading up to a generalised, covariant form of the standard Euclidean-space partial derivative, one can formulate a kind of derivative purely in terms of the Euclidean partial derivative, but covariant and without having to impose any additional structure on the manifold as will be necessary later; this is the *Lie derivative.*

2.2.1 The Lie Derivative

The Lie derivative describes the way in which a tensor field changes as it 'moves' along the integral curves of a vector field. In order to accomplish this, one needs to compare the tensor at a point $P = l(t_0)$ on an integral curve $l(t)$ of the vector field $\mathbf{u}$ with itself evaluated a small distance t along the curve from P at $Q = l(t_0 + t)$, i.e. an infinitesimal transformation. Hence one needs to define a mapping of the tensor from Q to P and look at the difference between this and the tensor evaluated at P divided by Δt as $\Delta t \to 0$. One can define this mapping as a diffeomorphism $\Phi_t : P \to Q = l(t_0 + t)$. For sufficiently small t this can be considered as $\Phi_t : \mathcal{M} \to \mathcal{M}$; hence one can generalise it as a map Φ_t^* on general tensors. Hence the Lie derivative of a tensor field $\mathbf{T}$ with respect to the vector field $\mathbf{u}$ is defined as:

$$\mathcal{L}_{\mathbf{u}}\mathbf{T} \equiv \lim_{t\to 0} \frac{1}{t}(\Phi_t^*\mathbf{T} - \mathbf{T})\bigg|_P \quad , \tag{2.30}$$

at any point P. One can most conveniently calculate the above using a co-ordinate basis. The result for a scalar f is:

$$\mathcal{L}_{\mathbf{u}} f = f_{,a}u^a \quad , \tag{2.31}$$

as can be expected; while for a vector $\mathbf{v}$ it is:

$$\begin{aligned} \mathcal{L}_{\mathbf{u}}\mathbf{v} &= (\mathbf{u}(\mathbf{v}))^b - (\mathbf{v}(\mathbf{u}))^b \\ &= u^b v^a{}_{,b} - v^b u^a{}_{,b} && (2.32) \\ &\equiv [\mathbf{u}, \mathbf{v}] \quad , && (2.33) \end{aligned}$$

where the notation in the last line is referred to as the *Lie bracket*; it is consequently trivial to verify the following result:

$$[\mathbf{u}, \mathbf{v}] = -[\mathbf{v}, \mathbf{u}] \tag{2.34}$$

However, for a covariant vector (one-form) $\mathbf{F}$ in a co-ordinate basis one can show that:

$$(\mathcal{L}_{\mathbf{u}}\mathbf{F})_b = u^a \frac{\partial F_b}{\partial x^a} + F_a \frac{\partial u^a}{\partial x^b} \quad . \tag{2.35}$$

It is also trivial to show that, if one chooses a co-ordinate system in which $\mathbf{u}$ is the co-ordinate basis vector $\frac{\partial}{\partial x^0}$ say, then:

$$\mathcal{L}_{\mathbf{u}}\mathbf{T} = \frac{\partial}{\partial x^0}\mathbf{T} \tag{2.36}$$

for any tensor $\mathbf{T}$. The above results are then easily extended to tensors of all types:

$$\begin{aligned}\mathcal{L}_{\mathbf{u}} T^{a_1\ldots a_m}_{b_1\ldots b_n} &= \frac{\partial}{\partial x^c} T^{a_1\ldots a_m}_{b_1\ldots b_n} u^c - T^{d a_2\ldots a_m}_{b_1\ldots b_n}\frac{\partial u^{a_1}}{\partial x^d} - \ldots - T^{a_1\ldots a_{m-1} d}_{b_1\ldots b_n}\frac{\partial u^{a_m}}{\partial x^d} \\ &\quad + T^{a_1\ldots a_m}_{d b_2\ldots b_n}\frac{\partial u^d}{\partial x^{b_1}} + \ldots + T^{a_1\ldots a_m}_{b_1\ldots b_{n-1} d}\frac{\partial u^d}{\partial x^{b_n}} \quad . \end{aligned} \tag{2.37}$$

The infinitesimal co-ordinate transformation derivation of the Lie derivative will be of particular importance later in formulating the notions of gauge invariance and transformations in Cosmological Perturbation Theory, particularly in the covariant approach.

Invariance

A vector field $\mathbf{v}$ is said to be *invariant* under another vector field $\mathbf{u}$ if the Lie bracket vanishes:

$$[\mathbf{u}, \mathbf{v}] = 0 \quad . \tag{2.38}$$

By the definition of the Lie derivative, this naturally implies that either field is completely dragged along by the other, the change (derivative) of one *along* the other being zero. Geometrically speaking, this results in the two fields fitting exactly together. Usually, one would associate some form of symmetry with this invariance; for example, regarding vector operation as a transformation, invariance under rotation implies axial symmetry. Hence one can associate such a symmetry directly with the vector field. The best illustration of this comes from consideration of the *Killing vector fields*: a vector $\mathbf{u}$ is said to be a Killing field if the following holds:

$$\mathcal{L}_{\mathbf{u}}\mathbf{g} = 0 \quad , \tag{2.39}$$

where $\mathbf{g}$ is the metric tensor; in component form this is:

$$u^c \frac{\partial g_{ab}}{\partial x^c} + g_{ac}\frac{\partial u^c}{\partial x^b} + + g_{cb}\frac{\partial u^c}{\partial x^a} = 0 \quad . \tag{2.40}$$

If one then chooses $\mathbf{u} = \frac{\partial}{\partial x^0}$, one has:

$$\frac{\partial g_{ab}}{\partial x^0} = 0 \quad , \tag{2.41}$$

implying that the metric is independent of the co-ordinate x^0, and thus the the manifold has a symmetry associated with the direction $\frac{\partial}{\partial x^0}$.

2.2.2 Frobenius' Theorem

Related to the Lie derivative and the notion of submanifolds, particularly hypersurfaces, is the theorem of Frobenius. This will be seen to provide a method for determining whether a set of vectors defined on a manifold form a hypersurface. This will have significant application later when considering space-time hypersurfaces in various Cosmological models. One sees intuitively that the requisite property of being a hypersurface and thus a submanifold is particularly important, as it thus admits application of many of the powerful, yet standard, results of Differential Geometry. It also provides a clear and unambiguous physical interpretation in many cases, as will be dealt with in detail later. The Theorem, which will not be proven here (see *Schutz* for a more thorough exposition), has two separate formulations: for vector fields and one-forms. The vector field version is stated as follows:

Consider an n-dimensional manifold $\mathcal{M}$ with p-dimensional submanifold $\mathcal{N}$. Then for every set of vectors V defined in some neighbourhood $\mathcal{U}$ in $\mathcal{N}$ having Lie Brackets with one another which are also linear combinations of the vector fields in the set (i.e. they form a Lie algebra), then the integral curves of these vector fields mesh smoothly to form a family of submanifolds each having the same dimension of the vector space the fields define at each point in $\mathcal{U}$, and with each such point lying in exactly one such submanifold. This set of submanifolds fills $\mathcal{U}$ completely, and forms what is termed a foliation of $\mathcal{U}$, each submanifold being a leaf of the foliation.

The theorem essentially rests on proving the result that for any such neighbourhood $\mathcal{U}$ there exist co-ordinates $\{x^1, \ldots x^p\}$ with a corresponding co-ordinate basis $\left\{\frac{\partial}{\partial x^1}, \ldots \frac{\partial}{\partial x^p}\right\}$ such that $\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0$. It follows naturally that the foliation of such a region $\mathcal{U}$ is not unique; for example, Euclidean three-space can be foliated by either concentric two-spheres, or parallel planes.

Before giving the one-form version, some definitions are required. Given the tangent space T_P defined at point P of a manifold $\mathcal{M}$, every set of one-forms $O_P = \{\alpha_i\}$ defines a vector subspace V_P, the elements of which each annihilate (i.e. zero inner product) each of O_P; V_P is thus called the *annihilator* of O_P. Similarly, all the one-forms at this same point P whose restriction to V_P vanishes is called the *complete ideal* of O_P. A complete ideal is also a *differential ideal* if for every α_i in the ideal, the differential $d\alpha_i$ is also contained therein. A set of one-forms $\{\alpha_i\}$ is called *closed* if each form $d\alpha_i$ is contained in the complete ideal generated by the α_i's. A useful consequence of the latter is that on an n-dimensional manifold, any linearly independent set of n or $(n-1)$ one-forms is closed. One can now state the one-form version of Frobenius theorem:

Given a linearly independent set of one-forms $F = \{\alpha_1, \ldots \alpha_p\}$ in an open neighbourhood $\mathcal{U}$ of an n-dimensional manifold $\mathcal{M}$ $(p < n)$. If and only if this set is closed, there exist functions A_{ij} and B_j, $i, j = 1 \ldots p$ such that:

$$\alpha_i = \sum_{j=1}^{p} A_{ij} dB_j \quad . \tag{2.42}$$

These forms α_i are termed surface forming. Consequently, this set of one-forms F define an $(n-p)$-dimensional submanifold, the tangent vectors of which are the annihilators of F; each point of $\mathcal{U}$ lies in exactly one such submanifold, and the set of all these submanifolds defined as such completely fills $\mathcal{U}$ as a foliation.

The equivalence between the above two forms of the theorem is, loosely speaking, that one is the 'dual' of the other (dual as in 'dual space', as expected). Firstly, the dual of the Lie algebra requirement of the vector form is the closure requirement of the one-form formulation; secondly, there is the dual correspondence between the tangent vector spaces: if the vector fields define a p-dimensional subspace of the tangent space at P, then they naturally define an $(n-p)$-dimensional subspace of the *dual* tangent space of one-forms by the requirement that these forms be annihilated by the original vectors.

2.2.3 The Exterior derivative

Having defined the notion of a p-form and the wedge product, one can associate with them a natural differential operator called the *exterior derivative*. This is an operator $\mathbf{d}$ which maps p-forms onto $(p+1)$-forms. Acting on a scalar function f it gives the one-form $\mathbf{d}f = \frac{\partial f}{\partial x^a} dx^a$, and is defined via the inner product as follows:

$$< \mathbf{d}f, \mathbf{u} >= \mathbf{u}(f) \quad , \tag{2.43}$$

for all vector fields $\mathbf{u}$. Hence the components are:

$$df_a = \frac{\partial f}{\partial x^a} \quad . \tag{2.44}$$

Acting on a p-form $\mathbf{Q} = Q_{a_1 \ldots a_p} dx^{a_1} \wedge dx^{a_2} \wedge \ldots dx^{a_p}$ it results in:

$$\mathbf{dQ} = \frac{\partial Q_{a_1 \ldots a_p}}{\partial x^b} dx^b \wedge dx^{a_1} \wedge dx^{a_2} \ldots \wedge dx^{a_p} \quad . \tag{2.45}$$

A direct consequence of the total skew-symmetric nature of the p-form is that the double exterior derivative vanishes:

$$\mathbf{d}(\mathbf{dQ}) = 0 \quad , \tag{2.46}$$

which can easily be proven using the above. Although this operation may appear abstract and remote, it is crucial in the formulation of integration over manifolds as will be seen shortly, and in the process yields many of the standard vector calculus results which are ubiquitous in Physics.

2.2.4 The Absolute and Covariant Derivatives

As a co-ordinate system is inherently defined on a manifold via the chosen atlas, and as a tensor is by definition a function of these co-ordinates, one may naturally define a standard partial derivative of a tensor with respect to this co-ordinate system. In doing so, one adopts the following notation:

$$\frac{\partial}{\partial x^b} T_{a_1 \ldots a_n} \equiv T_{a_1 \ldots a_n, b} \quad . \tag{2.47}$$

A simple calculation reveals that this quantity does not transform as a tensor under a co-ordinate transformation as in 2.15. However, one naturally would desire here a derivative, analogous to the partial derivative, which is still a tensorial quantity. The problem lies in the basic definition of a partial derivative in terms of a limit of the difference of the differentiated quantity *evaluated at two different points*. Note, however, that this problem is not inherent in the case of the derivative of a scalar function of the co-ordinates in the manifold, as is evidenced from equation 2.8.

One thus needs to define a differential operator which is a generalisation of the directional derivative, and which reduces to the partial derivative when acting on scalars. This is done by introducing an *affine connection* ∇ on the manifold $\mathcal{M}$. Such a connection ∇ at a point P of $\mathcal{M}$ is a rule which assigns to each vector field $\mathbf{v}$ at P a differential operator $\nabla_{\mathbf{v}}$ which maps an arbitrary vector field $\mathbf{u}$ into a vector field $\nabla_{\mathbf{v}}\mathbf{u}$, satisfying the following:

$$\nabla_{f\mathbf{u}+g\mathbf{v}}\mathbf{w} = f\nabla_{\mathbf{u}}\mathbf{w} + g\nabla_{\mathbf{v}}\mathbf{w} \tag{2.48}$$

$$\nabla_{\mathbf{u}}(\alpha\mathbf{v} + \beta\mathbf{w}) = \alpha\nabla_{\mathbf{u}}\mathbf{v} + \beta\nabla_{\mathbf{u}}\mathbf{w} \tag{2.49}$$

$$\nabla_{\mathbf{u}}(f\mathbf{v}) = \mathbf{u}(f)\mathbf{v} + f\nabla_{\mathbf{u}}\mathbf{v} \quad , \tag{2.50}$$

such that f, g are arbitrary scalar functions, while α, β are constants. The quantity $\nabla_{\mathbf{u}}\mathbf{v}$ is referred to as the *absolute* derivative of $\mathbf{v}$ with respect to ∇ in the direction $\mathbf{u}$ at P, and by construction is seen to be the generalisation of the ordinary total derivative. Analogously one can define the *covariant derivative* $\nabla\mathbf{v}$ of the vector $\mathbf{v}$ as being a type $(1,1)$ tensor which, when contracted with $\mathbf{u}$ gives $\nabla_{\mathbf{u}}\mathbf{v}$, and such that:

$$\nabla(f\mathbf{u}) = \mathbf{d}f \otimes \mathbf{u} + f\nabla\mathbf{u} \quad . \tag{2.51}$$

As the contraction of this with $\mathbf{u}$ yield the total derivative by construction, it follows that the covariant derivative is then the desired generalisation of the partial derivative; this will be clarified shortly.

If one now has some arbitrary contravariant basis $\{\mathbf{e}_a\}$ with its corresponding dual covariant basis $\{\mathbf{e}^a\}$ (not necessarily co-ordinate bases), then one can write the covariant derivative $\nabla\mathbf{u}$ in terms of components as follows:

$$\nabla\mathbf{u} = u^a{}_{;b}\mathbf{e}^b \otimes \mathbf{e}_a \quad , \tag{2.52}$$

such that the semi-colon indicates the covariant derivative consistent with the comma notation in 2.47; and where the connection is determined by differentiable functions of the co-ordinates $\Gamma^a{}_{bc}$ defined as:

$$\Gamma^a{}_{bc} =< \mathbf{e}^a, \nabla_{\mathbf{e}_b}\mathbf{e}_c > \Leftrightarrow \nabla\mathbf{e}_c = \Gamma^a{}_{bc}\mathbf{e}^b \otimes \mathbf{e}_a \quad . \tag{2.53}$$

Naturally then, for the absolute derivative one has:

$$\nabla_{\mathbf{v}}\mathbf{u} = u^a{}_{;b}v^b\mathbf{e}_a < \mathbf{e}^b, \mathbf{e}_b > \quad . \tag{2.54}$$

Hence, for any vector field $\mathbf{u}$, one has:

$$\nabla\mathbf{u} = \nabla(u^a\mathbf{e}_a) = du^a \otimes \mathbf{e}_a + u^a\Gamma^c{}_{da}\mathbf{e}^d \otimes \mathbf{e}_c \quad . \tag{2.55}$$

For co-ordinate bases 2.55 simplifies to:

$$u^a{}_{;b} = \frac{\partial u^a}{\partial x^b} + \Gamma^a{}_{bc}u^c \quad . \tag{2.56}$$

Hence, if one is given a continuous curve $l(t)$ in the manifold, then one can define the absolute derivative of a vector $\mathbf{u}$ *along the curve* as:

$$\frac{Du^a}{\partial t} \equiv \nabla_{\frac{\partial}{\partial t}} \mathbf{u} = \frac{\partial u^a}{\partial t} + \Gamma^a_{\ bc} u^c \frac{dx^b}{dt} \quad . \tag{2.57}$$

If then $\frac{Du^a}{\partial t} = 0$, one describes the vector field $\mathbf{u}$ as being *parallel propagated* along the curve; loosely speaking, this means that the vector 'maintains its direction' with respect to the curve as it moves along it. If the absolute derivative of a vector field with respect to its congruence of integral curves vanishes, the curves are called *geodesics*; these can naturally be seen as the generalisation of straight lines in Euclidean space, and are thus defined as:

$$u^a_{\ ;b} u^b = 0 = \frac{d^2 u^a}{dt^2} + \Gamma^a_{\ bc} \frac{dx^b}{dt} \frac{dx^c}{dt} \quad . \tag{2.58}$$

Notice that one can interpret $\Gamma^a_{\ bc}$ as being the 'a'th component of the covariant derivative of the 'c'th basis vector in the direction of the 'b'th basis vector. This most clearly illustrates the earlier assertion that the covariant derivative is the natural extension of the partial derivative, and that the extra terms arise from considering the change in the basis as well as the components of the vector. It is also easily verified that $\Gamma^a_{\ bc}$ is *not* a tensorial quantity, as under a co-ordinate transformation it does not satisfy 2.15.

One can extend the above definition to show that, for a general (m, n) tensor the covariant derivative becomes:

$$\begin{aligned} T^{a_1 \ldots a_m}_{b_1 \ldots b_n ; c} &= \frac{\partial}{\partial x^c} T^{a_1 \ldots a_m}_{b_1 \ldots b_n} + \Gamma^{a_1}_{\ cd} T^{d a_2 \ldots a_m}_{b_1 \ldots b_n} + \ldots \Gamma^{a_1}_{\ cd} T^{a_1 \ldots a_{m-1} d}_{b_1 \ldots b_n} \\ & \quad - \Gamma^d_{\ cb_1} T^{a_1 \ldots a_m}_{d b_2 \ldots b_n} \ldots - \Gamma^d_{\ cb_n} T^{a_1 \ldots a_m}_{b_1 \ldots b_{n-1} d} \quad . \end{aligned} \tag{2.59}$$

As will be seen in section 1.3, the connection encapsulates the physical curvature of the manifold; in the General Relativity context, this transpires as the physical curvature of space-time itself through the Riemann tensor. The quantity $\Gamma^a_{\ bc}$ will thus be the major perturbed quantity sought in later application.

The Metric Connection

Given a metric g_{ab} defined on a manifold, one can define a unique kind of connection by the stipulation:

$$g_{ab;c} = 0 \quad , \tag{2.60}$$

a condition which guarantees the preservation of scalar products (defined in terms of the metric) under parallel transfer along curves. Considering now the covariant derivative of basis vectors using the above condition, one can show that, given 2.60, the connections $\Gamma^a_{\ bc}$ take on the following form:

$$\Gamma^a_{\ bc} = \frac{1}{2} g^{ad} (g_{db,c} - g_{bc,d} + g_{cd,b}) \quad . \tag{2.61}$$

The unique connections thus associated with a metric are called the *Christoffel coefficients.*

Relative tensor differentiation

Using the same notions and definitions as in the previous section for tensor differentiation, one can show that the covariant derivative of a relative tensor transpires as:

$$\begin{aligned}\mathfrak{R}^{a_1\ldots a_m}_{b_1\ldots b_n;c} &= \mathfrak{R}^{a_1\ldots a_m}_{b_1\ldots b_n,c} + \Gamma^{a_1}_{dc}\mathfrak{R}^{da_2\ldots a_m}_{b_1\ldots b_n} + \ldots \Gamma^{a_m}_{dc}\mathfrak{R}^{a_1\ldots a_{m-1}d}_{b_1\ldots b_n} - \Gamma^{d}_{b_1c}\mathfrak{R}^{a_1\ldots a_m}_{db_2\ldots b_n} - \ldots \\ &\quad -\Gamma^{d}_{b_nc}\mathfrak{R}^{a_1\ldots a_m}_{b_1\ldots b_{n-1}d} - W\Gamma^{d}_{dc}\mathfrak{R}^{a_1\ldots a_m}_{b_1\ldots b_n} \quad .\end{aligned} \tag{2.62}$$

Note that in the case for a type $(1,0)$ tensor density one has:

$$\mathfrak{R}^{a}_{;a} = \mathfrak{R}^{a}_{,a} \quad , \tag{2.63}$$

that is, the covariant divergence is equal to the ordinary divergence. It also follows that both the quantities in 2.63 are scalar densities.

Application to the Lie Derivative

If one has a *symmetric* connection, *i.e.*: $\Gamma^{a}_{bc} = \Gamma^{a}{}_{cb}$, it follows immediately that the partial derivatives in the Lie derivative expression can be replaced by covariant derivatives:

$$\begin{aligned}\mathcal{L}_{\mathbf{u}}\mathbf{v} &= u^b v^a{}_{,b} - v^b u^a{}_{,b} \\ &= [\mathbf{u}, \mathbf{v}] \\ &= \nabla_{\mathbf{u}}\mathbf{v} - \nabla_{\mathbf{v}}\mathbf{u} \quad .\end{aligned} \tag{2.64}$$

Henceforth only symmetric connections will be assumed. In general though, one can define a *torsion tensor* as:

$$T^{a}_{bc} = \Gamma^{a}{}_{bc} - \Gamma^{a}{}_{cb} \quad . \tag{2.65}$$

Hence if the connection is symmetric, the torsion vanishes. Under such circumstances, the manifold is said to be *torsion-free*. The above result for the Lie derivative is now readily extended to tensors of all types:

$$\begin{aligned}\mathcal{L}_{\mathbf{u}}T^{a_1\ldots a_m}_{b_1\ldots b_n} &= T^{a_1\ldots a_m}_{b_1\ldots b_n;c}u^c - T^{da_2\ldots a_m}_{b_1\ldots b_n}u^{a_1}_{\ ;d} - \ldots - T^{a_1\ldots a_{m-1}d}_{b_1\ldots b_n}u^{a_m}_{\ ;d} \\ &\quad + T^{a_1\ldots a_m}_{db_2\ldots b_n}u^{d}_{\ ;x^{b_1}} + \ldots + T^{a_1\ldots a_m}_{b_1\ldots b_{n-1}d}u^{d}_{\ x^{b_n}} \quad .\end{aligned} \tag{2.66}$$

2.2.5 The Fermi Derivative

As will be seen in the formulation of General Relativity, one wishes to choose a *reference frame* of basis vectors at a point along a curve $l(t)$ in the manifold. One might then wish to investigate how this basis changes as one moves along the curve. If, for example, an orthonormal basis is chosen such that one of the basis vectors is set equal to the unit tangent vector of the curve, then one might wish to transport this basis along the curve in such a way so as to have a similar such basis at each point along the curve. This would amount to parallel transporting the basis vectors. However, if the basis vector originally set equal to the tangent vector to the curve were parallel transported, it would not remain tangent to the curve, unless the curve were a geodesic; thus the remaining vectors would not remain orthogonal to this vector. Bearing these geometrical considerations in mind, one can formulate another differential operation, called the *Fermi derivative*. It is defined as a *derivative along an integral curve $l(t)$ of a vector field* $\mathbf{v}$ as follows:

$$\frac{D_F \mathbf{u}}{\partial t} = \frac{D\mathbf{u}}{\partial t} - \mathbf{g}\left(\mathbf{u}, \frac{Dv}{\partial s}\right)\mathbf{v} + \mathbf{g}(\mathbf{u}, \mathbf{v})\frac{D\mathbf{v}}{\partial t} \quad , \tag{2.67}$$

or in component form:

$$\frac{D_F u^a}{\partial t} = \frac{Du^a}{\partial t} - u_b \frac{Dv^b}{\partial t} v^a + u_b v^b \frac{Dv^a}{\partial t} \quad . \tag{2.68}$$

One can see immediately that it satisfies the desired properties:

- If $l(t)$ is a geodesic, then:

$$\frac{D_F \mathbf{u}}{\partial t} = \frac{D\mathbf{u}}{\partial t}. \tag{2.69}$$

- The Fermi derivative of the tangent vector itself is zero:

$$\frac{D_F \mathbf{v}}{\partial t} = 0 \quad . \tag{2.70}$$

- If $\mathbf{u}, \mathbf{w}$ are vector fields along $l(t)$ such that $\frac{D_F \mathbf{u}}{\partial t} = \frac{D_F \mathbf{w}}{\partial t} = 0$, then $\mathbf{g}(\mathbf{u}, \mathbf{w})$ remains constant along $l(t)$;

- If $\mathbf{u}$ is a vector field along $l(t)$ orthogonal to $\mathbf{v}$ then:

$$\frac{D_F \mathbf{u}}{\partial t} = \frac{D\mathbf{u}}{\partial t}_{\perp} \quad , \tag{2.71}$$

where the sign $\perp$ indicates the projection of the derivative in the direction of $\mathbf{v}$. Hence one has the desired property that, if one propagates an orthonormal basis along an integral curve $l(t)$ of $\mathbf{v}$ such that the Fermi derivative of each basis vector vanishes, then one maintains an orthonormal basis *at each point along the curve*; this follows directly from the third property above.

2.3 Curvature

Considering the covariant derivative one notices that the second covariant derivatives of a tensorial quantity do not necessarily commute; that is, for a vector u^a:

$$u^a_{;bc} \neq u^a_{;cb} \quad . \tag{2.72}$$

A measure of this non-commutativity is given by the *Riemann* tensor $R^a_{\ bcd}$ defined by:

$$u^a_{;bc} - u^a_{;cb} = R^a_{\ ecb} u^e \quad . \tag{2.73}$$

The non-commutativity can be shown geometrically as corresponding to the fact that a vector parallel transported around a closed curve is not generally parallel to its original orientation. This gives rise to the notion of *curvature* of a space. If the second covariant derivatives *do* commute, the space is said to be *flat* or Euclidean, as evidenced by this geometric argument.

Curvature will be seen to be the most fundamental property of physical space and time, and the central feature of the Einstein Field Equations; a result in stark contrast with the absolute Euclidean assumption of classical Physics. In terms of a symmetric connection, one has:

$$R^a{}_{bcd} = \frac{\partial}{\partial x^c}\Gamma^a{}_{db} - \frac{\partial}{\partial x^d}\Gamma^a{}_{cb} + \Gamma^a{}_{cf}\Gamma^f{}_{db} - \Gamma^a{}_{df}\Gamma^f{}_{cb} \quad . \tag{2.74}$$

One can thus deduce the following symmetries:

$$R^a{}_{b(cd)} = 0 \tag{2.75}$$
$$R^a{}_{[bcd]} = 0 \tag{2.76}$$
$$R^a{}_{b[cd;e]} = 0 \quad , \tag{2.77}$$

the last identity being referred to as the *Bianchi identity*. One also defines the (symmetric) *Ricci tensor* as:

$$R_{bd} \equiv R^a{}_{bad} \quad , \tag{2.78}$$

and the *Ricci curvature scalar*:

$$R \equiv R^a{}_a \quad . \tag{2.79}$$

2.4 Integration

In order to synthesise the Variational Calculus with Differential Geometry which will be used in Part III, one requires a formal methodology for integration over a manifold. This is necessitated by the requirement of having an action integral in the variational formulation of General Relativity.

As a preliminary to the formulation of a general integral on a manifold, the notions of the Levi-Civita alternating symbol and determinants are first required, as well as a study of the metric determinant. This will formally be done using p-forms and the exterior product.

2.4.1 The Levi-Civita Symbol and Determinants

If one has a basis of 1-forms $\{\mathbf{e}^{a_1}, \ldots \mathbf{e}^{a_n}\}$, then one can define the following n-form:

$$\epsilon = n! \mathbf{e}^{a_1} \wedge \mathbf{e}^{a_2} \wedge \ldots \mathbf{e}^{e_n} \quad . \tag{2.80}$$

In a co-ordinate basis this gives, component-wise, the following:

$$\epsilon_{a_1 \ldots a_n} = n! \delta^1_{[a_1} \delta^2_{a_2} \ldots \delta^n_{a_n]} \quad , \tag{2.81}$$

which is an n-dimensional generalisation of the Krönecker delta δ^a_b. In the four dimensions of space-time this takes on the form of the *Levi-Civita alternating symbol*:

$$\varepsilon^{abcd} = \begin{cases} 1 \text{ if abcd is an even permutation of } 0,1,2,3 \\ -1 \text{ if abcd is an odd permutation of } 0,1,2,3 \\ 0 \text{ otherwise} \end{cases} \quad . \tag{2.82}$$

It can easily be shown that ϵ and hence the Levi- Civita alternating symbol, by definition, have to be tensor *densities*. Note that one can use ϵ to formulate the *determinant* of a two-tensor A^a_b as follows:

$$|A^a_b| = A^1_{a_1} A^2_{a_2} \ldots A^n_{a_n} \epsilon^{a_1 \ldots a_n} \quad . \tag{2.83}$$

2.4.2 Metric Tensor Determinants and Relations

The above can thus be used to evaluate the determinant of the metric tensor g_{ab}; this determinant shall be denoted as $|g^{ab}| = g$. One can show that the metric tensor determinant transforms under co-ordinate transformations as follows:

$$g' = \left| \frac{\partial x^a}{\partial x'^b} \right|^2 g \quad . \tag{2.84}$$

Hence it is a relative scalar of weight 2, and thus $\sqrt{|g|}$ is a scalar *density*: $\sqrt{|g'|} = J\sqrt{|g|}$. Consequently, by 2.63 the following important result holds for any *vector* A^a:

$$\left[\sqrt{|g|} A^a\right]_{;a} = \left[\sqrt{|g|} A^a\right]_{,a} \quad . \tag{2.85}$$

Similarly, one can derive:

$$g_{,c} = g g^{ab} g_{ab,c} \quad . \tag{2.86}$$

Hence, for the metric connection, the following holds:

$$\begin{aligned} g_{;a} &= g_{,a} - 2g\Gamma^c_{ca} \\ &= g_{,a} - g_{,a} \\ &= 0 \quad . \end{aligned} \tag{2.87}$$

Similarly,

$$[\sqrt{|g|}]_{;a} = 0 \quad , \tag{2.88}$$

so that, for any tensor $A^{a_1 \ldots a_m}_{b_1 \ldots b_n}$:

$$[\sqrt{|g|} A^{a_1 \ldots a_m}_{b_1 \ldots b_n}]_{;c} = \sqrt{|g|} A^{a_1 \ldots a_m}_{b_1 \ldots b_n ; c} \quad . \tag{2.89}$$

2.4.3 The Volume Element

As mentioned earlier, the generalised Levi-Civita symbol, as defined in 2.80, is a tensor density. In the following definitions this quantity will be required, but in a general tensor form, so that the subsequent expressions take on a co-ordinate invariant form and are thus fully covariant. This is achieved by multiplying 2.80 by a scalar factor of $\sqrt{|g|}$:

$$\boldsymbol{\eta} \equiv \sqrt{|g|}\boldsymbol{\epsilon} \quad , \tag{2.90}$$

which is easily verified as a tensor, having components:

$$\eta_{a_1 \ldots a_n} = \sqrt{|g|} n! \delta^1_{[a_1} \ldots \delta^n_{a_n]} \quad . \tag{2.91}$$

This quantity is called the *volume tensor*, because from it one can define an infinitesimal volume element $d^n V$, given arbitrary co-ordinate displacements $\mathbf{d}x^{a_i} = \mathbf{e}^{a_i}$ as a co-ordinate basis:

$$\begin{aligned} d^n V &= \frac{1}{n!}\boldsymbol{\eta} \\ &= \frac{1}{n!}\eta_{a_1 a_2 \ldots a_n} dx^{a_1} \wedge dx^{a_2} \wedge \ldots \wedge dx^{a_n} \\ &= \sqrt{|g|}\delta^1_{[a_1} \ldots \delta^n_{a_n]} dx^{a_1} \wedge \ldots \wedge dx^{a_n} \\ &= \sqrt{|g|} dx^1 dx^2 \ldots dx^n \\ &= \sqrt{|g|} d^n v \quad , \end{aligned} \tag{2.92}$$

where $d^n v$ is the co-ordinate volume. Consequently, as with conventional Euclidean geometry in curvilinear co-ordinates, the determinant of the metric, which then is simply the Jacobian, acts as a scaling factor relating co-ordinate volume $d^n v$ to physical volume $d^n V$. It also follows that this volume element, in being tensorial, is form-invariant under arbitrary co-ordinate transformations. When working in n-space, one refers to such 'volume' elements generically as 'n-surfaces', as the notion of volume is usually restricted to three-space. Note also that, directly analogous to the Euclidean case, one can associate an n-vector with any $(n-1)$-dimensional 'volume' element $d^{n-1}V$ in the following way:

$$d\sigma_a = n_a d^{n-1} V \quad , \tag{2.93}$$

which will henceforth be referred to as the *surface-element vector*; and such that n_a is a unit n-vector *orthogonal* to $(n-1)$-surface $d^{n-1}V$. This is the natural extension to the area vector used in 3-space multiple integrals. One is now in a position to define integrals over manifolds.

2.4.4 Integration over Manifolds

The foremost problem in formulating integrals of tensors over manifolds is that different tensors can only be added at the same point, as can naturally be expected from the problems encountered earlier in formulating differentiation. Thus, as a Riemann integral (to which case it will be restricted here) would, by definition, constitute the addition of infinitesimal quantities involving its integrand *at different points*, one is compelled to reformulate the notion of tensor integration. This is accomplished by noting that, unlike tensors, scalars can indeed be added at different points; Hence the problem of tensor integration can be obviated if one can formulate an integrand involving a tensor in some form analogous to a scalar. This is achieved by noting that an n-form defined on an n-dimensional manifold has a solitary component; when written in expanded form in terms of this component and a co-ordinate basis of differentials, one can therefore has:

$$G_{a_1 \ldots a_n} dx^{a_1} \wedge \ldots dx^{a_n} \equiv \frac{1}{n!} G_{1 \ldots n} dx^1 dx^2 \ldots dx^n \tag{2.94}$$

for an n-form $\mathbf{G}$. Given the fundamental characteristic property whereby manifolds are locally Euclidean in that one has by definition a set of homeomorphisms from 'sufficiently small' open neighbourhoods into Euclidean space, and are consequently covered by co-ordinate patches, the formulation of an integral over the entire manifold, or even a sufficiently large enough subset of it which needs to be covered by more than one co-ordinate patch, needs to incorporate a mechanism which globalises operations defined locally, taking into account the intersection of the covering open sets of the atlas. This is done by means of the *partition of unity* defined in appendix A. Bearing this and 2.94 in mind, one can now define the integral of the above n-form $\mathbf{G}$ over an n-dimensional manifold $\mathcal{M}$ with atlas $\{\mathcal{U}_\alpha, \phi_\alpha\}$ and partition of unity $\{f_\alpha\}$ as:

$$\int_{\mathcal{M}} \mathbf{G} = \frac{1}{n!} \sum_\alpha \int_{\phi_\alpha(\mathcal{U}_\alpha)} f_\alpha G_{12 \ldots n} dx^1 dx^2 \ldots dx^n \quad . \tag{2.95}$$

Note that this formulation results in the integration being performed in R^n, and can be shown to be independent of the partition of unity and atlas chosen. Consistent with the above definition, one then has for the volume element:

$$\int_{\mathcal{M}} \eta = \int_{\mathcal{M}} |g|^{\frac{1}{2}} dx^1 dx^2 \ldots dx^n \tag{2.96}$$

Hence η defines a positive-definite volume measure on the manifold, thus rendering an interpretation of the above integral as the volume of the manifold; volume in the sense generalised to n dimensions. One can now formulate the generalisation of Stokes' theorem for an $(n-1)$-form $\mathbf{F}$ using the exterior derivative:

$$\int_{\partial\mathcal{M}} \boldsymbol{F} = \int_{\mathcal{M}} \mathbf{d}\boldsymbol{F} \quad , \tag{2.97}$$

where $\partial\mathcal{M}$ is the $(n-1)$-dimensional boundary of $\mathcal{M}$. Using now 2.96, one can define the integral of a scalar function f over a submanifold $\mathcal{U}$ as:

$$\int_{\mathcal{U}} f dV = \int_{\mathcal{U}} f \eta \tag{2.98}$$

$$= \int_{\mathcal{U}} f \sqrt{|g|} dx^1 dx^2 \ldots dx^n \quad . \tag{2.99}$$

If one now has a general contravariant vector $\mathbf{X}$ with components X^a, then the contraction $\mathbf{X}.\boldsymbol{\eta}$ will be an $(n-1)$-form, hence admitting the following integral over the $(n-1)$-dimensional submanifold $\mathcal{U}$ of $\mathcal{M}$:

$$\begin{aligned}\int_{\mathcal{U}} X^a d\sigma_a &= \int_{\mathcal{U}} \boldsymbol{X}.\boldsymbol{\eta} \\ &= \frac{1}{(n-1)!}\int_{\mathcal{U}} X^a \eta_{a1\ldots(n-1)} dx^1..dx^{(n-1)} \quad ,\end{aligned} \tag{2.100}$$

such that $\boldsymbol{X}.\boldsymbol{\eta}$ has components:

$$(\boldsymbol{X}.\boldsymbol{\eta})_{a_2\ldots a_n} = X^{a_1}\eta_{a_1\ldots a_n} \quad , \tag{2.101}$$

which is consistent with 2.93. Equation 2.100 is then seen to define naturally a *surface integral* over the manifold. One can now derive Gauss' theorem using Stokes' theorem 2.97 on a compact neighbourhood $\mathcal{U}$ of $\mathcal{M}$:

$$\int_{\partial\mathcal{U}} X^a d\sigma_a = \int_{\mathcal{U}} X^a_{;a} dV \quad . \tag{2.102}$$

2.5 Co-Ordinate Systems and the Tetrad Formalism

As mentioned earlier, the covariant formulation of Cosmological Perturbation Theory can be conveniently expressed in terms of tetrads. Hence, for the sake of completion and insight, tetrads will briefly be touched upon here.

2.5.1 Tetrad Bases

In full generality, a *tetrad* can be defined as a basis set of four contravariant vectors *at a particular point in space-time* and which are thus formulated in terms of the relevant co-ordinate basis. Hence one can envisage a tetrad as being a particularly chosen *local* reference frame, such as the fluid co-moving frame defined later, a special case thereof; whereas a co-ordinate basis results from the co-ordinate system chosen to define the atlas associated with the manifold in question. The physical ramifications are self-evident, and need not be explained here.

As a particular case, in this context it will be assumed in addition that the tetrad basis fields are *constant* and *orthonormal*, and these will be denoted by $\{\mathbf{e}_A\}$ such that the tetrad index $A = 1, 2, 3, 4$ labels the *vectors*, and *not* their components; the convention of using capital Roman letters to label the tetrad indices will be maintained here. Each tetrad vector $\mathbf{e}_A$ can, as explained, be written in terms of a local co-ordinate basis:

$$\mathbf{e}_A = e^a_A \frac{\partial}{\partial x^a} \quad , \tag{2.103}$$

such that the e^a_A, the *tetrad components*, are functions of the local co-ordinates; i.e. $e^a_A = e^a_A(x^b)$. By definition one thus has for an arbitrary function f:

$$\mathbf{e}_A f = e^a_A \frac{\partial f}{\partial x^a} \quad , \tag{2.104}$$

while the tetrad components naturally must satisfy:

$$
\begin{aligned}
e^a_A e^B_a &= \delta^B_A && (2.105)\\
e^A_a e^b_A &= \delta^b_a \quad . && (2.106)
\end{aligned}
$$

Hence one can express an arbitrary tensor $T^{a_1 \ldots i_m}{}_{b_1 \ldots j_n}$ in terms of this tetrad basis as follows:

$$
T^{A_1 \ldots A_m}{}_{B_1 \ldots B_n} = e^{A_1}_{a_1} \ldots e^{A_m}_{a_m} e^{b_1}_{B_1} \ldots e^{b_n}_{B_n} T^{a_1 \ldots a_m}{}_{b_1 \ldots b_n} \quad . \tag{2.107}
$$

One notices from this formulation, that for a given fixed tetrad, any subsequent co-ordinate basis transformation will not effect the tetrad components of a tensor with respect to the tetrad basis. Similarly, for a fixed co-ordinate basis, a tetrad basis transformation will not effect the co-ordinate components of a tensor. This convenient property is the one of the reasons for choosing a tetrad basis as defined above.

As with co-ordinate basis formulations, it is useful to define a Levi-Civita alternating symbol. For tetrads this quantity is defined as with the co-ordinate case, except with tetrad indices: ϵ^{ABCD}. Hence one can calculate determinants of various quantities, specifically that of the tetrad itself (in four dimensions):

$$
e = \frac{1}{24} \epsilon^{abcd} \epsilon^{ABCD} e_{aA} e_{bB} e_{cC} e_{dD} \quad . \tag{2.108}
$$

Directional Derivatives

As a result of the above, one has:

$$
\begin{aligned}
g_{AB} &= g_{ab} e^a_A e^b_B \\
&= \mathbf{e}_A \cdot \mathbf{e}_B \\
&= \Pi_{AB} \quad , && (2.109)
\end{aligned}
$$

where Π_{AB} is the Minkowski metric, and follows from the orthonormality assumption. Bearing now in mind the tetrad definition 2.103, one can define a *tetrad directional derivative* of an arbitrary function f as follows:

$$
\begin{aligned}
f_{,A} &\equiv e^a_A \frac{\partial f}{\partial x^a} \\
&= e^a_A f_{,a} \quad . && (2.110)
\end{aligned}
$$

Generalising this to a type-two tensor:

$$
\begin{aligned}
T_{A,B} &\equiv e^b_B \frac{\partial}{\partial x^b} T_A \\
&= e^b_B \frac{\partial}{\partial x^b} e^a_A T_a \\
&= e^b_B [e^a_A T_{a,b} + T_c e^c_{A\ ,b}] \quad . && (2.111)
\end{aligned}
$$

Assuming a symmetric metric connection, one can replace the "commas" in the last line of the above with "semi-colons", indicating covariant differentiation with respect to the metric connection, resulting in:

$$T_{A,B} = T_{a,b} e^a_A e^b_B + e_{Ac;b} e^b_B e^c_C T^C \quad . \tag{2.112}$$

The Structure and Rotation Co-efficients

One can now define the *commutation relations* or *structure constants* $\gamma^A{}_{BC}$ related to the tetrad in terms of the commutator bracket as follows:

$$[\mathbf{e}_A, \mathbf{e}_B] = \gamma^C{}_{AB} \mathbf{e}_C \quad , \tag{2.113}$$

such that $\gamma^C{}_{AB} = \gamma^C{}_{AB}(x^a)$. Hence the following must hold:

$$\gamma^A{}_{BC} = -\gamma^A{}_{CB} \tag{2.114}$$

$$\begin{aligned} \gamma^A{}_{BC} &= e^A_a (e^b_B \partial_b e^a_C - e^b_C \partial_b e^a_B) \\ &= -2 e^a_B e^b_C e^A_{[b;a]} \quad . \end{aligned} \tag{2.115}$$

Naturally, the tetrad basis vectors must satisfy the Jacobi identities; this yields the following:

$$\mathbf{e}_{[A}(\gamma^D{}_{BC]}) + \gamma^E{}_{[AB} \gamma^D{}_{C]E} = 0 \quad . \tag{2.116}$$

One can also now define the *spin-connection components* or *Ricci rotation co-efficients* $\Sigma^A{}_{BC}$:

$$\Sigma^A{}_{BC} \equiv e^A_a e^b_C e^a_{B\ ;b} \quad , \tag{2.117}$$

from which it follows that:

$$\Sigma_{ABC} = -\Sigma_{BAC} \quad , \tag{2.118}$$

that is, the spin connection is skew-symmetric on the first and second indices. This is a useful property as the skew-symmetry results in the spin-connection having fewer components than the affine connection; this naturally simplifies calculations in most contexts. One can now calculate the tetrad form of the Riemann tensor:

$$R^A{}_{BCD} = \mathbf{e}_C(\Sigma^A{}_{BD}) - \mathbf{e}_D(\Sigma^A{}_{BC}) + \Sigma^A{}_{EC}\Sigma^E{}_{BD} - \Sigma^A{}_{ED}\Sigma^E{}_{BC} - \Sigma^A{}_{BE}\gamma^E{}_{CD} \quad . \tag{2.119}$$

Hence the Ricci tensor transpires as:

$$R_{BD} = \mathbf{e}_A(\Sigma^A{}_{BD}) - \mathbf{e}_D(\Sigma^A{}_{BA}) + \Sigma^A{}_{EA}\Sigma^E{}_{BD} - \Sigma^A{}_{ED}\Sigma^E{}_{BA} \quad , \tag{2.120}$$

and the Ricci scalar:

$$R = \Pi^{BD} R_{BD} \quad . \tag{2.121}$$

Returning now to the definition of the tetrad directional derivative 2.112, one can rewrite it as follows:

$$\begin{aligned} e^a_A T_{a;b} e^b_B &= T_{A,B} - \Sigma^C_{AB} T_C \\ &\equiv T_{A|B} \quad . \end{aligned} \tag{2.122}$$

The quantity $T_{A|B}$ is referred to as the *intrinsic derivative*, and is readily extended to tensors of all types, by the above definition.

Equivalently, one can define both the affine and spin-connections simultaneously by postulating the following vanishing derivative:

$$\mathcal{D}_a e_{bC} \equiv e_{bC,a} - \Gamma^d{}_{bc} e_{dC} + \Sigma_C{}^E{}_a e_{bE} = 0 \quad . \tag{2.123}$$

Performing this operation on the tetrad metric then yields the above definitive expressions for both connections. The rotation co-efficients and commutation relations then transpire as:

$$\Sigma_{ABC} = \frac{1}{2}\left(\Pi_{AD}\gamma^D_{CB} - \Pi_{BD}\gamma^D_{CA} + \Pi_{CD}\gamma^D_{AB}\right) \tag{2.124}$$

$$\gamma^A_{BC} = \Sigma^A_{CB} - \Sigma^A_{BC} \tag{2.125}$$

2.6 Bibliography

- SW Hawking and GFR Ellis (1973) *The Large Scale Structure of Space-Time* (Cambridge)
- R D'Inverno (1992) *Introducing Einstein's Relativity* (Oxford)
- RM Wald (1984) *General Relativity* (Chicago)
- B Schutz (1980) *Geometrical Methods of Mathematical Physics* (Cambridge)
- H Rund and D Lovelock (1975) *Tensors, Differential Forms and Variational Principles* (Wiley)
- GFR Ellis UCT lecture series: *Gravitation* (UCT)
- Courant and Hilbert (1933)*Methods in Mathematical Physics* Vol. 1,2 (Interscience)
- S Weinberg (1972) *Gravitation and Cosmology* (Wiley)
- CW Misner, KS Thorne and JA Wheeler (1973) *Gravitation* (Freeman)
- PAM Dirac *General Relativity* (Cambridge)
- GFR Ellis and RM Williams (1988) *Flat and Curved Space-Times* (Oxford)

Chapter 3

The Calculus of Variations

"Physics originally began as a descriptive macrophysics, containing an enormous number of empirical laws with no apparent connections. In the beginning of science, scientists may have been very proud to have discovered hundreds of laws. But, as the laws proliferate, they become unhappy with this state of affairs; they begin to search for underlying principles."

Rudolph Carnap

3.1 Variations

In this chapter the Calculus of Variations will be developed, both in a physical context from the Classical Mechanical formulation using the Least Action Principle derived from the principles of D'Alembert and Hamilton, and within an independent mathematical framework. Mathematical variations in general are discussed and, with reference to the subsequent chapters on cosmological perturbation theory, variations of various pertinent mathematical quantities relevant to General Relativity are derived.

The aim here will be not only to provide the requisite variational tools utilised in General Relativity and Cosmology, but also to inculcate a working knowledge of the Variational Calculus which, as will be justified, plays a crucially prominent unifying role in modern Physics.

In order to inculcate a conceptual appreciation and understanding of the Variational Principle, numerous equivalent derivations of the resultant Euler-Lagrange equations will be investigated together with their associated advantages.

3.1.1 The Notion of a Variation

The most lucid approach to understanding the notion of a 'variation' is through the idea of an infinitesimal increment of some particular quantity. In a co-ordinate system, for example, this increment can manifest itself through an *infinitesimal transformation*; this will be the regime utilised in the ensuing theory, not only for its generality, but also in that it is commensurate with the differential geometry framework adopted throughout.

Proceeding thus, consider some scalar functional quantity $Q(x^i)$ of a co-ordinate system x^i; one can define the following infinitesimal (scalar) transformation:

$$Q(x^i) \;\rightarrow\; \overline{Q(x^i)} = Q(x^i) + \epsilon P(x^i) \quad , \tag{3.1}$$

where $P(x^i)$ is a purely arbitrary function, while ϵ is a constant parameter of smallness such that $\epsilon \ll 1$. This enables one to define the "variation" $\delta Q(x^i)$ of the quantity $Q(x^i)$ naturally as follows:

$$\delta Q(x^i) = \overline{Q(x^i)} - Q(x^i) \tag{3.2}$$
$$= \epsilon P(x^i) \quad . \tag{3.3}$$

From this it is evident that δ behaves as an operator, the *variational operator*. From this and the above definition it follows that the variational operator *commutes* with any differential operator d:

$$d\overline{Q} = dQ + \epsilon dP \tag{3.4}$$

$$\begin{aligned} \delta dQ &= d\overline{Q} - dQ \\ &= \epsilon dP \\ &= d(\epsilon P) \\ &= d(\delta Q) \quad . \end{aligned} \tag{3.5}$$

This is an important property of the variational operator which will be used extensively later. In the above context, only variations in scalar quantities have been considered; however, this is naturally generalised to vector and tensorial quantities. The variation of particular fundamental differential geometry quantities, both tensorial and otherwise, can now be formulated.

3.1.2 Variations of Various Functional Quantities

In addition to providing the variations to quantities in use later, the contents of this section are also intended, through demonstration, to provide a working understanding and intuitive appreciation of the Variational Calculus.

In this section the derivation of the metric tensor variation will be given in full, due to its central importance in later applications. However, the details of the calculations for the variations of the other quantities mentioned will not be required here; these derivations can, however, be found in full in the references given at the end of the chapter.

The metric tensor

Consider the metric variation:

$$g_{ab} \;\rightarrow\; g_{ab} + \delta g_{ab} \quad , \tag{3.6}$$

and hence

$$g^{ab} \rightarrow g^{ab} + \delta g^{ab} \quad . \tag{3.7}$$

From these it follows that:

$$\begin{aligned} \delta^a_c &= g^{ab} g_{bc} \rightarrow (g^{ab} + \delta g^{ab})(g_{bc} + \delta g_{bc}) \\ &= \delta^a_c + \delta g^{ab} g_{bc} + g^{ab} \delta g_{bc} + \mathcal{O}(\delta^2) \quad . \end{aligned} \tag{3.8}$$

Ignoring higher order terms in δ, and bearing in mind that the Krönecker delta is constant, thus having *no* variation, it follows that:

$$\delta g^{ab} g_{bc} + g^{ab} \delta g_{bc} = 0 \quad , \tag{3.9}$$

so that, by contracting this with g^{cd}, one obtains the useful identity:

$$\delta g^{ad} = -g^{ab} g^{cd} \delta g_{bc} \quad . \tag{3.10}$$

Similarly, for the metric tensor determinant g one can show:

$$\delta\sqrt{-g} = \frac{1}{2}\sqrt{-g} g^{ab} \delta g_{ab} \quad . \tag{3.11}$$

The Christoffel Symbols

For the Christoffel symbols, one has by definition:

$$\delta\Gamma^a_{bc} = \frac{1}{2} g^{ad} (\delta g_{db,c} - \delta g_{bc,d} + \delta g_{cd,b}) \quad . \tag{3.12}$$

The Ricci tensor

Considering a variation in the connections:

$$\Gamma^a{}_{bc} \rightarrow \overline{\Gamma}^a{}_{bc} = \Gamma^a{}_{bc} + \delta\Gamma^a{}_{bc} \quad , \tag{3.13}$$

one can, using geodesic co-ordinates (choosing a point P at which the connection vanishes identically) and the definition of the Riemann tensor in terms of the (metric) connection, show (see D'Inverno) that:

$$\delta R_{bd} = (\delta\Gamma^a{}_{bd})_{;a} - (\delta\Gamma^a{}_{ba})_{;d} \quad , \tag{3.14}$$

which is known as the *Palatini equation.* This will be used later in Part III in the formulation of the Hilbert-Palatini action for General Relativity. Hence, using the foregoing, one also has:

$$\delta R^a_b = g^{ac} \delta R_{bc} - \delta g^{ac} R_{bc} \quad . \tag{3.15}$$

Variation of tetrad-defined quantities

The following tetrad variations will be used in Part III in the tetrad formulated action for General Relativity. For the tetrad e_a^A, one has the variation:

$$\delta e_A^a = -e_A^b e^{aC} \delta e_{bC} \quad , \tag{3.16}$$

and for the tetrad determinant e:

$$\delta e = e e^{aA} \delta e_{aA} \quad . \tag{3.17}$$

While the following Ricci and spin-connection component variations are also useful:

$$\delta R_{ab}{}^{AB} = 2\mathcal{D}_{[a} \delta \Sigma^{AB}_{\;\;b]} \tag{3.18}$$

$$\delta \Sigma^{AB}_{\;\;a} = e^{b[A} \left(\frac{1}{2} \mathcal{D}_{[a} \delta e^{B]}_{b]} + e^{f B]} e_a^C \mathcal{D}_f \delta e_{bC} \right) \quad , \tag{3.19}$$

where the operator $\mathcal{D}_a$ is as defined in section 2.5.1.

3.2 The Variational Principle

The Variational Principle in physics has its origin in classical mechanics. Consequently, in this section D'Alemert's principle of vanishing virtual work combined with Hamilton's principle of least action will be used to formulate the classical Lagrangian and optimisation of the associated action integral. The mathematical theory of the calculus of variations will then be used to derive the Euler-Lagrange equations in their full generality from Hamilton's principle. The primary motivation for indulging in the standard classical theory of this section is historical; however, it is intended to provide a complete basis for the material which follows.

3.2.1 The Principles of D'Alembert and Hamilton

Given a system of N particles in classical three-space, there exists $3N$ *degrees of freedom* $\{\mathbf{x}_1, \dots \mathbf{x}_N\}$, the number of parameters or variables needed to describe the system completely, where $\mathbf{x}_i$ is the set of co-ordinates of the i^{th} particle (i is thus not a vector component index). A *force of constraint* $\mathbf{F}^{(c)}$ is then defined to be a force which restricts the number of degrees of freedom of such a given system. An example of such a force of constraint would be the normal force restricting a body to move on a given surface; if this surface were a plane, then only two co-ordinates would be required to describe the motion of the body, the third being *fixed* by the normal force. If one has P such forces of constraint, these can be formulated functionally as:

$$C_i(\mathbf{x}_1, \dots \mathbf{x}_N) = 0, \; i = 1, \dots P \quad , \tag{3.20}$$

then the resulting number of degrees of freedom for the system will be $3N - P$. These remaining degrees of freedom will be referred to as *generalised co-ordinates*, and denoted by q_i, $i = 1, \dots (3N - P)$, while the above constraint equations 3.20 will be referred to as the *kinematic relations*. One further defines a *virtual displacement* $\delta \mathbf{q}_i$ as one which does not violate the kinematic relations.

Returning now to the previous example of a body constrained by a normal force to move on a given plane, one notices that the normal force *does no work*. This is on account of the fact that any motion of the body in the plane will be orthogonal to the line of action of the normal force, whence zero work. In terms of the above definitions this amounts to:

$$\mathbf{F}^{(c)} \cdot \delta \mathbf{q}_i = 0 \quad , \tag{3.21}$$

which is defined to be *virtual* work. Hence the virtual work *vanishes*. This principle can be extended generally, and extrapolated as D'Alembert's principle: *The forces of constraint in a mechanical system do no work.*

Using now Newton's third law, and splitting up the forces acting on a mechanical system into constraint forces $\mathbf{F}_i^{(c)}$ and external forces $\mathbf{F}_i$ one has the following:

$$\sum_i m_i \ddot{\mathbf{q}}_i = \sum_i \left(\mathbf{F}_i^{(c)} + \mathbf{F}_i \right) \quad . \tag{3.22}$$

Considering then a virtual displacement $\delta \mathbf{q}_i$, and using D'Alembert's principle, one has from 3.22 the following:

$$\sum_i (m_i \ddot{\mathbf{q}} \cdot \delta \mathbf{q}_i - \mathbf{F}_i \cdot \delta \mathbf{q}_i) = 0 \quad . \tag{3.23}$$

Integrating this equation with respect to time t from t_1 to t_2, and noting that $\delta \dot{\mathbf{q}}_i = \frac{d}{dt} \delta \mathbf{q}_i$, one obtains:

$$0 = \sum_i \int_{t_1}^{t_2} (m_i \ddot{\mathbf{q}} \cdot \delta \mathbf{q}_i - \mathbf{F}_i \cdot \delta \mathbf{q}_i) \, dt \tag{3.24}$$

$$= \sum_i m_i [\dot{\mathbf{q}}_i \cdot \delta \mathbf{q}_i]_{t_1}^{t_2} - \int_{t_1}^{t_2} \left(\delta T + \sum_i \mathbf{F}_i \cdot \delta \mathbf{q}_i \right) dt \tag{3.25}$$

$$= \int_{t_1}^{t_2} \left(\delta T - \sum_1 \nabla_i V_i \cdot \delta \mathbf{q}_i \right) dt \tag{3.26}$$

$$= \int_{t_1}^{t_2} (\delta T - \delta V) \, dt \tag{3.27}$$

$$= \delta \int_{t_1}^{t_2} (T - V) \, dt \tag{3.28}$$

$$= \delta \int_{t_1}^{t_2} \mathcal{L} dt \quad , \tag{3.29}$$

where i) zero end point variation has been assumed: $\mathbf{q}_i(t_1) = \mathbf{q}_i(t_2) = 0$; ii) $T = \frac{1}{2} \sum_i m_i \mathbf{q}_i \cdot \mathbf{q}_i$ is defined to be the *kinetic energy*; iii) the external forces are assumed to be conservative: $\mathbf{F}_i = \nabla_i V_i$ for some potential function V_i, and $V = \sum_i V_i$ is defined to be the total *potential energy* of the system; iv) $\mathcal{L} = T - V$ is called the *Lagrangian* of the mechanical system, and $\int_{t_1}^{t_2} \mathcal{L} dt$ is called the associated *action* for the system. Equation

3.29 is known as *Hamilton's variational principle:* the Lagrangian for a mechanical system is such that the associated action obtains a stationary value. This is also generically referred to as the *least-action principle.* The physicality of this principle is discussed in the following section.

3.2.2 Physical Context of the Variational Principle

At this point there arises a fundamentally pertinent issue: the relevance to physics of the principle of least action. The issue rests on two schools of thought: whether or not Hamilton's principle is to be taken as a *physical* principle, or whether one should regard it merely as a mathematical tool for obtaining the existing laws of physics in a generically unified way.

Advocates for the former allude to the apparent ubiquity and prevalence of the principle throughout physics - virtually every branch of physics can be formulated via a variational approach. Indeed, one cannot imagine doing Quantum Field Theory, String Theory or Quantum Electrodynamics without the notion of a variational principle. However, a noticeable exception is the theory of thermodynamics: the variational principle cannot yield inequalities. In addition, the approach is further motivated by the argument that the variational principle yields results which conform to explicit conservation laws and symmetries; properties which follow directly from the mathematical formalism, as will be demonstrated later in this chapter.

Advocates for the latter argue that it is merely a mathematical *trick* inherently contrived to yield existing physical results within the particularly useful mathematical framework of the calculus of variations. This stance is motivated technically by the argument that D'Alembert's principle of vanishing virtual work is *mathematically imposed in hindsight* so as to yield the known correct results, and thus is physically inane. The argument is given additional credibility by the lack of a plausible motivation as to why nature should be constrained to behaving in a manner compatible with 'least action'.

However, whatever the stance, the self-evident broad, cross-spectrum applicability and universality of the principle itself makes it perhaps the most unifying and thus perhaps most relevant notion in all of physics, regardless of its debatable physicality.

3.2.3 The Calculus of Variations

Treated within the mathematical context of the calculus of variations, the variational principle, in its simplest form, arises from the optimisation of the expression:

$$\mathcal{I} = \int_{c_1}^{c_2} \mathcal{F}\left(\dot{x}(t), x(t)\right) dt \quad , \tag{3.30}$$

where $\dot{x}$ is understood to be the total derivative of the variable $x(t)$ with respect to the independent variable t; and $\mathcal{F}$ is some generic function of x and $\dot{x}$. Optimisation of 3.30 amounts to solving for that $x(t)$, referred to as a *trajectory*, which optimises the *action* $\mathcal{I}$. The action $\mathcal{I}$ is referred to mathematically as a *functional*, the terminology intended to imply that the operative variable is itself a function as opposed to an arbitrary parameter.

The pertinent optimisation usually transpires, but not always, as a *minimisation*, whence the epithet *least* action for the above process. Consequently, one requires:

$$\delta\mathcal{I} = \int_{c_1}^{c_2} \delta\mathcal{F} dt = 0 \quad , \tag{3.31}$$

where the symbol δ is used to denote a *small variation* in the above integral. The nature of δ is as elucidated in section 3.1.1. In formulating the above action $\mathcal{I}$, zero boundary variation will firstly be assumed; that is, $\delta x = 0$ when $t = c_1$ or $t = c_2$. Here, one is referring to the limits of the definite integral as the *boundary* of the action; when generalised, as will be done subsequently, to a multiple integral this boundary takes on the form of an $(n-1)$-surface, where n refers to the number of integrand parameters t_i and hence the order of the integral in 3.30:

$$\mathcal{I} = \int_{c_1^{(1)}}^{c_2^{(1)}} \dots \int_{c_1^{(n)}}^{c_2^{(n)}} \mathcal{F}\left(\dot{x}(t_i), x(t_i)\right) dt_1 \dots dt_n \quad , \tag{3.32}$$

where $i = 1, \dots n$. Hence one is integrating in 3.32 over a n-*volume* bounded by a $(n-1)$-*surface*. Performing the variation 3.31, integrating by parts, and applying the vanishing boundary assumption to the one surface term which thus arises, one obtains:

$$\int_{c_1}^{c_2} \left(\frac{\partial\mathcal{F}}{\partial x} - \frac{d}{dt}\left(\frac{\partial\mathcal{F}}{\partial\dot{x}}\right)\right) \delta x dt = 0 \quad . \tag{3.33}$$

By Bliss' lemma, this implies:

$$\frac{\delta\mathcal{F}}{\delta x} \equiv \frac{\partial\mathcal{F}}{\partial x} - \frac{d}{dt}\left(\frac{\partial\mathcal{F}}{\partial\dot{x}}\right) = 0 \quad . \tag{3.34}$$

Equation 3.34 is known as the *Euler-Lagrange equation*. The notation $\frac{\delta\mathcal{F}}{\delta x}$ defined in the above equation is referred to as either the *Euler-Lagrange derivative*, *functional derivative*, or *variational derivative* of $\mathcal{F}$. This notation is motivated by the following:

$$\begin{aligned} \delta\mathcal{I} &= \int \delta\mathcal{F} dx \\ &= \int \frac{\delta\mathcal{F}}{\delta x} \delta x dt \quad , \end{aligned} \tag{3.35}$$

which, by Bliss' Lemma, yields the Euler-Lagrange equations in shorthand notation:

$$\frac{\delta\mathcal{F}}{\delta x} = 0 \quad . \tag{3.36}$$

One can extend the above analysis to its most general form by firstly considering a trajectory x^i in n-space, where $i = 1, \dots n$; secondly by treating the x^i each as functions of m parameters t^α; and thirdly by generalising $\mathcal{F}$ to a function of all derivatives of the x^i with respect to the t^α to p^{th} order, where each derivative will be a partial with respect to t^α. This gives:

$$\mathcal{F} = \mathcal{F}\left(t^\alpha, x^i, x_{,\alpha_1}, x^i_{,\alpha_1\alpha_2}, \dots x^i_{,\alpha_1\dots\alpha_p}\right) \quad . \tag{3.37}$$

With this generality, the Euler-Lagrange equations take on the form:

$$
\frac{\partial \mathcal{F}}{\partial x^i} - \frac{d}{dt^{\alpha_1}}\left[\frac{\partial \mathcal{F}}{\partial x^i_{,\alpha_1}} - \frac{d}{dt^{\alpha_2}}\left(\frac{\partial \mathcal{F}}{\partial x^i_{,\alpha_1\alpha_2}}\right) + \ldots \right.
$$

$$
\left.(-1)^{p+1}\frac{d^{(p-1)}}{dt^{\alpha_2}\ldots dt^{\alpha_{p-1}}}\left(\frac{\partial \mathcal{F}}{\partial x^i_{,\alpha_1\ldots\alpha_{p-1}}}\right)\right] = 0 \quad . \tag{3.38}
$$

Note also that here one has an action which is a multiple integral comprising integration with respect to m separate parameters; i.e. an m-tuple integral. In all the cases considered above it has implicitly been assumed that the space in which the actions have been considered is Euclidean; however, as the formulation of actions specific to gravity will be required later, a suitable generalisation to a non-flat manifold is desired. This is done trivially via the definition of integration over a manifold as in section 2.4.4, realising that, as mentioned, 2.95 results in the integration being performed in R^n. In physical contexts the functional form $\mathcal{F}$ is referred to as the *Lagrangian*, and will thus henceforth be denoted by convention as $\mathcal{L}$.

An alternative formulation of the Euler-Lagrange equations

For the sake of generalisation, the following convenient formulation of the variational problem will be utilised. In addition to being rather an elegant formulation, it is particularly practical in calculations and exhibits a notational simplicity.

Consider the following simplistic action:

$$
\mathcal{I}[\phi] = \int_{t_0}^{t_1} \mathcal{L}(\phi(t), \dot{\phi}, t)dt \quad . \tag{3.39}
$$

Recalling then the notion of optimising this action, let $u(t)$ be that trajectory which optimises 3.39; i.e. $\delta\mathcal{I}[u] = 0$. Maintaining then the prior fixed boundary assumption, one can express a general admissible trajectory $\phi(t)$ (not optimal) as follows:

$$
\phi(t) = u(t) + \epsilon\nu(t) \quad , \tag{3.40}
$$

such that $\nu(t_0) = \nu(t_1) = 0$, and ϵ is some parameter of smallness. Inserting 3.40 into 3.39, one transforms $\mathcal{I}$ strictly into a function of ϵ: $\mathcal{I} = \mathcal{I}[\epsilon]$. Consequently, one can now perform the following Taylor expansion of the functional:

$$
\begin{aligned}
\mathcal{I}[\phi(t)] &= \mathcal{I}[u(t)] + \epsilon\left(\frac{d\mathcal{I}}{d\epsilon}\right)_{\epsilon=0} + \frac{\epsilon^2}{2}\left(\frac{d^2\mathcal{I}}{d\epsilon^2}\right)_{\epsilon=0} + \ldots\frac{\epsilon^n}{n!}\left(\frac{d^n\mathcal{I}}{d\epsilon^n}\right)_{\epsilon=0} + \ldots \\
&\equiv \mathcal{I}[u(t)] + \epsilon\delta\mathcal{I} + \frac{\epsilon^2}{2}\delta^2\mathcal{I} + \ldots\frac{\epsilon^n}{n!}\delta^n\mathcal{I} + \ldots \quad ,
\end{aligned} \tag{3.41}
$$

where $\delta^n\mathcal{I}$ is called the n^{th} *variation of the functional* $\mathcal{I}$. The second variation will not be considered later; it is noted here however, and can indeed easily be proven, that if the second variation is greater than zero, the functional obtains a minimum as one would expect; this is indeed the case for classical mechanics, whence the term *Least* action principle.

3.2.4 The Hamiltonian Approach

In this section the problem of optimising an action will be tackled by using an approach due to Caratheodery. This will result in a set of first order equations, known as Hamilton's canonical equations of motion, being derived as an equivalent set of equations to the Euler-Lagrange Equations. In the process, the Hamilton-Jacobi relation will also be derived.

The Hamiltonian

Given a Lagrangian $\mathcal{L} = \mathcal{L}(t, x^a, \dot{x}^a)$ one can define a quantity called the *canonical momentum* as follows:

$$p_a = \frac{\partial \mathcal{L}}{\partial \dot{x}^a} \tag{3.42}$$

It can easily be shown from this definition that the canonical momentum p_a is a covariant vector with a one-to-one correspondence to the contravariant vector $\dot{x}^a$. Now, provided that

$$det\left(\frac{\partial^2 \mathcal{L}}{\partial \dot{x}^i \partial \dot{x}^j}\right) \neq 0 \quad , \tag{3.43}$$

one can solve equation 3.42 explicitly for $\dot{x}^i = \dot{x}^i(t, x^a, p_a)$. This can then be used to perform a Legendre transformation on $\mathcal{L}$ to obtain the *Hamiltonian* $\mathcal{H}$ as:

$$\mathcal{H}(t, x^b, p_b) := -\mathcal{L}[t, x^b, \dot{x}^b(t, x^c, p_c)] + p_a \dot{x}^a(t, x^c, p_c) \quad , \tag{3.44}$$

where the brackets denote the relevant functionality. Hence the Legendre transformation replaces $\dot{x}^a$ with the canonical momentum p_a. From this definition of the Hamiltonian, it is quite simple to verify the following identities:

$$\frac{\partial \mathcal{H}}{\partial p_a} = \dot{x}^a \tag{3.45}$$

$$\frac{\partial \mathcal{H}}{\partial x^a} = -\frac{\partial \mathcal{L}}{\partial x^a} \tag{3.46}$$

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t} \tag{3.47}$$

The canonical equations

From the above, one has that the Euler-Lagrange derivative can be transformed into:

$$\frac{dp_a}{dt} + \frac{\partial \mathcal{H}}{\partial x^a} = 0 \quad . \tag{3.48}$$

Hence, equivalent to the n second-order Euler-Lagrange equations, one has the following $2n$ *first*-order Hamilton canonical equations:

$$
\begin{aligned}
\dot{x}^a &= \frac{\partial \mathcal{H}}{\partial p_a} && (3.49)\\
\frac{dp_a}{dt} &= -\frac{\partial \mathcal{H}}{\partial x^a} \quad , && (3.50)
\end{aligned}
$$

where the latter is obtained directly from imposing the Euler-Lagrange equations. This formulation will be used later apropos of Noether's theorem which yields conserved quantities; the desire for such quantities in physics is self-evident.

The Hamilton-Jacobi relation

Suppose now that there exists a class C^2 function $S(t, x^a)$ defined along the curve C : $x^a = x^a(t)$. This quantity can then be used to formulate an alternative Lagrangian:

$$
\begin{aligned}
\mathcal{L}^\dagger &= \mathcal{L} - \frac{dS}{dt}\\
&= \mathcal{L} - \frac{\partial S}{\partial t} - \frac{\partial S}{\partial x^a}\dot{x}^a \quad . && (3.51)
\end{aligned}
$$

It then follows that the optimisation of the original action:

$$\mathcal{A}(C) = \int_{P_1}^{P_2} \mathcal{L}dt \tag{3.52}$$

amounts to the optimisation of the new action:

$$\mathcal{A}^\dagger(C) = \int_{P_1}^{P_2} \mathcal{L}^\dagger dt \quad . \tag{3.53}$$

This is due to the result:

$$\mathcal{A}^\dagger(C) - \mathcal{A}(C) = S_1 - S_2 \quad . \tag{3.54}$$

It can then be shown that the optimisation of the action $\mathcal{A}^\dagger$ amounts to determining $S(t, x^a)$ such that S satisfies the following two equations:

$$
\begin{aligned}
\frac{\partial S}{\partial x^a} &= p_a && (3.55)\\
\frac{\partial S}{\partial t} + \mathcal{H}\left(t, x^a, \frac{\partial S}{\partial x^a}\right) &= 0 \quad . && (3.56)
\end{aligned}
$$

The latter equation is known as the *Hamilton-Jacobi relation.*

Extension to higher dimensions

Considering an m-parameter action of the form:

$$I = \int_{c_1^{(1)}}^{c_2^{(1)}} \dots \int_{c_1^{(m)}}^{c_2^{(m)}} \dots \mathcal{L}(t^\alpha, x^a, \dot{x}^a_\alpha) dt_1 \dots dt_m \quad , \tag{3.57}$$

such that $\dot{x}^a_\alpha \equiv \frac{\partial x^a}{\partial t^\alpha}$ and $\alpha = 1 \ldots m$, one can define analogous to the above Hamiltonian the *Hamiltonian complex* $\mathcal{H}^\alpha_\beta$ as follows:

$$\mathcal{H}^\alpha_\beta = -\mathcal{L}\delta^\alpha_\beta + \frac{\partial \mathcal{L}}{\partial \dot{x}^a_\alpha}\dot{x}^a_\beta \tag{3.58}$$

such that the quantity $p_a = \frac{\partial \mathcal{L}}{\partial \dot{x}^a_\alpha}$ is easily verified to be the components of a covariant vector (the α subscript being a parameter index, not a vector component), the *canonical momentum* as before. The equivalent of Hamilton's equations then transpire as:

$$\frac{\partial \mathcal{H}^\alpha_\beta}{\partial p_a} = \dot{x}^a_\alpha \tag{3.59}$$

$$\frac{d\dot{x}^a_\alpha}{dt^\alpha} = -\frac{\partial \mathcal{H}^\alpha_\beta}{\partial x^a} \quad . \tag{3.60}$$

An associated geodesic field may then, analogously, be defined through the set of functions $S^\alpha(t^\beta, x^a)$, from which one may define the quantities:

$$c^\alpha_\beta = \frac{dS^\alpha}{dt^\beta} \tag{3.61}$$

$$\Delta = det(c^\alpha_\beta) \tag{3.62}$$

$$C^\alpha_\beta c^\beta_\gamma = \delta^\alpha_\gamma \Delta \quad . \tag{3.63}$$

Bearing these in mind, one can now write the Hamilton-Jacobi relation as:

$$\mathcal{H}^\alpha_\beta + C^\alpha_\delta \frac{\partial S^\delta}{\partial t^\beta} = 0 \quad . \tag{3.64}$$

3.3 Lagrangian Functional Forms and Invariance

In the preceding Lagrangian formulations, simplified functional dependence was assumed for the sake of clarity in elucidating the most natural generalisations of the Euler-Lagrange equations. In this section more complex, and ultimately physically relevant, functional forms will be investigated along with the consequences of invariance in the Variational Principle.

3.3.1 General Concepts

The general action integral concerned here will take on the form:

$$I[\mathbf{X}] = \int_{\mathcal{U}} \mathcal{L}(x^a, \mathbf{X}, \mathbf{X}_{,a}, \mathbf{X}_{,ab}) d^n x \quad , \tag{3.65}$$

where x^a denotes some arbitrary local co-ordinates, $\mathbf{X}$ is a general tensor, and the comma subscript denotes ordinary partial differentiation with respect to x^a as before. The invariance property considered here will be that of co-ordinate invariance; that is, *the action 3.65 will be assumed to be invariant under co-ordinate transformations of the form:*

$$x'^a = x'^a(x^b) \quad ; \tag{3.66}$$

that is,

$$I[\mathbf{X}(x^a)] = I[\mathbf{X}(x'^a)] \tag{3.67}$$

which is easily shown to imply that:

$$\begin{aligned} \mathcal{L}' &= \det\left(\frac{\partial x^a}{\partial x'^b}\right)\mathcal{L} \\ &= J\mathcal{L} \quad , \end{aligned} \tag{3.68}$$

where J is the Jacobian of the co-ordinate transformation 3.66. This then implies that the Lagrangian must be a scalar density. In the following sections two specific cases of the above will be considered: i) the case where $\mathbf{X}$ is a covariant vector field; ii) the case where $\mathbf{X}$ is a type $(0,2)$ symmetric tensor field. The aim will be to formulate the invariance 3.68 in terms of constraint equations to be satisfied by the Lagrangian. The Variational Principle will then take on the form of the Euler-Lagrange equations *along with these additional constraint equations.*

3.3.2 Vector Field Theory

Here, for the sake of relevance to subsequent applications in General Relativity and Cosmology, the Lagrangian form that will be studied will take on the following functional dependence:

$$\mathcal{L} = \mathcal{L}(x^a, X_a, X_{a,b}, S_{ab}) \quad , \tag{3.69}$$

where X_a are the components of a covariant vector field, and S_{ab} are the components of a symmetric type $(0,2)$ tensor field for which $s = |det(S_{ab})| \neq 0$. The S_{ab} here will be taken to be arbitrary, whereas the X_a are to be determined by the variational principle.

By considering a simple transformation $x'^a = x^a + c^a$ such that the c^a are constant, it can by shown that the above Lagrangian is explicitly *independent of* x^a. Using then the tensorial nature (tensor density of zero order) this result must then hold generically:

$$\mathcal{L} = \mathcal{L}(X_a, X_{a,b}, S_{ab}) \quad . \tag{3.70}$$

For the consideration of the transformation properties of the associated derivatives of $\mathcal{L}$ in 3.70 under 3.66, the following notation will be used for convenience:

$$\begin{aligned} \mathcal{L}^a &= \frac{\partial \mathcal{L}}{\partial X_a} \\ \mathcal{L}^{ab} &= \frac{\partial \mathcal{L}}{\partial X_{a,b}} \\ \mathcal{L}^{ab}_{(s)} &= \frac{\partial \mathcal{L}}{\partial S_{ab}} \quad , \end{aligned} \tag{3.71}$$

where the subscript (s) is a label corresponding to S_{ab}, and *not* a tensor index. The arguments of 3.70 themselves transform under 3.66 as follows:

$$
\begin{aligned}
X'_a &= J^b_a X_b \\
X'_{a,b} &= J^c_{ab} X_c + J^c_a J^d_b X_{c,d} \\
S'_{ab} &= J^c_a J^d_b S_{cd} \quad ,
\end{aligned}
\tag{3.72}
$$

such that, as in the preliminary chapters, the notation $J^a_b = \frac{\partial x^a}{\partial x'^b}$ and $J^a_{bc} = \frac{\partial^2 x^a}{\partial x'^b \partial x'^c}$ has been adopted. Consequently, 3.68 reduces to:

$$
\mathcal{L}'(J^b_a X_b, J^c_{ab} X_c + J^c_a J^d_b X_{c,d}, J^c_a J^d_b S_{cd}) = J\mathcal{L}(X_a, X_{a,b}, S_{ab}) \quad . \tag{3.73}
$$

Differentiating 3.73 with respect to $X_a, X_{a,b}$ and S_{ab} respectively, the following obtains:

$$
\mathcal{L}'^{ab} J^c_{ab} + \mathcal{L}'^a J^c_a = J\mathcal{L} \tag{3.74}
$$

$$
\mathcal{L}'^{ab} J^c_a J^d_b = J\mathcal{L}^{cd} \tag{3.75}
$$

$$
\mathcal{L}'^{ab}_{(s)} J^c_a J^d_b = J\mathcal{L}^{cd}_{(s)} \quad , \tag{3.76}
$$

which clearly show that the quantities $\mathcal{L}^{cd}, \mathcal{L}^{cd}_{(s)}$ are type $(2,0)$ tensor densities. Differentiating 3.73 now with respect to J^a_{bc} yields:

$$
\mathcal{L}'^{ab} X_c (\delta^d_a \delta^e_b - \delta^e_b \delta^d_a) = 0 \quad , \tag{3.77}
$$

which, by virtue of the tensorial nature of $\mathcal{L}^{ab}$, implies that:

$$
\mathcal{L}^{ab} + \mathcal{L}^{ba} = 0 \quad , \tag{3.78}
$$

which is the first constraint $\mathcal{L}$ must obey by virtue of 3.73. A notable consequence of this constraint is that one cannot consider gradient fields in $\mathcal{L}$; if gradient fields were to be considered, i.e fields X_a for which $X_{a,b} = X_{b,a}$ then this along with 3.78 would imply $X_{a,b} = 0$ resulting in $\mathcal{L} = \mathcal{L}(X_a, S_{ab})$.

Applying 3.78 to 3.74, the first term on the left vanishes yielding the definitive transformation for a type $(1,0)$ tensor density. Hence one has the second constraint, that $\mathcal{L}_{,a}$ is a type $(1,0)$ tensor density:

$$
\mathcal{L}'_a = J J^b_a \mathcal{L}_b \quad . \tag{3.79}
$$

Continuing by differentiating 3.73 with respect to J^a_b, and using the identity $\frac{\partial J}{\partial J^a_b} = J\overline{J}^b_a$, one obtains:

$$
\mathcal{L}'^a X_b + J^c_d \mathcal{L}'^{ad} F_{cb} + 2\mathcal{L}'^{ad}_{(s)} S_{cb} J^c_d = J\mathcal{L}\overline{J}^a_b \quad , \tag{3.80}
$$

where the notation:

$$
F_{ab} = X_{a,b} - X_{b,a} \tag{3.81}
$$

has been used for later convenience and analogy, as will be seen; and $\overline{J}^a_b = \frac{\partial \overline{x}^a}{\partial x^b}$. Also note the useful identity:

$$F_{ab;c} + F_{ca;b} + F_{bc;a} = 0 \quad . \tag{3.82}$$

Using now the identity transformation $\overline{x}^a = x^a$, 3.80 simplifies to:

$$\mathcal{L}^a X_b + \mathcal{L}^{ac} F_{cb} + 2\mathcal{L}^{ac}_{(s)} S_{cb} = \delta^a_b \mathcal{L} \quad , \tag{3.83}$$

which can be rewritten as:

$$\mathcal{L}^{ab}_{(s)} = \frac{1}{2}\left[S^{ab}\mathcal{L} - \mathcal{L}^a S^{bc} X_c - S^{bc}\mathcal{L}^{ad} F_{dc}\right] \quad , \tag{3.84}$$

which is the third constraint a Lagrangian of the form 3.70 must obey *if the action is to be invariant under co-ordinate transformations.* Because of this assumption, 3.78, 3.79 and 3.84 are referred to as the *invariance identities.* The constraint 3.84 can also written as a divergence equation using the Euler-Lagrange equations; first though, a few results are required from the latter. The Euler-Lagrange equations for 3.70, using the above notation, reduce to the following:

$$\begin{aligned} \mathcal{E}^a(\mathcal{L}) &= \mathcal{L}^{ab}_{,b} - \mathcal{L}^a \\ &= \mathcal{L}^{ab}_{;b} - \mathcal{L}^a \\ &= 0 \quad , \end{aligned} \tag{3.85}$$

where the notation:

$$\mathcal{E}^a = \frac{\partial}{\partial x^b}\frac{\partial}{\partial X_{a,b}} - \frac{\partial}{\partial X_a} \tag{3.86}$$

has been used to indicate the *Euler-Lagrange derivative.* Note also that the second line in 3.85 results from 3.78, 3.75 and 2.62 with respect to the Christoffel symbols defined in terms of S_{ab}. Hence *the Euler-Lagrange equations become type* (1,0) *tensor density conditions.* Secondly, from 3.86, 2.62 and 2.63 it follows that:

$$\begin{aligned} \mathcal{E}^a(\mathcal{L})_{;a} &= \mathcal{E}^a(\mathcal{L})_{,a} \\ &= -\mathcal{L}^a_{,a} \\ &= 0 \quad , \end{aligned} \tag{3.87}$$

which is referred to as the *generalised Lorentz condition,* a name which will be justified shortly. Returning now to constraint 3.84, differentiating it covariantly with respect to S_{ab}, and using 3.81, 3.82, 3.85 and 3.87, the following constraint obtains:

$$\begin{aligned} \mathcal{L}^{ab}_{(s);a} &= \frac{1}{2}\left[S^{ba} X_a \mathcal{E}^c(\mathcal{L})_{;c} + S^{ba}\mathcal{F}_{ca}\mathcal{E}^c(\mathcal{L})\right] \\ &= 0 \quad . \end{aligned} \tag{3.88}$$

The quantity 3.84 is usually referred to as the *energy-momentum tensor density.* This terminology is justified through consideration of the following example. Consider:

$$\mathcal{L} = \sqrt{s} S^{ab} \left(\frac{1}{2} S^{cd} F_{ac} F_{bd} + \alpha X_a X_b \right) \quad , \tag{3.89}$$

where α is a constant. This is clearly seen to satisfy the form 3.68, hence the above analysis applies. Proceeding thus, one obtains firstly the Euler-Lagrange equations:

$$\begin{aligned} \mathcal{E}^a(\mathcal{L}) &= 2 \left(\sqrt{s} F^{ab} \right)_{;b} - 2\alpha \sqrt{s} S^{ab} X_b \\ &= 0 \quad , \end{aligned} \tag{3.90}$$

and the energy-momentum tensor:

$$\mathcal{L}^{ab}_{(s)} = \sqrt{s} \left(F^{ac} F_{dc} S^{db} - \tfrac{1}{4} S^{ab} F_{cd} F^{cd} \right) - \alpha \sqrt{s} \left(S^{ac} S^{bd} X_c X_d - \tfrac{1}{2} S^{ab} S^{cd} X_c X_d \right) \quad ; \tag{3.91}$$

while the generalised Lorentz condition 3.87 becomes:

$$\frac{\partial}{\partial x^a} \left(\sqrt{s} S^{ab} X_b \right) = 0 \quad . \tag{3.92}$$

If one now imposes S^{ab} to be the Minkowskian metric, and X_a to be the electromagnetic 4-vector potential, then 3.82 and 3.90 reduce to the familiar Maxwell equations, while 3.92 becomes the Lorentz condition, whence its name.

3.3.3 Metric Field Theory

In this section a Lagrangian of the following form will be considered:

$$\begin{aligned} \mathcal{L} &= \mathcal{L}(x^a, g_{ab}, g_{ab,c}, g_{ab,cd}) \\ &= \mathcal{L}(g_{ab}, g_{ab,c}, g_{ab,cd}) \quad , \end{aligned} \tag{3.93}$$

where g_{ab} is a symmetric, invertible $(0,2)$ tensor thus generically representing a metric, whence the notational convention of using g_{ab}. As before, one requires that this Lagrangian be a scalar density:

$$\mathcal{L}' = J\mathcal{L} \quad . \tag{3.94}$$

As for the vector field approach, one now requires first to establish the tensorial nature of the Euler-Lagrangian equations associated with 3.93; and second, to formulate the constraint equations on 3.93 subject to the condition 3.94. Proceeding thus, one defines, as before, the following quantities:

$$\mathcal{L}^{ab}_{(g)} = \frac{\partial \mathcal{L}}{\partial g_{ab}} \tag{3.95}$$

$$\mathcal{L}^{abc}_{(g)} = \frac{\partial \mathcal{L}}{\partial g_{ab,c}} \tag{3.96}$$

$$\mathcal{L}^{abcd}_{(g)} = \frac{\partial \mathcal{L}}{\partial g_{ab,cd}} \quad , \tag{3.97}$$

noting the associated symmetries subject to the metric and its partial derivatives. In terms of these quantities the Euler-Lagrange equations associated with 3.93 become:

$$\begin{aligned} \mathcal{E}^{ab}(\mathcal{L}) &= \left[\mathcal{L}^{abc}_{(g)} - \mathcal{L}^{abcd}_{(g),d}\right]_{,c} - \mathcal{L}^{ab}_{(g)} \\ &= 0 \quad . \end{aligned} \tag{3.98}$$

The transformation properties of the quantities 3.95, 3.96 and 3.97 are derived via differentiation of 3.94 with respect to $g_{ab}, g_{ab,c}$ and $g_{ab,cd}$ as before, yielding respectively:

$$\begin{aligned} J\mathcal{L}^{abcd}_{(g)} &= \mathcal{L}'^{efgh}_{(g)} \frac{\partial g'_{ef,gh}}{\partial g_{ab,cd}} \\ &= \mathcal{L}'^{efgh}_{(g)} J^a_e J^b_f J^c_g J^d_h \end{aligned} \tag{3.99}$$

$$J\mathcal{L}^{abc}_{(g)} = \mathcal{L}'^{efgh}_{(g)} \frac{\partial g'_{ef,gh}}{\partial g_{ab,c}} + \mathcal{L}'^{efg}_{(g)} \frac{\partial g'_{ef,g}}{\partial g_{ab,c}} \tag{3.100}$$

$$J\mathcal{L}^{ab}_{(g)} = \mathcal{L}'^{cdef}_{(g)} \frac{\partial g'_{cd,ef}}{\partial g_{ab}} + \mathcal{L}'^{cde}_{(g)} \frac{\partial g'_{cd,e}}{\partial g_{ab}} + \mathcal{L}'^{cd}_{(g)} \frac{\partial g'_{cd}}{\partial g_{ab}} \quad ; \tag{3.101}$$

from which one can see that $\mathcal{L}^{abcd}_{(g)}$ is a tensor density, the other quantities not. For the sake of subsequent calculations one needs to introduce tensorial quantities closely related to the non-tensorial quantities $\mathcal{L}^{ab}_{(g)}, \mathcal{L}^{abc}_{(g)}$. This is done firstly by introducing a symmetric $(0,2)$ tensor S_{ab} which, along with its derivatives, transforms exactly as g_{ab} and its derivatives. This having been established, one can define a further quantity $\mathcal{F}$ as follows:

$$\mathcal{F} = \mathcal{L}^{abcd}_{(g)} S_{ab,cd} + \mathcal{L}^{abc}_{(g)} S_{ab,c} + \mathcal{L}^{ab}_{(g)} S_{ab} \quad , \tag{3.102}$$

which is trivially shown to transform as a tensor density:

$$\mathcal{F}' = J\mathcal{F} \quad . \tag{3.103}$$

Furthermore one would like to write this quantity purely in terms of other tensorial quantities, thereby enabling one ultimately to formulate tensorial quantities in terms of $\mathcal{L}^{ab}_{(g)}, \mathcal{L}^{abc}_{(g)}$. One therefore seeks tensors T_{ab} and T^{abc} such that:

$$\mathcal{F} = \mathcal{L}^{abcd}_{(g)} g_{ab;cd} + T^{abc} g_{ab;c} + T^{ab} g_{ab} \quad , \tag{3.104}$$

such that covariant differentiation is with respect to the metric connection via g_{ab}. This latter formulation will thus be fully tensorial apropos of its arguments. Expanding then 3.104 fully in terms of the Christoffel connections and equating with 3.102, one obtains the desired forms:

$$T^{abc} = \mathcal{L}^{abc}_{(g)} + 2\Gamma^a_{ef}\mathcal{L}^{ebcf}_{(g)} + 2\Gamma^b_{ef}\mathcal{L}^{eacf}_{(g)} + \Gamma^c_{ef}\mathcal{L}^{abef}_{(g)} \tag{3.105}$$

$$\begin{aligned} T^{ab} &= \mathcal{L}^{ab}_{(g)} + \Gamma^a_{cd,e}\mathcal{L}^{cbde} + \Gamma^b_{cd,e}\mathcal{L}^{cade} - \Gamma^c_{fe}\Gamma^a_{cd}\mathcal{L}^{fbde}_{(g)} - \Gamma^e_{cf}\Gamma^b_{ed}\mathcal{L}^{cadf}_{(g)} - \Gamma^a_{cd}\Gamma^b_{ef}\mathcal{L}^{cefd}_{(g)} \\ &= -\Gamma^b_{cd}\Gamma^a_{ef}\mathcal{L}^{cefd}_{(g)} - \Gamma^c_{ef}\Gamma^a_{dc}\mathcal{L}^{dbef}_{(g)} - \Gamma^f_{de}\Gamma^b_{cf}\mathcal{L}^{cade}_{(g)} + \Gamma^a_{cd}T^{cbd} + \Gamma^b_{cd}T^{acd} \quad . \end{aligned} \tag{3.106}$$

The tensorial nature (tensor densities) of the above two expressions can (tediously) be verified using the transformation properties of the Christoffel symbols, $\mathcal{L}^{ab}_{(g)}$, $\mathcal{L}^{abc}_{(g)}$ and $\mathcal{L}^{abcd}_{(g)}$. Consequently, one can rewrite 3.102 and 3.104 respectively in terms of T^{ab}, T^{abc} and $\mathcal{E}^{ab}(\mathcal{L})$ as:

$$\mathcal{F} = S_{ab}\mathcal{E}^{ab}(\mathcal{L}) + \left[S_{ab}\mathcal{L}^{abc}_{(g)} + S_{ab,e}\mathcal{L}^{abce}_{(g)} - S_{ab}\mathcal{L}^{abce}_{(g),e}\right]_{,c} \tag{3.107}$$

$$\mathcal{F} = -S_{ab}\left[-\mathcal{L}^{abcd}_{(g);cd} + T^{abc}_{;c} - T^{ab}\right] + \left[S_{ab}T^{abc} + S_{ab;e}\mathcal{L}^{abce}_{(g)} - S_{ab}\mathcal{L}^{abcd}_{(g);d}\right]_{;c} \quad . \tag{3.108}$$

Hence, equating the terms in square brackets in the above two equations, one obtains the following expression for the Euler-Lagrange equations 3.98:

$$\mathcal{E}^{ab}(\mathcal{L}) = T^{abc}_{;c} - T^{ab} - \mathcal{L}^{abcd}_{(g);cd} \quad , \tag{3.109}$$

which is thus fully tensorial; thus the first objective is achieved: establishing the tensorial nature of 3.98. Turning now to the principal objective, namely the formulation of the constraint equations, one proceeds as for the vector field theory case, by differentiating 3.94. Hence, differentiating first with respect to J^a_{bcd}, and using the identity transformation $x'^a = x^a$ as before, one obtains:

$$\mathcal{L}^{abcd}_{(g)} + \mathcal{L}^{cbda}_{(g)} + \mathcal{L}^{dbac}_{(g)} = 0 \quad , \tag{3.110}$$

which is the first constraint. Differentiating 3.94 with respect to J^a_{bc}, and again using the identity transformation, one obtains:

$$2\mathcal{L}^{abcd}_{(g)}g_{eb,d} + 2\mathcal{L}^{cbad}_{(g)}g_{eb,d} + \mathcal{L}^{gfac}_{(g)}g_{gf,e} + \mathcal{L}^{abc}_{(g)}g_{eb} + \mathcal{L}^{cba}_{(g)}g_{eb} = 0 \quad . \tag{3.111}$$

Choosing now a normal co-ordinate system in which, by definition, $\Gamma^a_{bc} = 0 = g_{ab,c}$ at the pole P, and using 3.105 equation 3.111 simplifies to:

$$T^{abc} + T^{cba} = 0 \quad , \tag{3.112}$$

which, along with the other symmetries of T^{abc}, imply:

$$T^{abc} = 0 \tag{3.113}$$

at the pole P, which thus must be true everywhere by the tensorial nature of T^{abc}. This is the second constraint on 3.93. Differentiating 3.94 now with respect to J^a_b, one obtains:

$$2\mathcal{L}^{abcd}_{(g)}g_{eb,cd} + 2\mathcal{L}^{bdac}_{(g)}g_{bd,ec} + 2\mathcal{L}^{adb}_{(g)}g_{ed,b} + \mathcal{L}^{bda}_{(g)}g_{bd,e} + 2\mathcal{L}^{ab}_{(g)}g_{eb} = \delta^a_e\mathcal{L} \quad . \tag{3.114}$$

Using once again a normal co-ordinate system, the definition of the Riemann curvature tensor R_{abcd} and equations 3.106 and 3.113, one can eventually reduce 3.114 to:

$$T^{ab} = \frac{1}{2}g^{ab}\mathcal{L} + \frac{2}{3}\mathcal{L}^{cdbe}_{(g)}R^a_{dce} \quad . \tag{3.115}$$

Using then constraints 3.113 and 3.115, one can rewrite the Euler-Lagrange equations as:

$$\begin{aligned}\mathcal{E}^{ab}(\mathcal{L}) &= -\frac{1}{2}g^{ab}\mathcal{L} - \frac{2}{3}\mathcal{L}_{(g)}^{cdae}R^{b}_{dce} - \mathcal{L}^{abef}_{;ef} \\ &= 0 \quad .\end{aligned} \tag{3.116}$$

One can thus calculate the divergence of the Euler-Lagrange equations to be:

$$\mathcal{E}^{ab}(\mathcal{L})_{;b} = -\mathcal{L}^{abcd}_{(g),cdb} + \mathcal{L}^{abc}_{(g),cb} - \mathcal{L}^{ab}_{(g),b} = 0 \quad , \tag{3.117}$$

where the above expression is shown to vanish using equations 3.105, 3.110, 3.113 and 3.115. Consequently, the above three constraints 3.110, 3.113 and 3.115 manifest themselves in the form of the Euler-Lagrange equations 3.116 with the associated *single* constraint 3.117.

Note that, using the above theory one can also derive further restrictions on admissible Lagrangian densities. The first follows from 3.113: suppose that one had the Lagrangian density form $\mathcal{L} = \mathcal{L}(g_{ab}, g_{ab,c})$; by 3.105 this would imply $T^{abc} = \mathcal{L}^{abc}_{(g)}$, which would in turn imply, by 3.113, that $\frac{\partial \mathcal{L}}{\partial g_{ab,c}} = 0$ which contradicts the assumed functionality of $\mathcal{L}$. Hence one has the important result stating that *there does not exist a Lagrangian density of the functional form* $\mathcal{L} = \mathcal{L}(g_{ab}, g_{ab,c})$.

the second restriction concerns the form of 3.116. Here one wishes to find the most general form of $\mathcal{E}^{ab}$. Omitting the somewhat involved derivation, it can be shown (Rund, Lovelock) that in a four-dimensional space the only tensor density $\mathcal{A}^{ab}$ which satisfies:

$$\mathcal{A}^{ab} = \mathcal{A}^{ba} \tag{3.118}$$

$$\mathcal{A}^{ab}_{;b} = 0 \tag{3.119}$$

$$\mathcal{A}^{ab} = \mathcal{A}^{ab}(g_{cd}, g_{cd,e}, g_{cd,ef}) \tag{3.120}$$

is one with the form:

$$\mathcal{A}^{ab} = \alpha\sqrt{g}\left(R^{ab} - \frac{1}{2}g^{ab}R\right) + \beta\sqrt{g}g^{ab} \quad , \tag{3.121}$$

such that α, β are arbitrary constants, and R, R^{ab} are the Ricci scalar and tensor respectively.

3.3.4 Combined Vector and Metric Field Theory

One way of synthesising the preceding two cases of vector and metric field theory is by formulating a Lagrangian which is the sum of two separate Lagrangians, each of which falls into one of the above cases. This particular form will be adopted as it will be seen to facilitate the formulation of General Relativity which will require a Lagrangian which is conveniently the sum of a geometric Lagrangian and a matter Lagrangian. One therefore postulates the following generalised vector-metric Lagrangian:

$$\mathcal{L} = \mathcal{L}_{(V)}(X_a, X_{a,b}, g_{ab}) + \mathcal{L}_{(M)}(g_{ab}, g_{ab,c}, g_{ab,cd}) \quad , \tag{3.122}$$

where the subscripts V and M stand for 'vector' and 'metric' respectively. Analogously to the preceding sections, and to avoid ambiguity, the following notation will henceforth be adopted:

$$\mathcal{L}^a = \frac{\partial \mathcal{L}}{\partial X_a} \tag{3.123}$$

$$\mathcal{L}^{ab} = \frac{\partial \mathcal{L}}{\partial g_{ab}} \tag{3.124}$$

$$\mathcal{L}^{ab}_{(v)} = \frac{\partial \mathcal{L}}{\partial X_{a,b}} \tag{3.125}$$

$$\mathcal{L}^{abc} = \frac{\partial \mathcal{L}}{\partial g_{ab,c}} \tag{3.126}$$

$$\mathcal{L}^{abcd} = \frac{\partial \mathcal{L}}{\partial g_{ab,cd}} , \tag{3.127}$$

where, as before, the subscript (v) is a label corresponding to the vector X_a, and is not a tensor index. One thus has two associated sets of Euler-Lagrange equations for the above combined Lagrangian:

$$\begin{aligned} \mathcal{E}^a(\mathcal{L}) &= \frac{d\mathcal{L}^{ab}_{(v)}}{dx^b} - \mathcal{L}^a \\ &= 0 \end{aligned} \tag{3.128}$$

$$\begin{aligned} \mathcal{E}^{ab}(\mathcal{L}) &= \frac{d}{dx^c}\left[\mathcal{L}^{abc} - \frac{d}{dx^d}(\mathcal{L}^{abdc})\right] - \mathcal{L}^{ab} \quad , \\ &= 0 \end{aligned} \tag{3.129}$$

which, due to the above form of the generalised Lagrangian, leads to the following identities:

$$\mathcal{E}^{ab}(\mathcal{L}) = \mathcal{E}^{ab}(\mathcal{L}_{(M)} - T^{ab} \tag{3.130}$$

$$\mathcal{E}^{a}(\mathcal{L}) = \mathcal{E}^{a}(\mathcal{L}_{(V)}) \quad ; \tag{3.131}$$

where $T^{ab} = \frac{\partial \mathcal{L}_{(V)}}{\partial g_{ab}}$ is the energy-momentum tensor as defined for vector field theory. Firstly, from 3.130 and 3.117 the following results:

$$\mathcal{E}^{ab}(\mathcal{L})_{;b} = -T^{ab}_{;b} \quad . \tag{3.132}$$

Secondly, from 3.88 one obtains:

$$T^{ab}_{;b} = -\frac{1}{2}\left[g^{ab}X_b\mathcal{E}^c(\mathcal{L})_{;c} - g^{ab}\mathcal{F}_{cb}\mathcal{E}^c(\mathcal{L})\right] \quad . \tag{3.133}$$

Using these two results, one obtains from the combined Euler-Lagrange equations 3.128 and 3.129 the following:

$$\begin{aligned} \mathcal{E}^{ab}(\mathcal{L}_{(M)}) &= T^{ab} && (3.134) \\ \mathcal{E}^{a}(\mathcal{L}_{(V)}) &= 0 \ . && (3.135) \end{aligned}$$

As an example, one may consider the following Lagrangian:

$$\mathcal{L} = \sqrt{-g}\left(R - 2\Lambda + \frac{\kappa}{2} g^{ab} g^{cd} F_{ca} F_{db}\right) \quad , \tag{3.136}$$

such that R is the Ricci scalar, and F_{ab} is defined by 3.81, while κ and Λ are constants. Hence, applying the Euler-Lagrange equations and using the above results, one obtains:

$$R^{ab} - \frac{1}{2} R g^{ab} + \Lambda g^{ab} = \kappa\left(\frac{1}{4} g^{ab} F^{cd} F_{cd} - F^{ac} F_{dc} g^{db}\right) \quad , \tag{3.137}$$

which, if F_{ab} is the Faraday tensor, yield the Einstein-Maxwell equations governing the interaction between gravitational and electromagnetic fields.

As another example one may consider a coupling between a Klein-Gordon scalar field and gravity by postulating the Lagrangian:

$$\mathcal{L} = \sqrt{-g}\left(R - \frac{1}{2} g^{ab} \phi_{;a} \phi_{;b} + m^2 \phi^2\right) \quad , \tag{3.138}$$

for a Klein-Gordon scalar field ϕ.

3.3.5 Conserved Quantities and Noether's Theorem

Relating to the notion of invariance as defined at the beginning of this section, one can derive a powerful result known as *Noether's theorem*, which yields conserved quantities associated with certain symmetries embodied in the invariant transformations of the action. To this end, consider as before a simplistic action of the form:

$$I = \int_C \mathcal{L}(x^a, \dot{x}^a, t) dt \quad , \tag{3.139}$$

such that $\dot{x}^a = \frac{d}{dt} x^a$. Now assume this action to be invariant under the following s-parameter group of transformations:

$$\begin{aligned} x'^a &= x^a + Y^a_b(x^a, t) u^b && (3.140) \\ t' &= t + Z_b(x^a, t) u^b \quad , && (3.141) \end{aligned}$$

such that u^b for $b = 1, \ldots s$ denotes the s parameters of the transformation group. Such a set of transformations is referred to as a *symmetry*. Hence one has that:

$$\begin{aligned} \left(\frac{\partial x'^a}{\partial u^b}\right)_{u^b=0} &= Y^a_b && (3.142) \\ \left(\frac{\partial t'}{\partial u^b}\right)_{u^b=0} &= Z_b \ . && (3.143) \end{aligned}$$

Imposing then the invariance 3.68 it is straight-forward, though tedious to derive (see Rund and Lovelock for details) the following constraint equations for the Lagrangian $\mathcal{L}$:

$$\mathcal{E}_a(\mathcal{L})(Y_b^a - \dot{x}^a Z_b) = -\frac{dN_b}{dt} \quad , \tag{3.144}$$

such that:

$$N_b \equiv \mathcal{H} Z_b - p_a Y_b^a \tag{3.145}$$

is the *Noether current*, and $\mathcal{H}$ and p_a are the Hamiltonian and canonical momentum respectively, as before. Hence, if x^a is an optimal trajectory of the action, the associated Euler-Lagrange equations will be satisfied resulting in the above becoming:

$$\frac{dN_b}{dt} = 0 \quad . \tag{3.146}$$

Consequently, *the Noether current N_b will be conserved along the optimal trajectory x^a of the action I.* This is Noether's theorem, which is equivalent to stating that there exists a conserved quantity, N_b, associated with the symmetry transformations 3.141. As an example, if one considers the simple single-parameter symmetry transformation:

$$x'^a = x^a \tag{3.147}$$

$$t' = t + u \quad , \tag{3.148}$$

then one obtains $N = \mathcal{H}$. In Classical Mechanics, as the Hamiltonian corresponds to the total energy of the physical system in question, one has from application of Noether's theorem that the total energy of the system is conserved, the associated symmetry being invariance under temporal translation (t naturally corresponding to time). Similarly, if one considers the transformation:

$$x'^a = x^a + u^a \tag{3.149}$$

$$t' = t \quad , \tag{3.150}$$

then one has $N_a = -p_a$, which by Noether's theorem yields conservation of linear momentum, the associated symmetry being invariance under spatial translation. One can generalise Noether's theorem to multiple-integral actions of the form 3.32. In this case one considers invariance under the group of symmetry transformations:

$$x'^a = x^a + Y_b^a(t^c, x^d)u^b \tag{3.151}$$

$$t'^\alpha = t^\alpha + Z_f^e(t^c, x^d)u^f \quad . \tag{3.152}$$

One then has, analogous to 3.144:

$$\mathcal{E}_a(\mathcal{L})(Y_b^a - \dot{x}_\alpha^a Z_b^\alpha) = -\frac{dN_b^\alpha}{dt^\alpha} \quad , \tag{3.153}$$

such that:

$$N_b^\alpha = \mathcal{H}_\beta^\alpha Z_b^\beta - \frac{\partial \mathcal{L}}{\partial \dot{x}_\alpha^c} Y_b^c \quad . \tag{3.154}$$

3.4 Bibliography

- H Rund and D Lovelock (1975) *Tensors, Differential Forms and Variational Principles* (Wiley)
- D Ter Haar (1971) *Elements of Hamiltonian Mechanics* (Pergamon)
- H Goldstein (1950) *Classical Mechanics* (Addison Wesley)
- TWB Kibble (1985) *Classical Mechanics* (Longman)
- RP Feynman (1990) *Feynman Lectures on Physics* (Penguin)
- IM Gelfand and SV Fomin (1963) *Calculus of Variations* (Prentice-Hall)
- M Giaquinta and S Hildebrandt (1996) *Calculus of Variations I* (Springer)
- H Sagan (1969) *Introduction to the Calculus of Variations* (McGraw-Hill)

Chapter 4

Gravitation

'The magic of this theory is such that almost no-one can escape it once he has understood it properly."
A. Einstein

In this chapter the physics of General Relativity will be formulated within the mathematical context of Differential Geometry. The formalism will then be readily conducive mathematically to a formal perturbative analysis, thus providing the requisite regime in which to study Cosmological Perturbation Theory later.

It will be attempted to cover all the major quantities utilised in Cosmology, although not all of them will be needed later. This, again, is done for the sake of completion.

4.1 Physics Within the Framework of Differential Geometry

Naturally, the motivation for pursuing differential geometry as the principle mathematical *modus operandi* lies in the inherent geometrical nature of General Relativity as postulated by the Einstein Field Equations. These equations explicitly demonstrate the direct dependence between geometry and matter, thereby renouncing the absolute Euclidean assumption apropos of the physicality of space. Adopting standard geometric nomenclature, one refers to Euclidean space as being *flat*, and *curved* otherwise, as formulated in Chapter 1; the best analogy illustrating this is in the two-dimensional case when comparing an infinite plane (flat space) with the surface of a sphere (curved space). Combining this notion of matter-geometry dependence with the postulates of Special Relativity concerning the absolute nature of the speed of light, one thus obtains the most tangible result of General Relativity, namely the manifest *curvature* of space and time together: *curved space-time.*

This generalisation of the nature of the physical geometry will then effec t, as earlier expounded in chapter 2, the notions of differentiation. Consequently, in order to interpret standard physical quantities, which are normally defined relative to other quantities via derivatives, in terms of a generalised curved geometry, one needs to make a few distinctions and identifications. These amount to the assumption of *minimal coupling*;

that is, making the simplest possible transition from known physical equations in a flat space-time to *covariant* equations in curved space-time.

The first distinction, which can be seen within the specific case of *Special* Relativity, is the difference between co-ordinate and proper time. In general Relativity one can *define* the proper time τ in terms of the *space-time arclength* ds between two points p and q as follows:

$$c\tau = \int_p^q \sqrt{|ds^2|}dt \quad , \tag{4.1}$$

where t is the *co-ordinate time* associated with a particular co-ordinate system, measuring the component x^0 of the four-space *position* vector x^a, and thus *variable* subject to the chosen co-ordinate system; while c is the speed of light.

As the arclength must be a physical quantity, and thus invariant under co-ordinate transformations, one sees that the proper time is a *physical invariant*. In the Special Relativity case one has:

$$d\tau = dt\sqrt{1 - \frac{v^2}{c^2}} \quad , \tag{4.2}$$

where v is the velocity of the observer. Hence, one would prefer to express 'physical' time derivatives in terms of the proper time τ, as this is itself invariant. Consequently, one has the first identification: $t \to \tau$.

The second identification concerns the nature of partial differentiation. As shown in chapter 2, the generalisation of the partial derivative in a curved space is the *covariant derivative*, which by formulation is *co-ordinate invariant*. Hence physical quantities, which in flat space are expressed in terms of partial derivatives of some other physical quantities, will now be generalised to covariant derivatives. Consequently, by the above one should express the time derivative in terms of the *absolute derivative*. Hence one now has the identifications: $\frac{\partial}{\partial x^a} \to \nabla_a$; and $\frac{d}{dt} \to \frac{D}{d\tau}$ such that the following hold:

$$u_a \equiv \frac{d}{d\tau}x_a \tag{4.3}$$

$$\nabla_b Y_a \equiv Y_{a;b} \tag{4.4}$$

$$\frac{D}{d\tau}Y_a \equiv Y_{a;b}u^b \quad . \tag{4.5}$$

4.2 General Relativity and the Field Equations

Having established the nature of Physics within a curved space-time, and utilised the mathematics of Differential Geometry, one now has an arena for General Relativity. Without digressing on the philosophical development culminating in the celebrated field equations, one can summarise by saying that, simply put, Einstein postulated that *the matter in the universe determines the geometry, or space-time structure.* For his General Relativity, Einstein eventually arrived thus at the following form of the Field Equations:

$$G_{ab} + \Lambda g_{ab} = \frac{8\pi\mathcal{G}}{c^4}T_{ab} \quad , \tag{4.6}$$

where $G_{ab} = R_{ab} - \frac{1}{2}Rg_{ab}$ is the *Einstein tensor*, containing the necessary definitive geometrical quantities of the space-time; g_{ab} is the metric of the space-time; Λ is the *cosmological constant* satisfying $\Lambda_{;a} = 0$, and is included for the sake of generality and completion; and T_{ab} is the *energy momentum tensor* containing the matter description of the universe, or region of space-time in question. Normally, geometrical units are chosen such that the coupling constant $\frac{8\pi G}{c^4} \equiv 1$ for convenience. Now, by construction, one has $G^{ab}{}_{;b} = 0$ and hence the field equations must necessarily satisfy:

$$T^{ab}_{\ ;b} = 0 \quad . \tag{4.7}$$

In addition, the Bianchi and Ricci identities have to be satisfied. These will be investigated in full in the next section. Equation 4.6, being the central, definitive equation of General Relativity, will thus form the core of all the subsequent analysis.

4.3 Decomposition of the Field Equations in the Comoving Frame

For most practical considerations it is necessary to select a reference frame with respect to which relevant quantities can be defined. In the subsequent analysis this will usually be chosen to be the rest frame of the matter fluid; i.e. comoving co-ordinates x^a will be utilised such that:

$$\begin{aligned} u^a &= \frac{dx^a}{d\tau} \\ &\doteq \delta^a_0 \qquad (4.8) \\ u^a u_a &= -1 \quad , \qquad (4.9) \end{aligned}$$

where τ is the proper time. Hence a tensorial quantity can be decomposed into its components parallel, and orthogonal to, the four velocity u^a; that is, the temporal and spatial components respectively apropos of the matter rest frame. To this end, one can thus formulate the temporal and spatial *projection tensors* U^{ab} and h^{ab} respectively as follows:

$$\begin{aligned} U^{ab} &\equiv -u^a u^b \qquad (4.10) \\ h^{ab} &\equiv g^{ab} + u^a u^b \\ &= g^{ab} - U^{ab} \quad , \qquad (4.11) \end{aligned}$$

such that:

$$\begin{aligned} h^b_a &= g_{ac}h^{cb} \qquad (4.12) \\ h^c_a h^b_c &= h^b_a \quad . \qquad (4.13) \end{aligned}$$

Hence, for example, the totally spatially projected components ${V_\perp}^{ab}$ of an arbitrary tensor V^{ab} will be:

$$V_{\perp}{}^{ab} \equiv h^a_c h^b_d V^{cd} \quad , \tag{4.14}$$

which thus guarantees:

$$V_{\perp}{}^{ab} u_a = 0 = V_{\perp}{}^{ab} u_b \quad ; \tag{4.15}$$

that is, it is orthogonal to u_a on *both* indices; as expected. It is important to note here that in the case of non-zero vorticity, the above decomposition is inane, as the resultant 'spatial projection' does not form a hypersurface, in terms of a submanifold, by Frobenius' Theorem (section 2.2.2). This issue will examined more closely later in the gauge-invariant formulation of Perturbation Theory (Chapters 5 and 6). Having thus established the above operating regime, one can formulate the full set of relativistic evolution equations therein. Firstly, however, one needs to consider what constitutes a full and complete set of evolution equations.

Such a set of equations must minimally, yet fully, determine the complete structure and nature of space-time. These are obtained by applying: 1) the Einstein Field equations; 2) the Ricci identities, which yield constraint and propagation equations for the kinematic quantities; 3) the Bianchi identities - this gives propagation and constraint equations for the gravito-electric and gravito-magnetic tensors; 4) the tetrad equations characterising the metric and connection. The first three considerations will be seen to yield twelve equations in total: six constraint equations, and six propagation equations. The remaining tetrad equations will be dealt with separately. The nature of these equations will be explained shortly. Firstly though, one needs to formulate the kinematic quantities.

4.3.1 The Kinematic Quantities

By application of the projection tensor h_{ab}, one can make the following decomposition of the covariant derivative of the co-moving velocity:

$$\begin{aligned}
u_{a;b} &= \delta^c_a \delta^d_b u_{c;d} \\
&= (h^c_a - u_a u^c)(h^d_b - u_b u^d) u_{c;d} \\
&= \frac{1}{2} h^c_a h^d_b (u_{c;d} - u_{d;c}) + \frac{1}{2} h^c_a h^d_b (u_{c;d} + u_{d;c}) - \dot{u}_a u_b \\
&\equiv \omega_{ab} + \Theta_{ab} - \dot{u}_a u_b \\
&= \omega_{ab} + (\Theta_{ab} - \frac{1}{3}\Theta^c_c h_{ab}) + \frac{1}{3}\Theta^c_c h_{ab} - \dot{u}_a u_b \\
&\equiv \omega_{ab} + \sigma_{ab} + \frac{1}{3}\Theta h_{ab} - \dot{u}_a u_b \quad ,
\end{aligned} \tag{4.16}$$

where the kinematic quantities defined in the above equations are the *vorticity* ω_{ab}, *expansion* Θ_{ab}, and *shear* σ_{ab} tensors. Quantities related to the vorticity and expansion tensors, but more useful in practical analysis will be the vorticity vector:

$$\omega^a \equiv \frac{1}{2}\eta^{abcd} u_b \omega_{cd} \quad , \tag{4.17}$$

which is sometimes plainly referred to as *the vorticity*; and the *expansion*:

$$\Theta \equiv \Theta^c_c = u^c_{\ ;c} \quad . \tag{4.18}$$

It will be convenient and consistent to define the magnitudes of the shear and expansion as follows:

$$\sigma^2 \equiv \frac{1}{2}\sigma_{ab}\sigma^{ab} \tag{4.19}$$

$$\begin{aligned} \omega^2 &\equiv \frac{1}{2}\omega_{ab}\omega^{ab} \\ &= \omega_a\omega^a \quad . \end{aligned} \tag{4.20}$$

$$\tag{4.21}$$

Roughly speaking, the vorticity vector is a rotational quantity denoting the rotation of matter, pointing along the axis of rotation; the expansion parameter (sometimes simply referred to as 'the expansion'), as its name suggests, denotes the uniform, symmetric spatial expansion of matter, that is, the uniform change in volume; while the shear tensor indicates, naturally, the spatial shearing of matter.

4.3.2 Matter Behaviour

Naturally one requires a tensorial formulation of the matter characteristics of the universe. By analogy with standard Classical Mechanics and Electromagnetism, this is achieved through the specification of a stress-energy, or energy-momentum, tensor. The most general form of the energy-momentum tensor, which describes the matter content of a particular type in the universe, is:

$$T_{ab} = \rho u_a u_b + 2q_{(a}u_{b)} + \pi_{ab} + ph_{ab} \quad , \tag{4.22}$$

where p is the isotropic pressure; ρ is the relativistic energy density; q_a is the relativistic momentum density; and π_{ab} represents the anisotropic (traceless) stresses. In general, the specific form thereof will be determined by the physical situation in question. For cosmological models, the stress-energy tensor will be simplified to that of a *perfect fluid*, as will be seen later. Naturally, as one would have such a tensor for each specified matter type, if a multiple matter situation were to be considered one would have a 'total' matter energy-momentum tensor as being the algebraic sum of the energy-momentum tensors for the constituent matter types. This will briefly be considered for multi-fluid matter scenarios in Chapter 5.

4.3.3 The Weyl Conformal Curvature Tensor

Of the twenty components of the Riemann tensor, only ten are encoded within the Ricci tensor and hence the Einstein Field Equations; the remaining ten components are described by the *Weyl conformal curvature tensor*, defined (in four-dimensional spacetime) as follows:

$$C_{abcd} = R_{abcd} - \frac{1}{2}(R_{ac}g_{bd} - R_{ad}g_{bc} + R_{bd}g_{ac} - R_{bc}g_{ad}) + \frac{R}{6}(g_{ac}g_{bd} - g_{ad}g_{bc}) \quad , \quad (4.23)$$

which consequently possesses the same symmetries as the Riemann tensor, but has the important property of a *vanishing trace*. The Weyl tensor can thus be thought of as representing the "free gravitational field", as it is not incorporated into the field equations, although still an essential part of the Riemann tensor which is necessary in its entirety to determine the space-time geometrical structure completely. It also characterises the conformal properties of space-time, and is analogous to the electromagnetic field tensor. Bearing the latter property in mind, it can be shown that the following decomposition of the Weyl tensor in terms of quantities analogous to electric and magnetic fields is possible:

$$C_{abcd} = (\eta_{abpq}\eta_{cdrs} + g_{abpq}g_{cdrs})u^p u^r E^{qs} - (\eta_{abpq}\eta_{cdrs} + g_{abpq}g_{cdrs})u^p u^r H^{qs} \quad , \quad (4.24)$$

with the definition

$$g_{abcd} = g_{ac}g_{bd} - g_{ad}g_{bc} \quad . \quad (4.25)$$

such that:

$$E_{ac} = C_{abcd}u^b u^d \quad (4.26)$$

is the *gravito-electric* tensor, and

$$H_{ac} = \frac{1}{2}\eta_{ab}{}^{gh}C_{ghcd}u^b u^d \quad (4.27)$$

is the *gravito-magnetic* tensor. As their names suggest, these two tensors are analogous to the electric and magnetic fields in Electromagnetism.

Naturally , the Weyl tensor can be related to the Ricci tensor by taking the covariant derivative of the Riemann tensor; this will then enable one to express the magnetic and electric tensors in terms of the kinematic quantities, eventually yielding:

$$\begin{aligned} E_{ab} &= h_{ab}(\frac{1}{3}\omega^2 - \frac{2}{3}\sigma^2 - \dot{u}^a_{;a}) + \dot{u}_a\dot{u}_b - \omega_a\omega_b - \sigma_a{}^f\omega_{fb} \\ &\quad + h_a^f h_b^g(\dot{u}_{(f;g)} - S^{-2}(S^2\sigma_{fg})^{\cdot} + \frac{1}{2}\kappa\pi_{ab} \quad (4.28) \\ H_{ad} &= 2\dot{u}_{(a}\omega_{d)} - h_a^t h_d^s(\omega_{(t}{}^{b;c} + \sigma_{(t}{}^{b;c})\eta_{s)fbc}u^f \quad . \quad (4.29) \end{aligned}$$

4.4 The Full set of Evolution Equations

One is now in a position to formulate the full set of twelve propagation and constraint equations using the Ricci and Bianchi identities. The propagation equations describe the way the various dynamical quantities behave from one spatial hypersurface to another in sequence according to the proper time definition, thus being interpreted as *time* evolution equations, while the constraint equations govern the relationship and behaviour of these quantities in a particular such hypersurface. As noted earlier in the chapter, such spatial hypersurfaces are uniquely defined only if there is vanishing vorticity; otherwise, in the case of non-zero vorticity, the propagation equations describe the change in proper time along the flow lines.

4.4.1 The Ricci Identities

For convenience, the following notation will be employed:

$$
\begin{aligned}
{}^{(3)}\nabla_a &\equiv h_a^b \nabla_b && (4.30)\\
v^{<a>} &\equiv h_b^a v^b && (4.31)\\
T^{<ab>} &\equiv \left(h_c^{(a} h_d^{b)} - \frac{1}{3} h^{ab} h_{cd} \right) T^{cd} && (4.32)\\
\text{curl}\, A^{ab} &\equiv \eta^{cd<a}\, {}^{(3)}\nabla_c A_d^{b>} \quad , && (4.33)
\end{aligned}
$$

where the second-last definition indicates the orthogonally projected symmetric and trace-free part of the tensor T^{ab}. From the Ricci identities:

$$2\nabla_{[a}\nabla_{b]} u^c = R_{ab\ d}^{\ \ c} u^d \quad (4.34)$$

one can derive three propagation equations for the kinematic quantities, as well as three constraint equations. The three propagation equations are:

$$
\begin{aligned}
\dot{\Theta} - {}^{(3)}\nabla_a \dot{u}^a &= -\frac{1}{3}\Theta^2 + \dot{u}^a \dot{u}_a - 2(\sigma^2 - \omega^2) - \frac{1}{2}(\rho + 3p) + \Lambda && (4.35)\\
\dot{\omega}^{<a>} - \frac{1}{2}\eta^{abc}\, {}^{(3)}\nabla_b \dot{u}_c &= -\frac{2}{3}\Theta\omega^a + \sigma_b^a \omega^b && (4.36)\\
\dot{\sigma}^{<ab>} - {}^{(3)}\nabla^{<a}\dot{u}^{b>} &= -\frac{2}{3}\Theta\sigma^{ab} + \dot{u}^{<a}\dot{u}^{b>} - \sigma_c^{<a}\sigma^{b>c} - \omega^{<a}\omega^{b>} - (E^{ab} - \frac{1}{2}\pi^{ab}) \quad , && (4.37)
\end{aligned}
$$

for the expansion, vorticity and shear respectively. The first equation is known as the *Raychaudhuri equation*, and is the basic equation of gravitational attraction. The three constraint equations are:

$$
\begin{aligned}
{}^{(3)}\nabla_b \sigma^{ab} &= \frac{2}{3}\, {}^{(3)}\nabla^a \Theta + \eta^{abc}(\, {}^{(3)}\nabla_b \omega_c + 2\dot{u}_b \omega_c) + q^a && (4.38)\\
{}^{(3)}\nabla_a \omega^a &= \dot{u}_a \omega^a && (4.39)\\
\text{curl}\, \sigma^{ab} &= H^{ab} + 2\dot{u}^{<a}\omega^{b>} + {}^{(3)}\nabla^{<a}\omega^{b>} \quad . && (4.40)
\end{aligned}
$$

Notice that for the case of vanishing vorticity, one has the following equation which is equivalent to the shear propagation equation for vanishing vorticity:

$$R^{(3)}_{ab} = h^f_a h^g_b[\dot{u}_{(f;g)} - S^{-3}(S^3\sigma_{fg})^{\cdot}] + \frac{2}{3}(-\frac{1}{3}\Theta^2 + \sigma^2 - \frac{1}{2}\dot{u}^c{}_{;c} + \Lambda + \rho)h_{ab} + \pi_{ab} \quad , \tag{4.41}$$

such that:

$$R^{(3)}_{abcd} = (R_{abcd})_\perp - \Theta_{ac}\Theta_{bd} + \Theta_{bc}\Theta_{da} \quad . \tag{4.42}$$

Equation 4.41 is known as the *Gauss-Codacci* equation, and shows how the matter tensor directly effects the Ricci curvature of the three-space.

4.4.2 The Twice-Contracted Bianchi Identities

These amount to $T^{ab}{}_{;b} = 0$, and yield the conservation equations:

$$\dot{\rho} + {}^{(3)}\nabla_a q^a = -\Theta(\rho + p) - 2\dot{u}_a q^a - \sigma^a_b \pi^b_a \tag{4.43}$$

$$\dot{q}^{<a>} + {}^{(3)}\nabla^a p + {}^{(3)}\nabla_b \pi^{ab} = -\frac{4}{3}\Theta q^a - \sigma^a_b q^b - (\rho + p)\dot{u}^a - \dot{u}_b \pi^{ab} - \eta^{abc}\omega_b q_c \quad . \tag{4.44}$$

These are respectively an evolution equation and a constraint equation for the matter. If the matter involved is of a *perfect fluid* form, *i.e.* if there are no anisotropic stresses and momentum densities:

$$T_{ab} = \rho u_a u_b + p h_{ab} \quad . \tag{4.45}$$

then the above constraint equations simplify to the standard *conservation equations*:

$$\dot{\rho} + (\rho + p)\Theta = 0 \tag{4.46}$$

$$(\rho + p)\dot{u}_a + h^c_a p_{,c} = 0 \quad . \tag{4.47}$$

These simplified relations will later form the basis of idealised background models in perturbation theory.

4.4.3 The Bianchi Identities

From the Bianchi Identities:

$$\nabla_{[a} R_{bc]de} = 0 \tag{4.48}$$

one obtains the following two propagation equations for the gravito-electric and gravito-magnetic tensors:

$$\dot{E}^{<ab>} + \frac{1}{2}\dot{\pi}^{<ab>} - \text{curl}\, H^{ab} + \frac{1}{2}\, {}^{(3)}\nabla^{<a}q^{b>} = -\frac{1}{2}(\rho+p)\sigma^{ab} - \Theta(E^{ab} + \frac{1}{6}\pi^{ab}) + 3\sigma_c^{<a}(E^{b>c} - \frac{1}{6}\pi^{b>c}) - \dot{u}^{<a}q^{b>} + \eta^{cd<a}\left[2\dot{u}_c H_d^{b>} + \omega_c(E_d^{b>} + \frac{1}{2}\pi_d^{b>})\right] \quad (4.49)$$

$$\dot{H}^{<ab>} + \text{curl}\, E^{ab} - \frac{1}{2}\text{curl}\, \pi^{ab} = -\Theta H^{ab} + 3\sigma_c^{<a}H^{b>c} + \frac{3}{2}\omega^{<a}q^{b>} - \eta^{cd<a}\left[2\dot{u}_c E_d^{b>} - \frac{1}{2}\sigma_c^{b>}q_d - \omega_c H_d^{b>}\right] \quad (4.50)$$

and the two constraint equations:

$$ {}^{(3)}\nabla_b(E^{ab} + \frac{1}{2}\pi^{ab}) = \frac{1}{3}\, {}^{(3)}\nabla^a\rho - \frac{1}{3}\Theta q^a + \frac{1}{2}\sigma_b^a q^b + 3\omega_b H^{ab} + \eta^{abc}\left[\sigma_{bd}H_c^d - \frac{3}{2}\omega_b q_c\right] \quad (4.51)$$

$$ {}^{(3)}\nabla_b H^{ab} = -(\rho+p)\omega^a - 3\omega_b(E^{ab} - \frac{1}{6}\pi^{ab}) + \eta^{abc}\left[\frac{1}{2}\, {}^{(3)}\nabla_b q_c + \sigma_{bd}(E_c^d + \frac{1}{2}\pi_c^d)\right] . \quad (4.52)$$

Hence, from the Bianchi identities one has six equations: a matter propagation equation and evolution equation; two propagation and two evolution equations for the gravito-electric and gravito-magnetic tensors. This, combined with the three propagation and three constraint equations for the kinematic quantities from the Ricci identities yields the requisite twelve equations.

4.5 Tetrad Formulation of the Field Equations in Cosmology

The twelve constraint and propagation equations as formulated in section 4.4, although geometrically and physically lucid and intuitive, do not represent a complete set in the sense of guaranteeing the existence of a corresponding metric and connection. For completeness, one has to consider the nature of the vectorial basis of the reference frame in which one describes the physical quantities in question. This requires the notion of tetrads, as defined in section 2.5.1, and wherein the above equations can be reformulated. The additional equations which then accrue, will essentially result from application of the Bianchi and Ricci identities to these basis vectors; these will not be derived here, but the underlying notions and relevant quantities will be given. In the above formulation of the Field Equations, the matter co-moving frame tetrad was implicitly assumed; however, only the equation components *relative to the tetrad basis* were given. One still needs to study the behaviour of the tetrad basis itself, which thus requires additional evolution equations.

4.5.1 The Comoving 1+3 Tetrad Formulation

In the cosmological application, it is useful to define the tetrad of a specific model by choosing $\mathbf{e}_0$ to be the unit tangent of the fluid flow, u^a, referred to as the 1+3 formalism. This is one of the most physically lucid and useful tetrad choices; consequently a more detailed analysis of the Field Equations formulated therein is warranted. Proceeding thus, one can derive the following form of the metric connections from equation 2.124:

$$\Sigma_{\alpha 00} = \dot{u}_\alpha \tag{4.53}$$

$$\Sigma_{\alpha 0\beta} = \frac{1}{3}\Theta\delta_{\alpha\beta} + \sigma_{\alpha\beta} - \epsilon_{\alpha\beta\gamma}\omega^\gamma \tag{4.54}$$

$$\Sigma_{\alpha\beta 0} = \epsilon_{\alpha\beta\gamma}\Omega^\gamma \tag{4.55}$$

$$\Sigma_{\alpha\beta\gamma} = 2a_{[\alpha}\delta_{\beta]\gamma} + \epsilon_{\gamma\delta[\alpha}n^{\delta}{}_{\beta]} + \frac{1}{2}\epsilon_{\alpha\beta\gamma}n^{\delta}{}_{\gamma} \quad , \tag{4.56}$$

where Greek letters have been used toindicate the *spatial frame* tetrad indices. The first two quantities in the above contain the kinematical variables, while the latter two encapsulate the rate of rotation Ω^α of the tetrad spatial frame $\{\mathbf{e}_\alpha\}$ with respect to a Fermi-propagated basis; while a^α and $n^{\alpha\beta} = n^{(\alpha\beta)}$ determine the nine spatial rotation co-efficients. Using equation 2.125, one can also derive the commutation relations in terms of the above quantities:

$$[\mathbf{e}_0, \mathbf{e}_\alpha] = \dot{u}_\alpha \mathbf{e}_0 - \left[\frac{1}{3}\Theta\delta^\beta_\alpha + \sigma_\alpha^{\ \beta} + \epsilon_\alpha{}^\beta{}_\gamma\left(\omega^\gamma + \Omega^\gamma\right)\right]\mathbf{e}_\beta \tag{4.57}$$

$$[\mathbf{e}_\alpha, \mathbf{e}_\beta] = 2\epsilon_{\alpha\beta\gamma}\omega^\gamma \mathbf{e}_0 + \left[2a_{[\alpha}\delta^\gamma_{\beta]} + \epsilon_{\alpha\beta\gamma}n^{\delta\gamma}\right]\mathbf{e}_\gamma \tag{4.58}$$

One can follow the foregoing procedure to derive constraint and propagations equations for the field equations, incorporating the above tetrad basis. Naturally, this will then yield equations which will incorporate derivatives of the tetrad components, ultimately resulting in equations which are considerably more complex than those in the previous section; these, however will not be needed later, and so will not be derived here. The full set of equations and affiliated analysis can be found in the Ellis Cargese Lectures [21].

However, from the above commutation relations one can make a crucial observation: the rest frame spatial basis vectors $\mathbf{e}_\alpha$ are themselves closed under the Lie bracket operation if and only if the vorticity vanishes. This follows immediately from equation 4.58, as a vanishing vorticity vector ω^γ eliminates the unwanted $\mathbf{e}_0$ basis vector which is parallel to the four-velocity. Upon application of the theorem of Frobenius (section 2.2.2), one then deduces that *the spatial rest frame itself forms a submanifold, thereby defining a unique foliation of space-time with respect to the comoving velocity, if and only if the vorticity vanishes.* Hence, if one has a scenario including a non-vanishing vorticity, the rest frames are only local as opposed to being global hypersurfaces, and are not unique.

Now, in the case where the spatial rest frame does constitute a hypersurface, one can naturally obtain a circular trajectory in this spatial frame which is orthogonal to all the world lines it intersects; the fact that it is circular means that it always meets up, and this embodies the notion of the integral curves 'meshing up' to form a hypersurface. However, in the case of non-zero vorticity, this is not the case. One can visualise this by considering a world-line congruence under pure vorticity - i.e. no expansion or shear. Depicting these in a space-time diagram with the conventional suppression of one spatial dimension, one notices that the emergent picture is reminiscent of the individual fibre strands of a vertical twisted rope. All the strands wind around each other in a helical structure, maintaining a constant distance between them - this embodies the two fundamental properties of pure vorticity: rotation and fixed distance. Now, one would imagine a rest-frame hypersurface to correspond to a horizontal slicing of this rope, as this should be perpendicular to the strands. However, this is not the case, as elementary Euclidean geometry shows that the only perpendicular curve to a congruence of helices is another helix curving in the opposite direction - that is, one can move continually from strand to strand in the rope in a direction always orthogonal to the strands by following another helical path. As this orthogonal trajectory is itself a vertical helix, following it one would eventually travel right around the perimeter of the rope, BUT one would then be either above or below to where one originally started out - in space-time this would mean one would have either advanced or regressed in time respectively. As one normally uses the notion of a spatial hypersurface in Cosmology to define a global time uniquely with respect thereto, one now notices that in the case of vorticity this is clearly impossible.

4.5.2 The Arnowit-Deser-Misner Formalism

An alternative to the afore-mentioned tetrad formalism is the Arnowit-Deser- Misner approach (ADM), referred to sometimes as the $3+1$ formalism. This will be used later in the Bardeen formalism and in the variational formulation of Cosmological Perturbation Theory. The ADM formalism entails considering space-time as an ordered sequence of $3-$geometries, or space-like hypersurfaces. This foliation is only applicable to space-times which are *globally hyperbolic*; that is, space-times which contain a *Cauchy surface* (Wald chapter 5). One can label each such hypersurface with a co-ordinate time t and indicate space-like points within the surface as $x^\alpha = (x^1, x^2, x^3)$. The *intrinsic* geometry of the hypersurface will then be described by a spatial $3-$metric $o_{\alpha\beta}$ with inverse $o^{\alpha\beta}$.

One can now define a *lapse function* N which measures the proper time interval $\Delta\tau$ between hypersurfaces at t and $t+\Delta\tau$ along some worldline normal to the hypersurfaces:

$$\Delta\tau = N\Delta t \quad . \tag{4.59}$$

Similarly, spatial co-ordinates may vary in a continuous way from one hypersurface to another, this change being described in terms of a *shift vector* N^α; in terms of a change from, as above, a hypersurface at t to one at $t+\Delta\tau$, this change transpires as:

$$\Delta x^\alpha = N^\alpha \Delta t \quad . \tag{4.60}$$

Hence it is evident that the shift vector N^α is the co-ordinate $3-$velocity of an observer at rest in the hypersurface. One can thus regard the essence of the ADM formalism as being the *choice of a global foliation of spatial hypersurfaces from which co-ordinate and proper time variables are induced.* Note that this is distinct from the $1+3$ covariant formalism whereby a proper time is defined, from which *local* spatial hypersurfaces are induced. One can now write the full space-time metric g_{ab} in terms of the lapse function N and the shift vector N^α:

$$\begin{aligned} g_{00} &= N^\alpha N_\alpha - N^2 && (4.61)\\ g_{0\alpha} &= -N_\alpha && (4.62)\\ g_{\alpha\beta} &= o_{\alpha\beta} \quad , && (4.63) \end{aligned}$$

where $N_\alpha = o_{\alpha\beta} N^\beta$. Hence, using the relation $g_{\beta\gamma} g^{\gamma\alpha} = \delta^\alpha_\beta$, one can calculate the components of the inverse metric g^{ab}:

$$\begin{aligned} g^{00} &= -\frac{1}{N^2} && (4.64)\\ g^{0\alpha} &= -\frac{N^\alpha}{N^2} && (4.65)\\ g^{\alpha\beta} &= o^{\alpha\beta} - \frac{N^\alpha N^\beta}{N^2} \quad , && (4.66) \end{aligned}$$

where $o_{\alpha\beta}$ is the intrinsic metric tensor; *intrinsic* in that it defines the geometry of the *spatial* hypersurfaces, but does not define how these surfaces are embedded in the space-time. The above now yields the unit normal to the hypersurface as:

$$\begin{aligned} n^a &= \frac{1}{N}(1, N^1, N^2, N^3) && (4.67)\\ n_a &= -N(1,0,0,0) \quad . && (4.68) \end{aligned}$$

This leads to the *projection tensor* P_{ab} which projects quantities onto the three-space constant-time hypersurfaces:

$$p_{ab} = g_{ab} + n_a n_b \quad . \qquad (4.69)$$

One can then define the *extrinsic curvature* tensor $K_{\alpha\beta}$ as being a measure of the local bending of the spatial hypersurfaces as they are stacked together, forming the 4−dimensional space-time. This tensor thus describes how the normal 4−vectors go from one point on a hypersurface to another; i.e. how these normals to the spatial hypersurfaces diverge. It is thus defined as:

$$\begin{aligned} K_{\alpha\beta} &= -n_{\alpha;\beta} && (4.70)\\ &= -N\Gamma^0{}_{\alpha\beta} && (4.71)\\ &= \frac{1}{2N}(2N_{(\alpha;\beta)} - o_{\alpha\beta}{}') \quad . && (4.72) \end{aligned}$$

where the primed sign indicates differentiation with respect to proper time. From this definition it is evident that, in the case of vanishing vorticity, the extrinsic curvature tensor is equal to the negative of the expansion tensor θ_{ab}. Hence one immediately notices that the negative trace $-K$ thereof describes the rate of expansion of the normal world lines. Likewise one has a shear quantity:

$$\overline{K_{\alpha\beta}} \equiv K_{\alpha\beta} - \frac{1}{3}K o_{\alpha\beta} \quad . \tag{4.73}$$

Having defined the projection and extrinsic curvature tensors, one can now perform a $3+1$ split of any tensorial quantity, analogous to that carried out in the $1+3$ formalism. Most importantly, one can perform the necessary splitting of the energy-momentum tensor, defining thus the energy density, momentum density, and stress tensor respectively:

$$2E \equiv n_a n_b T^{ab} = NT^{00} \tag{4.74}$$
$$J_\alpha \equiv -n_a P_{b\alpha} = NT^0_\alpha \tag{4.75}$$
$$S_{\alpha\beta} \equiv P_{\alpha a}P_{\beta b}T^{ab} = T_{\alpha\beta} \quad . \tag{4.76}$$

This splitting thus leads to the following form of the energy and momentum conservation equations:

$$E_{,0} = NKE + NK^{\alpha\beta}S_{\alpha\beta} - \frac{1}{N}(N^2J^\alpha)_{;\alpha} - N^\alpha E_{,\alpha} \tag{4.77}$$
$$J_{\alpha,0} = NKJ_\alpha - (E\delta^\gamma_\alpha + S_\alpha^{\ \gamma})N_{,\gamma} - NS_\alpha{}^\gamma{}_{;\gamma} - N^\gamma J_{\alpha,\gamma} - N^\gamma{}_{,\alpha}J_\gamma \quad , \tag{4.78}$$

where, as before, the semi-colon indicates covariant differentiation with respect to the intrinsic metric tensor $o_{\alpha\beta}$. Similarly splitting up the Einstein Field Equations, one has the constraint equations:

$$E = R - K^{\alpha\beta}K_{\alpha\beta} + K^2 \tag{4.79}$$
$$J_\alpha = K^{\ \beta}_{\alpha;\beta} - K_{,\alpha} \quad , \tag{4.80}$$
$$\tag{4.81}$$

such that the coupling constant has been normalised; and the dynamical equations:

$$K^\alpha_{\beta,0} + N^\gamma K^\alpha_{\beta,\gamma} - N^\alpha_{,\gamma}K^\gamma_\beta + N^\gamma_{,\beta}K^\alpha_\gamma = -N^{;\alpha}_{;\beta} + N\left(R^\alpha_\beta + KK^\alpha_\beta - S^\alpha_\beta + \tfrac{1}{2}\delta^\alpha_\beta(S^\gamma_\gamma - E)\right) \quad . \tag{4.82}$$

Note that there are no evolution equations for N and N^α, as these only entail the arbitrary choice of space-time co-ordinates. In this light, the choice of a time co-ordinate amounts to nothing less than a *hypersurface condition* specifying the way in which the space-time is foliated.

4.6 Cosmology

Throughout the treatment of cosmology that will be developed here, the *standard model* will be presupposed, as this is the simplest and most frequently used. This essentially amounts to the classifications of Friedmann-Lemaître-Robertson-Walker universe models (hereafter denoted in shorthand by "FLRW"); models which obey the underlying assumption of the *Cosmological Principle*: the universe is both *spatially isotropic* (the Copernican Principle) about every point, and spatially homogeneous. The former property implies that the universe is geometrically the same in any direction, thus eliminating the possibility of there being any privileged point (such as "the centre of the universe") or direction in space; while the latter implies the uniformity and equivalence of physical properties on *all* space-like surfaces orthogonal to the matter fluid flow.

4.6.1 The Standard Model in Cosmology

As stated above, the standard model is here taken to be the class of FLRW models. The assumption of geometric isotropy naturally implies, in terms of the dynamical variables previously formulated:

$$\sigma^{\alpha\beta} = 0, \quad \omega^{\alpha} = 0, \quad \dot{u}^{a} = 0 \quad . \tag{4.83}$$

A consequence is that there exists a normalised *proper* time t such that:

$$u^{a} = -t_{,a} \quad , \tag{4.84}$$

which is unique up to a constant; and yields the surfaces of spatial homogeneity t = *constant* which are thus orthogonal to the fluid flow. A further important consequence of the Cosmological Principle is that *the spatial curvature at any point is a constant* $\mathcal{R}$, otherwise all points therein would not be equivalent; Also, spatial isotropy about a point naturally implies spherical symmetry about a point; that is, in terms of the *spatial* Riemann, Ricci and metric tensors:

$$R_{\alpha\beta\gamma\delta} \equiv \mathcal{R}(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}) \tag{4.85}$$

$$R_{\beta\delta} = 2\mathcal{R}g_{\beta\delta} \quad . \tag{4.86}$$

In co-moving co-ordinates the above assumptions imply the following simple line element form for an FLRW model in spherical co-ordinates:

$$ds^2 = -dt^2 + S^2\left(\frac{dr^2}{1-Kr^2} + r^2 d\Omega^2)\right) \quad , \tag{4.87}$$

where r is the spatial radial measure, while $S(t)$ is the *scale factor* characterising radial distances between any pair of observers in the space; and such that:

$$d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\phi^2 \tag{4.88}$$

$$u_a = -t_a = \delta^a_0 \tag{4.89}$$

$$\frac{\dot{S}}{S} = \frac{1}{3}\Theta \tag{4.90}$$

$$K = \begin{cases} 1 & \text{closed} \\ -1 & \text{open} \\ 0 & \text{flat} \end{cases} \quad (4.91)$$

$$\mathcal{R} = \frac{K}{S^2} \quad , \quad (4.92)$$

where $\mathcal{R}$ is the curvature of the three-space, such that $K = 1, -1, 0$ corresponds to a *closed, open* or *flat* spatial geometry respectively.

4.6.2 Geometrical Quantities Associated With the FLRW Models

For practical convenience, one now defines the *conformal time* τ by:

$$d\tau = \frac{dt}{S} \quad . \quad (4.93)$$

This reduces the above line element to the following functional form:

$$ds^2 = S(\tau)^2(-d\tau^2 + dr^2 f^2 d\Sigma^2) \quad . \quad (4.94)$$

Having formulated the line element for FLRW universe models, and thus the associated metric, one can calculate the connection coefficients, yielding:

$$\Gamma^0{}_{00} = \frac{S'}{S} \quad (4.95)$$

$$\Gamma^\alpha{}_{0\beta} = \frac{S'}{S}\delta^\alpha_\beta \quad (4.96)$$

$$\Gamma^0{}_{\alpha\beta} = \frac{S'}{S}o_{\alpha\beta} \quad (4.97)$$

$$\Gamma^\alpha{}_{\beta\gamma} = \Gamma^{(3)\alpha}{}_{\beta\gamma} \quad , \quad (4.98)$$

where here and henceforth a primed sign indicates differentiation with respect to conformal time τ, and a superscript indicates differentiation with respect to the proper time; $o_{\alpha\beta}$ is the spatial part of the metric; and $\Gamma^{(3)\alpha}{}_{\beta\gamma}$ is the Christoffel symbol associated with $o_{\alpha\beta}$. Defining now the quantity $\mathcal{H} = \frac{S'}{S} = S\mathcal{H}$, where $H = \frac{\dot{S}}{S}$ is Hubble's parameter, one has the Following Ricci and Einstein tensor components in conformal time:

$$R^\beta_\alpha = S^{-2}(\mathcal{H}' + 2\mathcal{H}^2 + 2K)\delta^\beta_\alpha \quad (4.99)$$

$$R^0_0 = 3S^{-2}\mathcal{H}' \quad (4.100)$$

$$R^0_\alpha = 0 \quad (4.101)$$

$$R = 6kS^{-2}(\mathcal{H}' + \mathcal{H}^2 + K) \quad (4.102)$$

$$G^0{}_0 = -3S^{-2}(\mathcal{H}^2 + K) \quad (4.103)$$

$$G^0{}_\alpha = 0 \quad (4.104)$$

$$G^\alpha_\beta = -S^{-2}(2\mathcal{H}' + \mathcal{H}^2 + K)\delta^\alpha_\beta \quad . \quad (4.105)$$

(4.106)

Similarly one has, by the co-moving system assumption, the following expression for the co-moving velocity u^a defined with respect to conformal time τ:

$$u^a = S^{-1}\delta_0^a \quad . \tag{4.107}$$

Variation of the Ricci components

For these models one can now calculate the variations of the above quantities. The calculations are rather tedious but elementary. The results are:

$$\begin{aligned}
\delta R_\alpha^\beta &= \frac{1}{2S^2}\left(\delta g_\alpha{}^{\gamma;\beta}{}_{;\gamma} + \delta g^\beta{}_{\gamma;\alpha}{}^{;\gamma} - \delta g_\alpha{}^{\beta;\gamma}{}_\gamma - \delta g_{;\alpha}{}^{;\beta}\right) \\
&\quad + \frac{1}{2S^2}\delta g_\alpha^\beta + \frac{S'}{S^3}\delta g_\alpha^\beta + \frac{S'}{2S^3}(\delta g)'\delta_\alpha^\beta + K\frac{2}{S^2}\delta g_\alpha^\beta && (4.108) \\
\delta R_0^0 &= \frac{1}{2S^2}(\delta g)'' + \frac{S'}{2S^3}(\delta g)' && (4.109) \\
\delta R_\alpha^0 &= \frac{1}{2S^2}\left((\delta g)'_{;\alpha} - \delta g^b_{\alpha;\beta}\right) && (4.110) \\
\delta R &= \frac{1}{S^2}\left(\delta g_\alpha{}^{\alpha;\beta}{}_{;\beta} - \delta g_{;\alpha}{}^{;\alpha}\right) + \frac{1}{S^2}(\delta g)'' + \frac{3S''}{S^3}(\delta g)' + K\frac{2\delta g}{S^3} \\
&, && (4.111)
\end{aligned}$$

such that $\delta g = \delta g_a{}^a$; and the covariant derivative is defined on, and with respect to, the three-dimensional spatial metric. Note also that in the course of the above derivations, all higher order terms in δf_{ab} have been dropped, as only linear variations are being considered here. One can also calculate the following derivative for the the spatial components of the metric tensor:

$$\begin{aligned}
d_{\alpha\beta} &= \frac{d}{dt}g_{\alpha\beta} \\
&= \dot{g_{\alpha\beta}} \\
&= \frac{2}{S^2}\frac{dS}{d\tau}g_{\alpha\beta} \\
&= \frac{2S'}{S^2}g_{\alpha\beta} \quad . && (4.112)
\end{aligned}$$

Similarly one obtains:

$$d_\alpha^\beta = \frac{2S'}{S^2}\delta_\alpha^\beta \quad . \tag{4.113}$$

Matter description

For the sake of simplicity, in cosmological models it is often assumed that the matter conforms to a perfect fluid description, thus yielding an energy-momentum tensor with components:

$$T_\alpha^\beta = p\delta_\alpha^\beta \tag{4.114}$$
$$T_\alpha^0 = 0 \tag{4.115}$$
$$T_0^0 = -\rho \quad . \tag{4.116}$$

It will be useful in later developments (Parts II and III) to have the variation of the perfect fluid tensor. The result follows easily from 4.45:

$$\delta T_a^b = (p+\rho)(u_a\delta u^b + u^b\delta u_a) + (\delta p + \delta\rho)u_a u^b + \delta p \delta_a^b \quad , \tag{4.117}$$

such that the relationship between the components of δu_a are found from the variation of the quantity $g^{ab}u_a u_b = -1$:

$$\delta g^{ab}u_a u_b + g^{ab}(u_a\delta u_b + u_b\delta u_a) = 0 \quad . \tag{4.118}$$

This will be elaborated upon later. Bearing the above in mind, it is important to note that the contravariant and covariant variations will differ in form: this is a major area of confusion encountered in the literature, where different components are used. As a result of this difference, it is preferable to select a particular form, and then maintain that throughout the subsequent calculations. This will be observed later.

4.6.3 Dynamics of The Field Equations for the FLRW Universes

Assuming a perfect fluid form as motivated above, one can now solve directly for the pressure p and energy density ρ using the above via the Einstein Field Equations. For the FLRW models the field equations simplify to the *Raychaudhuri* and *Friedmann* equations (in geometrised units) respectively:

$$3\frac{\ddot{S}}{S} + \frac{1}{2}(\rho + 3p) - \Lambda = 0 \tag{4.119}$$
$$3\dot{S}^2 - \rho S^2 - \Lambda S^2 = K \tag{4.120}$$

where K is a constant of integration. From which one solves for p and ρ :

$$p = \frac{\frac{1}{3}K - 2S\ddot{S} - \dot{S}^2}{S^2} + \Lambda \tag{4.121}$$
$$\rho = \frac{3\dot{S}^2 - K}{S^2} - \Lambda \quad . \tag{4.122}$$

With the above assumptions the energy conservation equation becomes:

$$\dot{\rho} = -3(\rho + p)\frac{\dot{S}}{S} \quad . \tag{4.123}$$

One can then easily solve the above equations, once an equation of state has been specified. To illustrate this, a simple example will now be considered: a closed universe with two separate equations of state $p = 0$ and $p = \frac{1}{3}\rho$. For simplicity the Cosmological constant Λ will be set equal to zero.

Following Hawking [6], one obtains the results:

Case $p = 0$

$$\rho = \frac{M}{S^3} \tag{4.124}$$
$$0 = \frac{3}{M}\frac{\ddot{S}}{S} - \frac{1}{2S^3} \tag{4.125}$$
$$E = \frac{3}{M}(\dot{S})^2 - \frac{1}{S} \quad , \tag{4.126}$$
$$\tag{4.127}$$

where M and E are constants of integration, and differentiation is with respect to t. The solutions to these equations are:

$$E > 0 \; : \; S = \frac{1}{2E}[\cosh(\tau\sqrt{\frac{EM}{3}}) - 1] \tag{4.128}$$
$$t = \frac{1}{2E}[\sqrt{\frac{3}{EM}}\sinh(\tau\sqrt{\frac{EM}{3}}) - \tau] \tag{4.129}$$
$$E = 0 \; : \; S = \frac{M}{12}\tau^2 \tag{4.130}$$
$$t = \frac{M}{36}\tau^2 \tag{4.131}$$
$$E < 0 \; : \; S = \frac{-1}{2E}[1 - \cos(\tau\sqrt{\frac{-EM}{3}})] \tag{4.132}$$
$$t = \frac{-1}{2E}[\tau - \sqrt{\frac{-3}{EM}}\sin(\tau\sqrt{\frac{-EM}{3}})] \quad , \tag{4.133}$$

which, when $E \neq 0$, one can perform the normalisation: $M = \frac{3}{|E|}$.

Case $p = \frac{1}{3}\rho$

$$\rho = \frac{M}{S^4} \tag{4.134}$$
$$-\rho = 3\frac{\ddot{S}}{S} \tag{4.135}$$
$$E = 3\frac{\dot{S}^2}{M} - \frac{1}{S^2} \quad , \tag{4.136}$$

yielding:

$$E > 0 \quad : \quad S = \frac{1}{E}\sinh\tau \tag{4.137}$$

$$t = \frac{1}{E}[\cosh\tau - 1] \tag{4.138}$$

$$E = 0 \quad : \quad S \ \tau \tag{4.139}$$

$$t = \frac{1}{2}\tau^2 \tag{4.140}$$

$$E < 0 \quad : \quad S = -\frac{1}{E}\sin\tau \tag{4.141}$$

$$t = \frac{1}{E}[\cos\tau - 1] \quad , \tag{4.142}$$

$$\tag{4.143}$$

where the additional normalisation $\frac{M}{3} = 1$ has been imposed. In both the above cases the constant of integration E can be interpreted as the total kinetic plus potential energy; depending on whether this quantity is positive, negative or zero, different equations result for S and t. Note that for the open model universe one can derive the following form of the Robertson-Walker metric:

$$d\eta^2 = S(\tau)^2\left[-d\tau^2 + d\chi^2 + \sinh^2\chi(\sin^2\theta d\phi^2 + d\theta^2)\right] \quad . \tag{4.144}$$

Following Lifshitz and Khalatnikov [4], this form of the metric can be formally obtained from the above closed model metric via the transformations:

$$\tau \to i\tau \tag{4.145}$$

$$\chi \to i\chi \tag{4.146}$$

$$S \to iS \quad . \tag{4.147}$$

Consequently, all the equations for the open model can be obtained from those of the closed model via these transformations.

4.7 Bibliography

- SW Hawking and GFR Ellis (1973) *The Large Scale Structure of Space-Time* (Cambridge)
- RM Wald (1984) *General Relativity* (Chicago)
- S Weinberg (1972) *Gravitation and Cosmology* (Wiley)
- CW Misner, Ks Thorne and JA Wheeler (1973) *Gravitation* (Freeman)
- W Rindler (1969) *Essential Relativity* (van Nostrand Reinhold)
- P Coles and F Lucchin (1995) *Cosmology* (Wiley)

- AS Eddington (1920) *Space-Time and Gravitation* (Cambridge)
- H Stephani (1982) *General Relativity* (Cambridge)

Part II

Cosmological Perturbation Theory

"But the concept of a physical law is that it enables us to treat a system that is being studied as one particular case among many. The universe, however, is a unique system, and so this aspect of physical law seems to lose all meaning when we attempt to apply it to the universe."

W.H. McCrae

Cosmological perturbation Theory has its origins in the study of gravitational stability; that is, the study of the physical and mathematical stability of the solutions to the Einstein Field Equations. In Cosmological contexts, the physical motivation for this stability analysis amounts to the desire for the origins of structure formation, such as galaxy evolution, from initial 'perturbations' to an idealised Cosmological model.

As will be seen in the ensuing chapters, there are several ways of formulating and interpreting such studies. The principal approaches covered here will be the metric and covariant formalisms. The former concentrates on the form, and perturbations to, the metric tensor; while the latter focuses on a covariant formulation of relevant perturbed physical quantities. Both formalisms utilise the various attributes of gauge-dependence inherent in such geometric manifold-related formulations; consequently, a detailed treatment of the gauge issue will be given in both contexts.

The bulk of the material in this section will be taken from seminal papers produced in the field of cosmological perturbation theory over the past half-century. The intent behind sectioning the chapters according to these papers is to provide a chronological progression of historic developments in the field, contextualising the results, and in the process maintain a sense of scientific continuity. However,it should be emphasised here, as mentioned in the introduction, that the material of this section is restricted to the period 1960 – 1993, and that the subsequent years have witnessed an increase in interest in *observational* Cosmology in the light of improved observational surveys and data pertaining to large-scale structure and the like.

Chapter 5

Metric Perturbation Theory

"It has become increasingly evident, however, that Nature works on a different plan. Her fundamental laws do not govern the world as it appears in our picture in any direct way, but instead they control a substratum of which we cannot form a mental picture without introducing irrelevancies."

P.A.M. Dirac

5.1 Motivation

Observational evidence to date seems to substantiate the assumption that the universe is *almost* isotropic and spatially homogeneous on the large s cale. Consequently, as the principal aim of perturbation theory is to determine structural growth in terms of small variations from an idealised mathematical universe model, it follows naturally that the class of Friedman-Robertson-Walker (FRW) models be chosen to fulfill such a capacity. Hence one would perturb FRW calculated quantities to obtain the corresponding real, physically observable quantities in the real 'lumpy' universe. Henceforth in this section the FRW assumption will be maintained in the underlying analysis. One refers to such an assumed idealised structure as a *background model*, and it is envisioned as a mathematical template against which to study observational evidence concerning the structure of the universe.

Having thus established a mathematical framework within which to operate, one can proceed to the consideration of an exact formulation of a general perturbed quantity. In the approach of this section, the perturbation formulation is considered strictly within the geometrical description of a universe model, namely via the metric, as this describes the immediate physical structure of such a mathematical model. One is consequently concerned primarily with an initial perturbation to an FRW background model metric. However, all constituent matter in the universe is written into the geometry thereof via the Einstein Field Equations, effectively coupling the metric, contained in the Einstein tensor, with the matter expressed through the energy-momentum tensor. It follows thus that one should separately perturb the matter from some idealised background form to that of the real universe; for simplicity, the background matter description will norΦrmally be assumed to be that of a perfect fluid.

The approach adopted here will be to formulate a generic metric perturbation theory in terms of fundamental perturbation variables, and to proceed thence to a study of the gauge problem, the underlying difficulty in classical perturbation theory. A number of approaches to the solution of the ensuing perturbation equations will then be highlighted, focusing on several seminal works by prominent researchers in the field over the past thirty years. The literature thus covered will include a number of specific physical applications, all of which ultimately provide useful insight into the nature of structure formation in the universe, this being the primary motivation of perturbation theory.

5.2 Metric Perturbations

The term 'metric perturbations' is used in this approach, as the fundamental object which is 'perturbed' is the metric tensor.

In the subsequent analysis, it will thus be necessary to utilise certain generic decomposition characteristics of vectors and tensors in the spatial part of the manifold before proceeding to the form of the perturbed metric. This will be used, in particular, to classify perturbations as scalar, vector or tensorial in later applications; hence these characteristics will be derived first. The following analysis closely follows that of Stewart.

5.2.1 Vector and Tensor Decomposition

Vectors

Consider an arbitrary scalar field ϕ such that

$$\Delta\phi \equiv \phi^{;\alpha}_{;\alpha} = 0 \quad , \tag{5.1}$$

that is, ϕ is *harmonic*. Here, the covariant derivative is with respect to the spatial metric. Using Gauss' theorem on 5.1, one obtains:

$$\begin{aligned} 0 &= \int_{\mathcal{M}} \phi\Delta\phi d\tau \\ &= \int_{\partial\mathcal{M}} \phi\phi_{;\alpha} dS^{\alpha} - \int_{\mathcal{M}} \phi_{;\alpha}\phi^{;\alpha} d\tau \quad . \end{aligned} \tag{5.2}$$

In the above, the surface term vanishes if $\mathcal{M}$ is a closed, unbounded space; in this case 5.2 would imply that ϕ is constant. For the open and flat case scenarios, the known asymptotic form of harmonic functions together with boundedness allow one to assume that the surface term vanishes: here $\partial\mathcal{M}$ is essentially a 'sphere at infinity'. Hence, if one were now to construct a vector from a scalar field:

$$A^{\alpha} = \phi^{;\alpha} \tag{5.3}$$

it follows from the above that ϕ couldn't be harmonic, as this would yield $A^{\alpha} = 0$. Now let A^{α} be *any* smooth vector field on $\mathcal{M}$, and consider the solutions ϕ to the equation:

$$\Delta\phi = A^{\alpha}_{;\alpha} \quad . \tag{5.4}$$

Such solutions are known to exist, and are unique up to $\phi \to \phi + cst$. Having established this, one can now define the unique, solenoidal vector B^α (i.e $B^\alpha_{;\alpha} = 0$):

$$B^\alpha = A^\alpha - \phi^{;\alpha} \quad . \tag{5.5}$$

Hence one has the unique (non-local) decomposition for an arbitrary vector field A^α:

$$A^\alpha = \phi^{;\alpha} + B^\alpha \quad , \tag{5.6}$$

where B^α is solenoidal, and ϕ is determined, up to a constant, from equation 5.4.

Tensors

Similarly, one would like to form an analogous decomposition for all rank-two tensors. As skew-symmetric tensors are equivalent to vectors (same number of components for three-space) allowing application of the above analysis, one need consider here only symmetric tensors. Associated with such tensors is a natural scalar, the trace, which is usually non-zero; as one wishes to consider a decomposition which will classify all scalar parts of the tensor generically, here only the trace-free part of spatial symmetric tensors $T_{\alpha\beta}$ will be considered, i.e.

$$T^{\alpha\beta} - \frac{1}{3} o^{\alpha\beta} T^\epsilon_\epsilon \quad , \tag{5.7}$$

where $o^{\alpha\beta}$ is the spatial metric, as before. In the course of the subsequent decomposition, *transverse* symmetric tensors will also be considered; namely, tensors $T^{\alpha\beta}$ which satisfy:

$$T^{\alpha\beta}_{\ \ ;\beta} = 0 \quad . \tag{5.8}$$

One starts by defining the following derivative:

$$\mathcal{D}_{\alpha\beta} = \nabla_\alpha \nabla_\beta - o_{\alpha\beta}(\Delta + 2K) \quad , \tag{5.9}$$

where $K = \pm 1, 0$ depending on whether the space is open, closed or flat. This derivative is easily verifiable as transverse when acting on scalars ϕ, by using the Ricci identity; i.e. $(\mathcal{D}^{\alpha\beta}\phi)_{;\beta} = 0$. Having defined this derivative one first looks for the following decomposition of a smooth, symmetric and trace-free tensor $T^{\alpha\beta}$:

$$T^{\alpha\beta} = \mathcal{D}^{\alpha\beta}\omega + 2A^{(\beta;\alpha)} + W^{\alpha\beta} \quad , \tag{5.10}$$

where A^α is an arbitrary vector, and $W^{\alpha\beta}$ is a symmetric, *transverse* and trace-free tensor. The divergence of this equation then yields:

$$(\Delta + 3K)\omega = A^\alpha_{;\alpha} \quad , \tag{5.11}$$

which thus uniquely determines ω. Having established all this, one can now define, as with the vector case:

$$W^{\alpha\beta} = T^{\alpha\beta} - \mathcal{D}^{\alpha\beta}\omega - 2A^{(\beta;\alpha)} \quad , \tag{5.12}$$

which is thus unique, trace-tree and transverse, and yields equation 5.10. Hence, substituting in full into 5.10 for $\mathcal{D}^{\alpha\beta}$; using the decomposition 5.6 for A^α; noting that $(\Delta + 3K)\omega = \Delta\phi$ from 5.4 and 5.11; and defining:

$$\psi \equiv \omega + 2\phi \tag{5.13}$$

$$\Delta_{\alpha\beta} \equiv \nabla_\alpha \nabla_\beta - \frac{1}{3} o_{\alpha\beta} \Delta \quad , \tag{5.14}$$

one obtains the final desired decomposition:

$$T^{\alpha\beta} = \Delta^{\alpha\beta}\psi + 2B^{(\beta;\alpha)} + W^{\alpha\beta} \quad , \tag{5.15}$$

where B^α is solenoidal as before. It transpires that B^α will be unique if $K \leq 0$; otherwise it is defined up to $B^\alpha \rightarrow B^\alpha + k^\alpha$, where k^α is a Killing vector. Hence one can now define:

- *Scalar terms*: terms which are derived from a scalar potential ψ via *linear* operations involving only $o^{\alpha\beta}$ and ∇_α;
- *Vector terms*: terms which are derived from a solenoidal vector B^α via linear operations of only $o^{\alpha\beta}$ and ∇_α;
- *Tensor terms*: tensors which are symmetric, trace-free and transverse.

In all of the preceding analysis, the covariant derivative ∇_α was taken with respect to the spatial part of the metric, namely $o_{\alpha\beta}$. Bearing the above in mind, one can now study the effect of generic metric tensor perturbations by considering separately tensor, vector and scalar perturbations in terms of the above decompositions.

5.2.2 Metric Decomposition

One commences the study by postulating the metric of 'real' space-time g_{ab} in terms of the the background FRW metric f_{ab} and the perturbation δf_{ab} to this metric:

$$g_{ab} = f_{ab} + \delta f_{ab} \quad , \tag{5.16}$$

where the FRW metric f_{ab} is derived from the line element:

$$ds^2 = S(\tau)^2[d\tau^2 - o_{\alpha\beta}dx^\alpha dx^\beta] \tag{5.17}$$

$$o_{\alpha\beta} = \delta_{\alpha\beta}[1 + \frac{K}{4}(x^2 + y^2 + z^2)]^{-1} \quad , \tag{5.18}$$

as in the previous chapter, where $K = -1, 0, 1$ depending on whether the 3−space is open, flat or closed; and x, y and z are the spatial co-ordinates. As shown in the preceding section, a metric can be constructed from scalars, vectors and arbitrary 2−tensors; consequently one can classify the metric perturbation in terms of separate generic scalar, vector and tensor perturbations, a classification which follows from the way in which the fields from which δf_{ab} is constructed transform under spatial co-ordinate transformations on a constant-time hypersurface, as will be shown in due course.

As will subsequently be seen, there are numerous approaches to the formulations of these three basic classifications. In this section, a fairly simplistic, generic and transparent formulation will be given in order to highlight most effectively the relevant features of the theory.

5.2.3 Scalar Perturbations

Scalar perturbations can enter into δf_{ab} through $o_{\alpha\beta}$ either as a multiplicative scalar factor γ, or through a covariant derivative (with respect to $o_{\alpha\beta}$) of some scalar function κ. Similarly one scalar ϕ is required for f_{00}, and a scalar ψ via a covariant derivative for $f_{0\alpha}$. These all yield the following form for δf_{ab}:

$$\delta f_{ab}^{(s)} = S^2 \begin{bmatrix} 2\phi & -\psi_{;\alpha} \\ -\psi_{;\alpha} & 2(\gamma o_{\alpha\beta} - \kappa_{;\alpha\beta}) \end{bmatrix} , \tag{5.19}$$

where the factors of two and negative signs are not crucial, but rather are inserted for later computational convenience. Note that full generality is assumed for the scalars in the above expression in that they are all functions of space *and* time. In terms of the above, it is useful to define the following associated scalar quantities:

$$L = \mathcal{H}' - \frac{1}{S}\psi \tag{5.20}$$

$$K = -\left(3\left(\frac{\dot{\gamma}}{S^2} + 2S\mathcal{H}\phi\right) + L^{;\alpha}_{\;;\alpha}\right) , \tag{5.21}$$

such that the covariant derivative is taken with respect to the spatial part of the background metric, as before. Note that, for any chosen hypersurface foliation, K is the perturbation to the trace of the extrinsic curvature as defined in section 3.5, while L generates the perturbation to the traceless part; this is readily seen in the ADM formalism, as outlined in section 3.5.2.

5.2.4 Vector Perturbations

Vector perturbations can enter δf_{ab} via a vector A_α for $f_{0\alpha}$, and the covariant derivative of a vector B_α with respect to $o_{\alpha\beta}$ for $f_{\alpha\beta}$; where both A_α and B_α are solenoidal as previously motivated, so as to ensure that each vector cannot be split into the sum of a divergenceless vector and the gradient of a scalar, thus ensuring a pure vector perturbation. One can construct $\delta o_{\alpha\beta}$ as:

$$\delta o_{\alpha\beta} = 2B_{(\alpha;\beta)} \tag{5.22}$$

in order to guarantee the necessary symmetry. Hence one has the net vector perturbed metric as:

$$\delta f_{ab}^{(v)} = -S^2 \begin{bmatrix} 0 & -A_\alpha \\ -A_\alpha & 2B_{(\alpha;\beta)} \end{bmatrix} . \tag{5.23}$$

5.2.5 Tensor Perturbations

The tensor perturbation is written in terms of a symmetric , transverse and trace-free three-dimensional tensor $h_{\alpha\beta}$ as follows:

$$\delta f_{ab}^{(m)} = -S^2 \begin{bmatrix} 0 & 0 \\ 0 & h_{\alpha\beta} \end{bmatrix} \quad . \tag{5.24}$$

The properties of symmetric, transverse and trace-free, as before, ensure that $h_{\alpha\beta}$ does not contain parts which transform as scalars or vectors. Combining the scalar, vector and tensor perturbed metrics, one obtains the full, generic metric perturbation thus as follows:

$$\delta f_{ab} = S^2(\tau) \begin{bmatrix} 2\phi & -\psi_{;\alpha} - A_\alpha \\ -\psi_{;\alpha} - A_\alpha & 2(\gamma o_{\alpha\beta} - \kappa_{;\alpha\beta}) + 2B_{(\alpha;\beta)} + h_{\alpha\beta} \end{bmatrix} \quad . \tag{5.25}$$

Counting thus the total number of independent functions in the full metric perturbation 5.25, one obtains ten, which in turn equals the total number of independent components of f_{ab}, as one would expect. It transpires that both vector and tensor perturbations exhibit no instability (see [4]): vector perturbations decay kinematically in an expanding universe while tensor perturbations generate gravitational waves which do not couple significantly to energy density and pressure inhomogeneities. On the other hand, scalar perturbations may generate inhomogeneities which affect the dynamical behaviour of matter, and thus ultimately lead to the growth of structure in the universe. Hence in the subsequent perturbational theory analysis the focus will be primarily on scalar perturbations. As will be motivated in detail later, the above perturbation scalar, vector and tensor components in the metric can be expanded as the product of a time-dependent 'amplitude' and spatially-dependent harmonic functions; the details behind harmonic decomposition are provided in appendix B.

The Perturbed Velocity

Now that the complete form of the perturbed metric has been obtained, one can derive an expression for the perturbation δu^a to the four velocity as follows; considering the variation of the quantity $g_{ab}u^a u^b = -1$, one obtains:

$$\delta g_{ab} u^a u^b + g_{ab}(u^a \delta u^b + u^b \delta u^a) = 0 \quad . \tag{5.26}$$

Setting $a = b = 0$ in this expression, one obtains:

$$\delta u^0 = -\frac{\delta f_{00} u^0}{2 g_{00}} \quad , \tag{5.27}$$

which, upon substituting in for the FRW metric 00 components and co-moving velocity, together with 5.25, one obtains:

$$\delta u^0 = -\frac{\phi}{S} \quad . \tag{5.28}$$

For the three-velocity perturbation δu^{α} however, one needs the matter description via the energy density tensor, as δu^{α} would describe the perturbation th the fluid flow. This will be seen more explicitly later when a perfect fluid matter description is assumed.

5.2.6 The Perturbed Field Equations

Now that the form of the perturbed metric has been established, one can formulate the perturbed field equations. Taylor expanding both sides of the field equations, one has:

$$
\begin{aligned}
G^a_b &= G^{(0)a}_b + \delta G^a_b + \dots && (5.29)\\
T^a_b &= T^{(0)a}_b + \delta T^a_b + \dots \quad , && (5.30)
\end{aligned}
$$

where, as before, the 'zero' superscript indicates the relevant background quantities which are calculated as in equations 4.103, 4.104 and 4.105. For scalar perturbations, as these will be considered predominantly later, the perturbations to the field equations defined by the above through:

$$\delta G^a_b = \delta T^a_b \tag{5.31}$$

become, after some straight-forward though tedious calculations (see [17]):

$$
\begin{aligned}
\delta G^0_0 &= \frac{2}{S^2}\left\{-3\mathcal{H}(\mathcal{H}\phi + \gamma') + \nabla^2\left[\gamma - \mathcal{H}(\Psi - \kappa')\right] + 3K\gamma\right\}\\
&= 8\pi G \delta T^0_0 && (5.32)\\
\delta G^0_\alpha &= \frac{2}{S^2}\left[\mathcal{H}\phi + \gamma' - K(\psi - \kappa')\right]_{;\alpha}\\
&= 8\pi G T^0_\alpha && (5.33)\\
\delta G^\alpha_\beta &= -\frac{2}{S^2}\left\{\left[(2\mathcal{H}' + \mathcal{H}^2)\phi + \mathcal{H}\phi' + \gamma'' + 2\mathcal{H}\gamma' - K\gamma + \frac{1}{2}\nabla^2 D\right]\delta^\alpha_\beta - \frac{1}{2}D^{;\alpha}_{;\beta}\right\}\\
&= 8\pi G \delta T^\alpha_\beta \quad , && (5.34)
\end{aligned}
$$

such that:

$$D = \phi - \gamma + 2\mathcal{H}(\psi - \kappa') + (\psi - \kappa')' \quad . \tag{5.35}$$

The form of δT^a_b depends on the matter description assumption. The gauge-invariant form of these equations will be derived later.

5.3 The Gauge Problem

Following Mukhanov *et al.* [17], one can describe gauge transformations via two different approaches: an *active* and a *passive* approach. Starting with the passive approach, one considers a single *physical manifold* $\mathcal{M}$, and chooses a set of co-ordinates x^a thereon. A *background model* is then defined by assigning to all functions $F(x^a)$ (scalar or tensor) on $\mathcal{M}$ a previously given value $F^{(0)}(x^a)$. Note that both these functions are fixed functions of the co-ordinates and thus not geometrical quantities. The perturbation of $\delta F(p)$ at the point p is then defined as:

$$\delta F^{(x)}(p) = F\left(x^a(p)\right) - F^{(0)}\left(x^a(p)\right) \quad . \tag{5.36}$$

Consequently, in a new co-ordinate system y^a, but at the *same point* p one would have:

$$\delta F^{(y)}(p) = F\left(y^a(p)\right) - F^{(0)}\left(y^a(p)\right) \quad . \tag{5.37}$$

One can then make the following definition:

"The transformation $\delta F^{(x)}(p) \rightarrow \delta F^{(y)}(p)$ is called the gauge transformation associated with the co-ordinate system change $x^a(p) \rightarrow y^a(p)$ on the manifold $\mathcal{M}$".

In the *active* approach one instead considers two separate manifolds: a physical manifold $\mathcal{M}$, and a background space time $\mathcal{N}$ on which co-ordinates $x^a_{(b)}$ are rigidly fixed (the index b stands for "background"). The introduction of a diffeomorphism $\mathcal{D}: \mathcal{N} \rightarrow \mathcal{M}$ then naturally induces a set of co-ordinates on $\mathcal{M}$ via $\mathcal{D}: x^a_{(b)} \rightarrow x^a$. For such a chosen diffeomorphism $\mathcal{D}$ one can define a perturbation δF of an arbitrary function F, defined on $\mathcal{M}$ at point p, as:

$$\delta F^{(D)}(p) = F^{(D)}(p) - F^{(0)}\left(\mathcal{D}^{-1}(p)\right) \quad , \tag{5.38}$$

such that $F^{(0)}$ is a fixed function defined on $\mathcal{N}$. If one now considers a different diffeomorphism $\mathcal{E}$, which thus induces a different set of co-ordinates on $\mathcal{M}$, then the following perturbation transpires at p:

$$\delta F^{(E)}(p) = F^{(E)}(p) - F^{(0)}\left(\mathcal{E}^{-1}(p)\right) \quad . \tag{5.39}$$

Hence one defines the gauge transformation, as before, as $\delta F^{(D)}(p) \rightarrow \delta F^{(E)}(p)$; this being generated by the change of diffeomorphism correspondence $\mathcal{D} \rightarrow \mathcal{E}$ between the manifolds $\mathcal{M}$ and $\mathcal{N}$. Stated simply, one can then say that a gauge transformation changes the point in the background space-time corresponding to a point in the physical space-time. This is precisely the *gauge dependence* inherent in perturbation theory, as will be seen subsequently.

Both the active and passive approaches can be seen to be equivalent. The passive approach allows one to relate choice of gauge with different co-ordinate systems, while the second permits the study of perturbation amplitudes in terms of background and physical manifold correspondence. In the next chapter covering covariant perturbation theory, a much more lucid definition of the gauge problem (see equation 6.57) will be given which firstly doesn't directly involve a computation between two quantities in two different manifolds (equation 6.57); and secondly, provides a clear and unambiguous test for gauge-invariance (see equation 6.59). This will certainly be seen as a motivational factor for the covariant approach over the metric approach to Perturbation Theory.

To comprehend the nature of gauge transformations more fully, consider the following infinitesimal transformation:

$$x^a \rightarrow \overline{x}^a = x^a + \zeta^a \quad , \tag{5.40}$$

such that ζ^a is a infinitesimal co-ordinate change. This results in the following change in an arbitrary co-ordinate-dependent quantity F:

$$\Delta F = \delta\overline{F} - \delta F = \mathcal{L}_\zeta F \quad , \tag{5.41}$$

such that $\mathcal{L}_\zeta$ is the Lie derivative (section 1.2.1) in the direction of ζ^a. The transformations thus formed by 5.40 and 5.41 form a group structure of infinitesimal transformations: *the gauge group of gravitation.* As any three-vector can be decomposed into a potential-defined and divergenceless part as previously derived, one can express ζ^a in the following form:

$$\begin{aligned} \zeta^a &= (\zeta^0, \zeta^\alpha) & (5.42)\\ &= (\zeta^0, \zeta^\alpha_{(sol)} + o^{\alpha\beta}\xi_{;\beta}) \quad , & (5.43) \end{aligned}$$

such that ξ is a solution to the equation $\xi^{;\alpha}_{;\alpha} = \zeta^\alpha_{;\alpha}$, and $\zeta^\alpha_{(sol)}$ is solenoidal, as before, with the covariant derivative being taken with respect to the background spatial co-ordinates. Hence it follows that $\zeta^\alpha_{(sol)}$ will only affect the vector, and not scalar perturbations. This then naturally implies that the scalar metric perturbations will be affected only by ζ^0 and ξ. The most general diffeomorphism which thus preserves the scalar nature of metric perturbations will be:

$$\begin{aligned} \tau \rightarrow \overline{\tau} &= \tau + \zeta^0(\tau, \mathbf{x}) \\ x^\alpha \rightarrow \overline{x}^\alpha &= x^\alpha + o^{\alpha\beta}\xi_{;\beta}(\tau, \mathbf{x}) \quad . \end{aligned} \tag{5.44}$$

Under this generalised infinitesimal transformation one can calculate the associated transformation of the metric, using the notation of section 5.2:

$$\delta f_{ab} \rightarrow \delta\overline{f_{ab}} = \delta f_{ab} + \Delta f_{ab} \quad . \tag{5.45}$$

Note that one *cannot* call this a gauge transformation, as the original stipulation for arbitrary functions on the manifold was co-ordinate functional invariance: this clearly does not apply to the metric as this is a geometric quantity. However, one would none-the-less like to express metric perturbations in terms of gauge-invariant quantities, a process which will facilitate calculations and physical interpretation of associated physical quantities. To this end one first calculates the transformations of the associated scalar metric perturbations from 5.19 under 5.45:

$$\begin{aligned} \overline{\phi} &= \phi - \left(\frac{S'}{S}\right)\zeta^0 - \zeta^{0\prime} & (5.46)\\ \overline{\gamma} &= \gamma + \left(\frac{S'}{S}\right)\zeta^0 & (5.47)\\ \overline{\psi} &= \psi + \zeta^0 - \xi' & (5.48)\\ \overline{\kappa} &= \kappa - \xi \quad . & (5.49) \end{aligned}$$

Now, by taking various functional combinations of ϕ, γ, ψ and κ one can formulate gauge-*independent* variables; that is quantities which preserve their functional form under the above infinitesimal transformation. The simplest such linear functions are constructed from the following two linearly independent quantities (see [17]):

$$A = \phi + \left(\frac{1}{S}\right)[(\psi - \kappa')S]' \tag{5.50}$$

$$B = \gamma - \left(\frac{S'}{S}\right)(\psi - \kappa') \quad . \tag{5.51}$$

These are equal to the negative of the Bardeen [10] variables, which will be seen later. These quantities play a rôle analogous to the electric and magnetic fields in electromagnetism. Naturally, any linear combination of the above variables will also be gauge-invariant. As will be seen later, after formulating the matter components in terms of the perturbation variables via the perturbed field equations, one can analogously establish gauge invariant quantities for these; this will be developed later.

At this point, returning to equation 5.19 it is useful to note a number of specific gauge choices which are most frequently used in perturbation theory. The gauge freedom here essentially amounts to the arbitrariness of the perturbation variables: one can at will set some of these, or specific combinations of them, equal to zero.

5.4 Choice of Gauge

For the practicalities of calculation, it is convenient to consider a gauge transformation as a 'mapping' from the background model into the physical universe. As such, following Ellis and Bruni [13], such a map can be considered to comprise four parts:

- The definition of a separate family of world lines in the physical and background universes respectively;
- The definition of a correspondence between these two families;
- The definition of a separate family of space-like hypersurfaces in the physical and background universes respectively;
- The definition of a correspondence between these two families of hypersurfaces.

Consequently, the choice of a specific gauge amounts to the individual specification of the above properties; this is referred to as *gauge fixing*. In the following, a number of specific such choices will be described. Practically, this amounts to imposing various constraints on the perturbation variables defined through the perturbed metric of 5.25.

5.4.1 The Synchronous Gauge

This gauge is defined through the constraints:

$$\delta f_{00} = 0 \quad (5.52)$$
$$\delta f_{0\alpha} = 0 \ , \quad (5.53)$$

which is thus equivalent to setting:

$$\phi = \psi = 0 \quad (5.54)$$

in 5.19; and amounts to defining equivalent proper times in the background and physical models. Hence, starting with an arbitrary initial co-ordinate system one can perform an infinitesimal transformation to a *synchronous-gauge* co-ordinate system. This entails setting $\overline{\phi} = \overline{\psi} = 0$, and solving for ζ^a from 5.46 and 5.48. This yields:

$$\zeta^0 = \frac{1}{S}\int S\phi \quad (5.55)$$
$$\zeta^\alpha = o^{\alpha\beta}\left(\int \psi d\tau + \int \frac{1}{S}(\int S\phi d\tau)d\tau\right)_{;\beta} \ . \quad (5.56)$$

The name 'synchronous gauge' arises from the fact that in this gauge, from 5.19 there are no perturbations to the time-like components Note, however, in the above synchronous gauge the residual co-ordinate freedom due to the integration constants; if one labels these constants $C_1(x)$, $C_2(x)$, then the synchronous gauge conditions are maintained under the additional set of transformations ([4]):

$$\tau \rightarrow \overline{\tau} + \frac{1}{S}C_1(x) \quad (5.57)$$
$$x^\alpha \rightarrow \overline{x^\alpha} + o^{\alpha\beta}\left(C_{1;\beta}(x)\int \frac{1}{S}d\tau + C_{2;\beta}(x)\right) \ . \quad (5.58)$$

This additional gauge freedom leads to the appearance of unphysical gauge modes which result in difficulty in the interpretation of synchronous gauge calculations; it essentially amounts to a freedom in choice of initial constant time hypersurface. This is a fundamental problem in classical perturbation theory which makes extensive use of the synchronous gauge, and it will be seen later that careful consideration needs to be paid to gauge freedom in general when considering the physicality of obtained results, especially those derived within a specific fixed gauge. However, as Bardeen [12] notes, in an expanding background the synchronous gauge is always mathematically well-behaved, with no spurious co-ordinate singularities.

Another problem is that, to define the perturbation of an arbitrary physical quantity such as the energy density, and determine its present day value, one has to integrate the field equations all the way back to the initial singularity: this will be seen later.

5.4.2 The Longitudinal Gauge

The longitudinal or conformal-Newtonian gauge is defined by imposing the following constraints on the perturbed metric:

$$\delta f_{0\alpha} = 0 \quad (5.59)$$
$$\delta f_{\alpha\beta} \propto o_{\alpha\beta} \quad . \quad (5.60)$$

In equation 5.19 this is equivalent to setting $\psi = \kappa = 0$. Hence one obtains, in the same way as for the synchronous gauge:

$$\zeta^0 = \kappa' - \psi \quad (5.61)$$
$$\zeta^\alpha = o^{\alpha\beta}\kappa_{;\beta} \quad , \quad (5.62)$$

yielding a metric of the form:

$$ds^2 = S^2(\tau)\left[(1+2A)d\tau^2 - (1-2B)o_{\alpha\beta}dx^\alpha dx^\beta\right] \quad , \quad (5.63)$$

where A, B are the gauge-invariant variables as defined before. This form explains the name 'conformal-Newtonian' gauge; as, when the spatial part of the energy-momentum tensor is diagonal ($\delta T^\alpha_\beta \sim \delta^\alpha_\beta$), one has $A = B$, thereby having A as the equivalent of the Newtonian gravitation potential. Note that there is no residual co-ordinate freedom in the longitudinal gauge, as there is for the synchronous gauge; hence the co-ordinates remain fixed.

5.4.3 The Uniform Expansion Gauge

This is obtained by setting the quantity K, as defined in 5.21, equal to zero, which then results in perturbations which are determined purely in terms of elliptic equations. Hence this amounts to equating the expansion parameters in the background and physical models, which thus defines space-like surfaces in the background model: surfaces of constant expansion. This particular gauge corresponds to a Newtonian gauge when the perturbation is well within the particle horizon, and in which the co-ordinate velocities become Newtonian peculiar velocities.

5.4.4 The Zero-Shear Gauge

The zero-shear gauge is defined by setting the quantity L, as defined in 5.20, equal to zero. This is referred to as the zero-*shear* gauge in that the traceless part of the extrinsic curvature, the perturbation to which is defined through L, is indeed the shear of the chosen hypersurface normals (*cf.* section 4.5.2). This gauge is also a Newtonian gauge when the perturbation is well within the particle horizon.

5.4.5 The Co-Moving Gauge

This amounts to choosing surfaces of constant time as surfaces orthogonal to the fluid flow; this will be seen to be equivalent to setting the perturbation to the energy flux equal to zero, which is only possible if there is no fluid vorticity.

From the above gauge examples the meaning of 'gauge transformation' becomes more lucid: roughly speaking, one can regard a gauge transformation as a particular set of co-ordinate transformations which form a particular algebraic group; for example the group of all spatial rotation s. Consequently, choosing a *particular* gauge amounts to specifying such a group of transformations, the group in question being referred to as the gauge. One is now in a position to study various cosmological models and scenarios within different gauge choices.

5.5 Direct Metric Perturbation

It naturally follows that one should begin with a study which assumes the fairly generic isotropic background models with a relatively simple and convenient equation of state. Hence, in investigating the gravitational stability of the isotropic model, Some of the pioneering work, particularly that of Lifshitz and Khalatnikov [4], paid close attention to closed models with equations of state $p = 0$ and $p = \frac{\mu}{3}$. Also, Lifshitz in 1946 [1] and later Lifshitz and Khalatnikov [4] in their seminal work of 1963 sought to utilise the notion of four-dimensional spherical harmonics in the perturbed metric form. As will be seen, this idea rather simplifies the calculations and provides a clear interpretation of the results. However, the one drawback is the assumption of the synchronous gauge throughout. Some details of the calculations, results and interpretations will be given here in order to provide an understanding of the application of the foregoing theory, as well as to generate an appreciation for the underlying physical relevance and motivation.

The following formulation then closely follows that of Lifshitz and Khalatnikov [4], wherein, for simplicity, no cosmological constant is assumed. One starts with the basic Robertson-Walker metric for the closed model:

$$d\eta^2 = dt^2 - S(t)^2 \left[d\chi^2 + \sin^2\chi(\sin^2\theta d\phi^2 + d\theta^2)\right] \quad . \tag{5.64}$$

Making the transformation to conformal time τ via

$$dt = S(t)d\tau \quad , \tag{5.65}$$

one thus obtains:

$$d\eta^2 = -S(\tau)^2 \left[-d\tau^2 + d\chi^2 + \sin^2\chi(\sin^2\theta d\phi^2 + d\theta^2)\right] \quad . \tag{5.66}$$

A perturbation of the isotropic model is described here by the changes (i.e. a variation) in the metric tensor, δg_{ab}, in the four-velocity of matter, δu^a, and in the energy-density $\delta\rho$. Apropos of this, the notation of the previous section will be used:

$$g_{ab} = f_{ab} + \delta f_{ab} \quad . \tag{5.67}$$

Now, in this linear approximation, one will thus have the small perturbations described in terms of the variations of section 3.1.2, which would thus have to satisfy equation 5.154.

One can now make use of the gauge freedom to impose a particular gauge for computational convenience. This, however, comes at a price: one has to keep track of the gauge choice to the extent of identifying and eliminating spurious co-ordinate gauge modes which are unphysical, and removable by a suitable change of co-ordinates; this will be considered later. In this context the simplest gauge choice would be the *synchronous gauge*, which *ipso facto* imposes the four constraints on f_{ab} contained in 5.52 and 5.53. This results in the system no longer being co-moving; i.e.

$$\delta u^{\alpha} \neq 0 \quad . \tag{5.68}$$

However, using 5.52, 5.53 and equation 4.118, one obtains:

$$\delta u^{0} = 0 \quad . \tag{5.69}$$

Focusing first on the variation to the energy-momentum equations, and assuming a perfect fluid, one has from equation 4.117 and the synchronous gauge assumption the following perturbed quantities:

$$\begin{aligned} \delta T^{\beta}_{\alpha} &= \delta^{\beta}_{\alpha}\delta p & (5.70)\\ \delta T^{\beta}_{0} &= -S(\rho + p)\delta u^{\beta} & (5.71)\\ \delta T^{0}_{0} &= -\delta\rho \quad . & (5.72) \end{aligned}$$

Due the 'smallness' assumption of the perturbations, and the assumption of an equation of state $p = p(\rho)$, one can write:

$$\delta p = \frac{dp}{d\rho}\delta\rho \quad , \tag{5.73}$$

which yields, from the above perturbed matter components:

$$\delta T^{\beta}_{\alpha} = -\delta^{\beta}_{\alpha}\frac{dp}{d\rho}\delta T^{0}_{0} \quad , \tag{5.74}$$

which thus enables one to introduce and utilise the equation of state to the subsequent analysis. Returning now to the Einstein tensor, one has the Ricci tensor variations as in section 3.1.2 through equations 4.108, 4.109, 4.110, and 4.111; by replacing g therein with f to obtain consistent notation. Hence, combining these variations for the energy momentum and Ricci components via 5.154, one obtains from equation 5.74 the equations for the perturbation to the metric tensor:

$$\left(\delta f_{\alpha\ ;\gamma}^{\ \gamma;\beta} + \delta f^{\beta\ ;\gamma}_{\ \gamma;\alpha} - \delta f_{\alpha\ \gamma}^{\ \beta;\gamma} - \delta f_{;\alpha}^{\ ;\beta}\right) + \delta f^{\beta\ \prime\prime}_{\alpha} + \frac{2S'}{S}\delta f_{\alpha}^{\ \beta} \mp \delta f_{\alpha}^{\ \beta} = 0 \quad \alpha \neq \beta \tag{5.75}$$

and:

$$\begin{aligned} &\frac{1}{2}\left(\delta f_{\ \alpha}^{\ \alpha} - \delta f_{\alpha\ ;\alpha}^{\ \alpha;\alpha}\right) - \delta f'' - 2\frac{S'}{S}\delta f' \pm \delta f \\ =\ & 3\frac{dp}{d\rho}\left[\frac{1}{2}\left(\delta f_{\alpha\ ;\alpha}^{\ \alpha;\alpha} - \delta f_{;\alpha}^{\ ;\alpha}\right) + \frac{S'}{S}\delta f' \mp \delta f\right] \quad , \qquad (5.76) \end{aligned}$$

where, as before, a primed sign indicates differentiation with respect to conformal time. These are the equations fundamental to the subsequent analysis which essentially comprises i) assuming the general functional form of separately scalar, vector and tensor perturbations to $f_{\alpha\beta}$ as in the foregoing sections, considering each case individually, solving for this general functional form through substitution into 5.75 and 5.76; and ii) applying the results to various different FRW models (i.e. open, closed, flat universes etc.) each with arbitrary equations of state (i.e. dust, radiation etc.). Once this has been done one desires to know the behaviour of the density and velocity perturbations, as stated earlier in order to obtain insight into the possible growth of structure with time in the universe. To this end it is useful to formulate the relative density perturbation $\frac{\delta\rho}{\rho}$ directly using equations 5.154, 5.72, 4.109, 4.111 and 4.122:

$$\frac{\delta\rho}{\rho} = \frac{S^2}{6(S'^2 \pm S^2)} \left(\delta f_{\alpha\ ;\beta}^{\ \beta;\alpha} - \delta f_{;\alpha}^{\ ;\alpha} + \frac{2S'}{S}\delta f' \mp 2\delta f \right) \quad . \tag{5.77}$$

Similarly, the change in velocity δu^α can be determined directly from equations 5.71, 4.121, 5.154 and 4.110:

$$\delta v^\alpha \equiv S\delta u^\alpha = \frac{S^2}{4(\pm S^2 + 2S'^2 - SS'')} (\delta f^{;\alpha} - \delta f_{\beta}^{\ \alpha;\beta})' \quad , \tag{5.78}$$

where v^α is the *physical* velocity. At this point one notices the inherent gauge-dependence of the formalis m, in this instance related to the synchronous gauge assumption. Here, the gauge-dependence manifests itself simply in the form of a subset of solutions to 5.75 and 5.76 which merely amount to co-ordinate transformations; i.e. their associated perturbations can be *removed* through an appropriate co-ordinate transformation, as mentioned earlier, and are thus not physical. This was shown in section 5.4.1. For the synchronous gauge the non-physical gauge modes that result thus transpire as, as in section 5.4.1:

$$\delta f^\alpha_\beta = c_{0;\beta}^{\ ;\alpha} \int \frac{d\tau}{S} + \frac{S'}{S^2} c_0 \delta^\alpha_\beta + (c_\beta^{\ ;\alpha} + c^\alpha_{\ ;\beta}) \tag{5.79}$$

associated with the infinitesimal transformation:

$$x^a \rightarrow x^a + c^a \quad . \tag{5.80}$$

One can now proceed with the analysis by considering separately scalar, vector and tensor perturbations, as done in section 5.2; however, for mathematical convenience, one can opt here for a representation in terms of the scalar, vector and tensor harmonics. This is motivated by the following geometrical argument.

One can reason that, as the metric of a four-space with constant positive curvature corresponds to the geometry of the surface of a hypersphere in four-space, one should be able to expand any arbitrary perturbation in terms of the four-dimensional spherical harmonics and their derivatives. Similarly for a space of constant *negative* curvature one can utilise the 'pseudo-harmonics' of a 'pseudosphere' (sphere with imaginary radius). The associated scalar, vector and tensor harmonics $\mathcal{Q}$, $\mathcal{S}_\alpha$ and $\mathcal{G}_\alpha^{\ \beta}$ and their derivatives $\mathcal{Q}_\alpha^{\ \beta}$ and $\mathcal{Q}_\alpha^{(t)\ \beta}$ are as defined in appendix B. By assuming this form one obtains solutions which are *ipso facto* Fourier decomposed. This gives an infinite sequence of periodic terms, most of the terms of which can be neglected on physical grounds (e.g. wavelengths exceeding that of the Hubble horizon). The analysis then centres around the magnitude of τ and the order of the various harmonics. Recalling the assumption of a synchronous gauge, the three separate ansätze for f_α^β that are assumed are:

Case I: Scalar Perturbations

One thus begins by reformulating the scalar perturbation variables κ, γ in 5.19 in terms of scalar harmonic functions, as motivated above; this can be done together with the following transformation in terms of two new scalar variables $a(\tau)$, $b(\tau)$ defined as follows:

$$\gamma \equiv \frac{1}{6S^2}(a+b)Q \tag{5.81}$$
$$\kappa^{;\alpha}_{;\beta} \equiv \frac{a}{2S^2(n^2 \mp 1)} Q^{;\alpha}_{;\beta} \quad , \tag{5.82}$$

which has the purpose of simplifying the form of the perturbed metric in the synchronous gauge to the following:

$$\delta f_\alpha^\beta = a(\tau)Q_\alpha^\beta + b(\tau)Q_\alpha^{(t)\,\beta} \tag{5.83}$$
$$\delta f = bQ \quad , \tag{5.84}$$

such that $a(\tau)$, $b(\tau)$ are determined by substituting this ansatz into 5.75 and 5.76. Note that the harmonic functions defined above satisfy the necessary conditions as stipulated in the previous section for metric perturbation decomposition; this is derived in appendix B. These perturbations involve a direct change in the density of matter, as can be seen from equation 5.77. Consequently, one obtains for equations 5.77 and 5.78:

$$\frac{\delta\rho}{\rho} = \frac{S^2}{9(S'^2 - S^2)}\left[(n^2+4)(a+b) + 3\frac{S'}{S}b'\right] Q \tag{5.85}$$
$$\delta v^\alpha = \frac{S^2}{6(S^2 - 2S'^2 + SS'')}\left[(n^2+1)b' + (n^2+4)a'\right] Q^\alpha \quad . \tag{5.86}$$

As an application, one can then solve these for an expanding universe scenario, using the most physical equations of state: dust $p = 0$ for late evolution (post-decoupling), and radiation $p = \frac{1}{3}\rho$ for the early stage; the results are:

Case $p = \frac{\rho}{3}$:

$$\delta f_\alpha^\beta = \frac{3A_1}{\tau} Q_\alpha^\beta + A_2(Q_\alpha^{(t)\beta} + Q_\alpha^\beta) \tag{5.87}$$
$$\frac{\delta\rho}{\rho} = \frac{n^2+4}{9}(A_1\tau + A_2\tau^2) \tag{5.88}$$
$$\delta v^\alpha = \frac{n^2+4}{12}\left[3A_1 + A_2\frac{n^2+1}{9}\tau^3\right] Q^\alpha \quad , \tag{5.89}$$

such that $n\tau \ll 1$; and A_1, A_2 are constants; and:

$$\delta f_\alpha^\beta = \frac{D}{n^2\tau^2}(Q_\alpha^\beta - 2Q_\alpha^{(t)\beta})\exp\left(\frac{in\tau}{\sqrt{3}}\right) \tag{5.90}$$

$$\frac{\delta\rho}{\rho} = -\frac{D}{9}\exp\left(\frac{in\tau}{\sqrt{3}}\right)Q \tag{5.91}$$

$$\delta v^\alpha = \frac{Din}{12\sqrt{3}}\exp\left(\frac{in\tau}{\sqrt{3}}\right)Q^\alpha \quad , \tag{5.92}$$

such that $n^{-1} \ll \tau \ll 1$, and D is a complex constant such that $|D| \ll 1$. These are the results for the early stages of the universe. From these results is is easy to show that for radiation dominated stages of such an expanding universe, the perturbations remain small, and are either constant, or decay with time; secondly, the amplitudes of the metric perturbations decrease proportionally to $\frac{1}{S^2}$. For the later expansions of the universe, i.e. pressure free ('dust'), it is found that the density perturbations are damped, decreasing proportionally with respect to $S^{-\frac{2}{3}}$ and then S^{-1}. However, the metric perturbations involving the higher order harmonics for this case exhibit an instability. For intermediate equations of state it is found that the perturbations slow down with the expansion of the universe, but have a reasonably stable amplitude. It is found in general that long wavelength perturbations (higher order harmonics) in an expanding universe lead to the changes in matter density increasing with time; short wavelength perturbations transpire as being nothing more than hydrodynamical sound waves with damped amplitude. In the contracting universe however, these perturbations are unstable, becoming increasingly large.

For the late stages of the universe (case $p = 0$)one obtains two sets of solutions corresponding to two constants B_1 and B_2:

For $\tau \ll 1$ and n small:

$$\delta f_\alpha^\beta = \frac{B_1}{\tau^3}\left[(n^2+1)Q_\alpha^\beta - (n^2+4)Q_\alpha^{(t)\beta}\right] \tag{5.93}$$

$$\frac{\delta\rho}{\rho} = \frac{2B_1(n^2+1)}{\tau^3}Q \tag{5.94}$$

$$\delta f_\alpha^\beta = B_2(Q_\alpha^\beta + Q_\alpha^{(t)\beta}) \tag{5.95}$$

$$\frac{\delta\rho}{\rho} = \frac{n^2+4}{30}B_2\tau^2 Q \tag{5.96}$$

$$\delta v^\alpha = 0 \quad . \tag{5.97}$$

For $\tau \gg 1$ and n small:

$$\delta f_\alpha^\beta = \frac{B_1}{2}\exp\left(-\tau[(n^2+1)Q_\alpha^\beta - (n^2+4)Q_\alpha^{(t)\beta}]\right) \tag{5.98}$$

$$\frac{\delta\rho}{\rho} = \frac{n^2+4}{8}B_1\exp(-\tau)Q \tag{5.99}$$

$$\delta f_\alpha^\beta = \frac{2n^2+5}{3} B_2 (Q_\alpha^\beta - Q_\alpha^{(t)\beta})$$
$$+4B_2\tau \exp\left(-\tau[(n^2+1)Q_\alpha^\beta - (n^2+4)Q_\alpha^{(t)\beta}]\right) \quad (5.100)$$
$$\frac{\delta\rho}{\rho} = \frac{n^2+4}{3} B_2 Q \quad , \quad (5.101)$$

and for the intermediate stage $n^{-1} \ll \tau \ll 1$:

$$\delta f_\alpha^\beta = \frac{B_2 n^2}{15} \tau^2 (Q_\alpha^\beta - Q_\alpha^{(t)\beta}) \quad (5.102)$$
$$\frac{\delta\rho}{\rho} = \frac{B_2 n^2}{30} \tau^2 Q \quad . \quad (5.103)$$

It is important to note here that the perturbations were considered to be *adiabatic* that is, they occur at constant entropy thus excluding dissipative forces; *i.e.* one assumed *ab initio* that the equation of state was not also a function of the entropy. The changes that might occur due to such dissipative forces cannot necessarily be assumed to be stable.

Case II: Vector (Rotational) Perturbations

As with the scalar perturbation case, one can transform the vector perturbation variable B_α of 5.23 in terms of a new function $c(\tau)$ and the vector harmonics via the following stipulation:

$$B_\alpha = 2\frac{c}{S^2} S_\alpha \quad (5.104)$$

to obtain the following simple form of the perturbed metric:

$$\delta f_\alpha^\beta = c(\tau) S_\alpha^\beta \quad . \quad (5.105)$$

As the trace of S_α^β vanishes, 5.76 vanishes completely, thus giving no density perturbation in 5.77, and corresponding to an effective 'rotational' perturbation. Applying the same procedure as before, one solves for the velocity perturbation from the above, obtaining:

$$\delta v^\alpha = \frac{S^2 c'}{4(2S'^2 - S^2 - SS'')} S^\alpha \quad , \quad (5.106)$$

such that for $\tau \ll 1$:

$$\delta v^\alpha = L S^\alpha \; ; \; p = \frac{\rho}{3} \quad (5.107)$$
$$\delta v^\alpha = \frac{M}{\coth\tau - 1} S^\alpha \quad , \quad (5.108)$$

for some constants M and L. This expression thus does not contain the harmonic order, thus imposing no restriction thereto and hence being applicable to all orders. Considering all the above scenarios, it is found that the velocity perturbation is either damped with time, or decreases as S^{-1}; while the metric perturbation are always damped with time (see [4]).

Case III: Tensor Perturbations

As before, one transforms the tensor perturbation variable $h_{\alpha\beta}$ of 5.24 in section 5.2 into a new variable $d(\tau)$ and the tensorial harmonics via:

$$h_{\alpha\beta} = \frac{d}{S^2} G_{\alpha\beta} \tag{5.109}$$

to obtain the simple perturbed metric:

$$\delta f_{\alpha}^{\ \beta} = d(\tau) G_{\alpha}^{\ \beta} \quad . \tag{5.110}$$

Solving as before, one obtains:

$$d(\tau) \ = \ N_1 \frac{\exp(in\tau)}{\sinh\tau} \ \ p \ = \ \frac{rho}{3} \tag{5.111}$$

$$d(\tau) \ = \ N_2 \frac{in}{\sinh^2(\frac{\tau}{2})} \exp(in\tau) \ \ p \ = \ 0 \ n\tau \ \gg \ 1 \quad . \tag{5.112}$$

In this case only the metric changes, while the matter remains static and uniformly distributed throughout space. This classification yields a periodic function for $d(\tau)$, thus resulting in *gravitational waves* propagating at the speed of light, and damped as S^{-1}. This results in the corresponding wave energy density (as in terms of amplitude) decreasing as S^{-4}, as one would expect.

Advantages and Disadvantages

The above pioneering approach to perturbation theory is readily seen to be conceptually simple and technically straight-forward. However, it should be noted that, although one may argue the assumptions to be quite restrictive and even over-simplistic, in doing so one is able to obtain a heuristic and intuitive feel for gravitational stability analysis and structure formation. This is indeed the major advantage of the approach.

However, the noticeable disadvantages are the evident gauge dependence - the synchronous gauge was assumed throughout, and spurious gauge modes were unambiguously alluded to - and the notion of perturbing a relatively unphysical quantity such as the metric. The latter concern will be dealt with through the covariant approach of the following chapter, while the former will be attended to through the gauge-invariant formulation of Bardeen [10] later in this chapter. In the meantime, an alternative method of solution to the above perturbation equations will be explored through the Sachs-Wolfe direct integration technique.

5.6 Solution By Direct Integration

Thus far the approach adopted entailed an *ab initio* ansatz form for the perturbed metric in terms of scalar, vector and tensorial quantities determined through substitution into the field equations. The approach of this section as illustrated through the pioneering work of Sachs and Wolfe [7], utilises what may in theory appear as the most logical solution to the problem, namely direct integration of the quantity δf_{ab}. However, in practice the technicalities of this method involve Fourier transforms, and ultimately arbitrary constants (scalars, vectors and tensors) arising from the integration which necessarily satisfy properties analogous to those of the ansatz quantities in the previous approach.

This section thus closely follows the Sachs-Wolfe paper of 1967 [7], not only as a reference to the direct integration method, but also the important result derived therefrom, namely the well-known Sachs-Wolfe Effect.

The direct integration method is used in the case of a flat ($K = 0$) FRW model incorporating equations of state $p = 0$ and $p = \frac{1}{3}\mu$, and ultimately used to investigate angular variations of the cosmic microwave background.

Instead of implicitly including the scale factor into the perturbed metric as in 5.67, one can define the perturbed metric h_{ab} in terms of the above as:

$$h_{ab} = \frac{1}{S^2}\delta f_{ab} \quad , \tag{5.113}$$

so that:

$$ds^2 = S^2(\tau)[\eta_{ab} + h_{ab}]dx^a dx^b \quad , \tag{5.114}$$

where η_{ab} is the Minkowski metric due to the flat space scenario $K = 0$. Secondly, assuming co-moving co-ordinates x^a, one has the comoving velocity as:

$$u^a = \frac{\delta^a_0}{S} \quad . \tag{5.115}$$

Starting then, as before, with the perturbed Einstein Field Equations 5.154, and making the perfect fluid assumption, one obtains for the 'flat' universe ($K = 0$):

$$\begin{aligned}
&(p+\rho)(u_a\delta u^b + u^b\delta u_a) + (\delta p + \delta\rho)u_a u^b + \delta p\delta^b_a \; = \\
&\tfrac{1}{2S^2}\left((h^a_{\ b} - h\delta^a_{\ b})^{;c}_{\ ;c} + h^{;a}_{\ ;b} + h^{ed}_{\ \ ;ed}\delta^a_{\ b} - h_{bd}^{\ \ ;ad} - h^{ac}_{\ \ ;bc}\right) \\
&+\left(4\tfrac{S'^2}{S^4} - 2\tfrac{S''}{S^3}\right)\delta^0_{\ b}h^{a0} - \tfrac{S'}{S^3}\left(h^{0\ ;a}_{\ b} + h^{0a}_{\ \ ;b} - (h^a_{\ b})'\right) \\
&+\tfrac{S'}{S^2}\left(2h^{0\alpha}_{\ \ ;\alpha} - h'\right)\delta^a_{\ b}
\end{aligned} \tag{5.116}$$

One now wishes to solve for the perturbed quantity h_{ab} from the above equations. As mentioned, the approach adopted here effectively amounts to integrating the above equations, which thus ultimately leads to the introduction of arbitrary scalar, tensor and vector quantities, although in a much less generalised and indirect sense than that outlined in section 5.2. As these equations are linear, one proceeds to solve for them by assuming the following Fourier transforms:

$$h_{\alpha 0} = \int H_{\alpha 0}(k\tau)e^{ik\cdot x}d^3k \qquad (5.117)$$

$$h_{\alpha\beta} = \int H_{\alpha\beta}(k\tau)e^{ik\cdot x}d^3k \quad , \qquad (5.118)$$

where $k = |\mathbf{k}|$, $\mathbf{k} = k^\alpha$ being the wave vector; and such that the tensors $H_{\alpha\beta}$, $H_{\alpha 0}$ and k^α all obey the following regularity conditions:

$$H_{\alpha\beta}k^\alpha k^\beta = k^4 o(k,\tau) \qquad (5.119)$$

$$H_{\alpha\beta}k^\beta = k^2 p_\alpha(k,\tau) \qquad (5.120)$$

$$H^\alpha_{\ \alpha} = k^4 q(k,\tau) \qquad (5.121)$$

$$H_{\alpha 0}k^\alpha = -ik^2 m(k,\tau) \qquad (5.122)$$

$$k^2 = \mathbf{k}\cdot\mathbf{k}$$

$$= -k_\alpha k^\alpha \quad , \qquad (5.123)$$

where o, p, q and m are to be generalised functions to coincide with continuous, ordinary functions in a small neighbourhood of $k = 0$. These functions are called 'moment functions', and the above constraints are called the 'moment conditions'. These conditions guarantee the uniqueness of the following splitting:

$$H(k\tau)_{\alpha 0} = n_\alpha + imk_\alpha \qquad (5.124)$$

$$n_\alpha k^\alpha = 0 \quad , \qquad (5.125)$$

and similarly for the the transverse and longitudinal tensor decompositions of the remaining components, as highlighted in the previous section. As mentioned, in solving for 5.116 using this method will generate some arbitrary constants, the exact form of which will depend on the initial conditions (from separate scalar, vector and tensor perturbations) of the system under consideration. These will be different according to which equation of state is chosen. For the equation of state $p = 0$ one requires two arbitrary scalars $A(x^\alpha)$, $B(x^\alpha)$; a vector function $C(x^\alpha)^\beta$ and tensor function $D(x^\alpha)^{\beta\gamma}$ which satisfy:

$$C^\alpha_{\ ;\alpha} = 0 \qquad (5.126)$$

$$D_{\alpha\beta} = D_{\beta\alpha} \qquad (5.127)$$

$$D^\alpha_{\ \alpha} = 0 \qquad (5.128)$$

$$D_{\alpha\beta}^{\ \ ;beta} = 0 \qquad (5.129)$$

$$\left(\frac{\partial^2}{\partial\tau^2} - \partial_\mu\partial^\mu\right)D_{\alpha\beta} = 0 \quad . \qquad (5.130)$$

For the equation of state $p = \frac{1}{3}\rho$, one also requires the above functional vector and tensor, but a single scalar function $E(x^\alpha)$ which satisfies:

$$\left(3\frac{\partial^2}{\partial\tau^2} - \partial_\mu\partial^\mu\right) E = 0 \quad . \tag{5.131}$$

One can then prove, using the above results, that all solutions to 5.116 and which also obey the moment conditions have the form:

Case I: $p = 0$

$$\begin{aligned} h_{\alpha\beta} &= \frac{1}{\tau}\frac{\partial}{\partial\tau}\left(\frac{1}{\tau}D_\alpha\beta\right) - 2\left(\frac{\tau}{8} - \frac{\partial_\mu\partial^\mu}{\tau}\right)(C_{\alpha;\beta} + C_{\beta;\alpha}) \\ &\quad + \frac{A_{;\alpha\beta}}{\tau^3} + \eta_{\alpha\beta}B - \frac{\tau^2}{10}B_{;\alpha\beta} \end{aligned} \tag{5.132}$$

$$h_{\alpha 0} = -2\frac{\partial_\mu\partial^\mu C_\alpha}{\tau^2} \tag{5.133}$$

$$\delta\rho = \frac{{H_0}^2}{4}\partial_\mu\partial^\mu\left(\frac{6A}{\tau^9} - \frac{3B}{5\tau^4}\right) \quad , \tag{5.134}$$

where H_0 is the present day value for Hubble's constant.

Case II: $p = \frac{1}{3}\rho$

$$\begin{aligned} h_{\alpha\beta} &= \frac{D_{\alpha\beta}}{\tau} - \left(\tau\partial_\mu\partial^\mu + \frac{8}{\tau}\right)(C_{\alpha;\beta} + C_{\beta;\alpha}) \\ &\quad + \frac{\tau^2}{2}\frac{\partial}{\partial\tau}\left(\frac{E_{;\alpha\beta}}{\tau^2} - \frac{\eta_{\alpha\beta}}{\tau^2}\frac{\partial E}{\partial\tau}\right) \end{aligned} \tag{5.135}$$

$$h_{\alpha 0} = -\partial_\mu\partial^\mu C_\alpha + \frac{\tau^2}{4}\frac{\partial}{\partial\tau}\left(\tau^{-2}\frac{\partial E_{;\alpha}}{\partial\tau}\right) \tag{5.136}$$

$$\delta\rho = 3\frac{{H_0}^2}{\tau^4}\frac{\partial}{\partial\tau}\left[\tau^2\frac{\partial}{\partial\tau}\left(\frac{1}{\tau^2}\frac{\partial E}{\partial\tau}\right)\right] \quad . \tag{5.137}$$

Moreover, this form of the solutions is true *only* up to a gauge transformation; the relevant gauge and co-ordinate frame in which the solutions hold being fixed up to the group of transformations:

$$\overline{x}^a = U^a_b x^b + c^a \quad , \tag{5.138}$$

where c^a is a constant vector, and U^a_b is a constant unitary transformation. Hence these results are not gauge-invariant, and one is thus obliged to keep track of the relevant gauge in all subsequent results.

In the above, the tensor $D_{\alpha\beta}$ corresponds to a generating function for gravitational waves having two degrees of freedom by the restriction 5.130. One can show that, for $k\tau \geq 1$, such a gravitational wave emitted at time τ_E and received at time τ_R is redshifted by an amount:

$$Z + 1 = \frac{S(\tau_R)}{S(\tau_E)} \quad . \tag{5.139}$$

In addition, one can also show that the remaining terms in the solutions for $p = 0$ have direct Newtonian analogues in the Navier-Stokes equations; while the vector C_α is related to the vorticity tensor $\omega_{\alpha\beta}$ as follows:

$$\begin{aligned} p = 0: \quad \omega_{\alpha\beta} &= \tfrac{1}{H_R}\nabla^2(C_{\beta,\alpha} - C_{\beta,\alpha}) \\ p = \tfrac{\varrho}{3}: \quad \omega_{\alpha\beta} &= \tfrac{\tau}{2H_R}\nabla^2(C_{\beta,\alpha} - C_{\beta,\alpha} \\ \omega_{\alpha 0} &= 0 \end{aligned} \quad . \tag{5.140}$$

These relations, along with the moment conditions and the divergenceless assumption of C_α, show that the vorticity tensor at any fixed τ and C_α uniquely determine each other, thus giving rise to rotational perturbations.

Considering energy density perturbations for case I via equation 5.134, one observes that the energy density can either increase or decrease, depending on the scalar functions A and B. One can argue here for the former, based on observational evidence for galaxy clusters; however these results are only accurate for calculating the gravitational field of an energy 'lump', but not of its internal structure. For the density perturbations of case II via equation 5.137 one obtains density *waves* which are either increasing *or* decreasing, and are redshifted in the same way as for gravitational and electromagnetic waves. One can now proceed to integrate the geodesic equations using equations 5.114 and 5.115, obtaining:

$$\frac{d}{dv}\left(\eta_{ab}\frac{dx^b}{dv} + h_{ab}\frac{dx^b}{dv}\right) = \frac{1}{2}h_{bc,a}\frac{dx^b}{dv}\frac{dx^c}{dv} \quad . \tag{5.141}$$

This yields, to zero order:

$$\frac{d^2}{dv^2}x^a_{(0)} = 0 \quad , \tag{5.142}$$

where v is an affine parameter, and the subscript (0) indicates that this is an unperturbed quantity. Considering then a light signal emitted and received at (τ_E, x^α), $(\tau_R, 0)$ respectively; equation 5.142 is integrated to yield:

$$\tau_{(0)} = \tau_R - v \tag{5.143}$$

$$x^\alpha_{(0)} = vl^\alpha \quad , \tag{5.144}$$

(5.145)

such that $l^\alpha l_\alpha = -1$; $l^\alpha = cst$ being a constant unit vector representing to zero order the spatial direction of the light signal as seen by an observer moving with the fluid. This then yields the zero order tangent vector as:

$$j^a = (-1, l^\alpha) \quad . \tag{5.146}$$

The first order conformal time correction $\tau_{(1)}$ to $\tau_{(0)}$ (which is all that will be used from this in subsequent analysis) then becomes, from equation 5.141:

$$\frac{d\tau_{(1)}}{dv} = -\left(h_{\alpha 0}e^\alpha\right)_{(0)} + \frac{1}{2}\int_0^v\left(\frac{\partial h_{\alpha\beta}}{\partial\tau}l^\alpha l^\beta - 2\frac{\partial h_{\beta 0}}{\partial\tau}l^\beta\right)_{(0)} du \tag{5.147}$$

such that all the quantities in the integral are evaluated at the unperturbed $x^a_{(0)}(u)$. This leads to a redshift correct to first order terms of:

$$z+1=\frac{S(\tau_R)}{S(\tau_E)}\left[1-\frac{1}{2}\int_0^{\tau_R-\tau_E}\left(\frac{\partial h_{\alpha\beta}}{\partial\tau}l^\alpha l^\beta-2\frac{\partial h_{\beta 0}}{\partial\tau}l^\beta\right)_{(0)}du\right] \quad . \tag{5.148}$$

On account of this being a physically observable quantity, one would expect it to be invariant under the gauge transformation defined in terms of infinitesimal translations $\overline{x}^a = x^a - \zeta^a$ as defined in the preceding literature. A simple calculation shows this indeed to be the case.

One can now proceed to estimations of angular variations in the microwave radiation, culminating in the highly significant *Sachs-Wolfe effect.* This can be shown to be caused by the density fluctuations $\delta\rho$ which in turn contribute to the gravitational field by equation 5.134 which consequently causes a change in the redshift as in equation 5.148; if this microwave radiation is then cosmological, it shows a corresponding variation of temperature with angle. In demonstrating this result, one needs to make the following assumptions:

- i)the present value of Hubble's parameter is taken to be 10^{-10} year^{-1};
- ii)density fluctuations on the scale of less than 10^9 light years are ignored, and it is assumed that at present ($\eta_R = 1$) for some scale $L \simeq 10^9 - 10^{10}$ light years there are density variations $\frac{\delta\rho}{\rho}$ of the order of ten percent;
- iii)The assumed appropriate background model has $p = 0$, and with the microwave radiation giving a negligible contribution to ρ.
- iv) Only density perturbations of the *relatively increasing* type are assumed relevant.
- v) It is assumed that at some $\tau_E \leq \frac{1}{2}$ in the gauge frame of equation 5.134 the microwave radiation as measured by observers moving with a gas of $p = 0$ was isothermal with temperature T_E independent of the position x^a. It is further assumed that since the 'time' τ_E no significant Thompson scattering of the microwave background has taken place; it will be further assumed that $\tau_E \simeq \frac{1}{30}$; this all being done using the relationship:

$$\frac{T_R}{T_E}=\frac{S(\tau_E)}{S(\tau_R)}=\frac{{\tau_E}^2}{{\tau_R}^2}\sim 10^{-2} \quad , \tag{5.149}$$

  which holds exactly in the background model.

The rather conservative nature of the above assumptions is intended at gaining a *lower* limit of the microwave radiation anisotropy. A *caveat* regarding the physicality of intrinsic uniformity is in assumption v), stating that any intrinsic variations in temperature emissions could easily dominate the subsequent analysed effects. Secondly, it is noted that this assumption is not gauge-invariant under the relevant fixed gauge of equation 5.134; hence this assumption only becomes meaningful when analysed within the same gauge of 5.134.

Proceeding thence, and using results from standard geometric optics, one can derive the relationship between the emitted and observed (received) temperatures T_E and T_R respectively, and the redshift z; obtaining the following exact expression:

$$T_R = \frac{T_E}{z+1} \quad , \tag{5.150}$$

demonstrating the inverse relationship between the observed temperature and the redshift expression $z+1$. Using this and equations 5.148, 5.149 one thus derives, to *first order*:

$$\begin{aligned} T_R &= T_E \frac{{\tau_E}^2}{{\tau_R}^2} \left(1 + \frac{1}{2} \int_0^{\tau_R - \tau_E} \left(\frac{\partial h_{\alpha\beta}}{\partial \tau} l^\alpha l^\beta - 2 \frac{\partial h_{\beta 0}}{\partial \tau} l^\beta \right)_{(0)} du \right) \\ &= T_R \left(1 + \frac{1}{2} \int_0^{\tau_R - \tau_E} \left(\frac{\partial h_{\alpha\beta}}{\partial \tau} l^\alpha l^\beta - 2 \frac{\partial h_{\beta 0}}{\partial \tau} l^\beta \right)_{(0)} du \right) \quad , \end{aligned} \tag{5.151}$$

so that one must have, to first order, the following relation:

$$\begin{aligned} \frac{\delta T_R}{T_R} &= \frac{1}{2} \int_0^{\tau_R - \tau_E} \left(\frac{\partial h_{\alpha\beta}}{\partial \tau} l^\alpha l^\beta - 2 \frac{\partial h_{\beta 0}}{\partial \tau} l^\beta \right)_{(0)} du \\ &= \frac{1}{10} \left[(B_{,\alpha} l^\alpha)_R \tau_R - (B_{,\alpha} l^\alpha)_E \tau_E + B_R - B_E \right] \quad , \end{aligned} \tag{5.152}$$

where in the last step equation 5.134 has been used together with the assumptions iii) and iv) above which themselves result in $A = C_\alpha = D_{\alpha\beta} = 0$ in 5.134. One can now proceed to analyse *heuristically* the four separate terms in equation 5.152. It thus transpires that the first two terms are Doppler-type terms; the second being negligible small, and essentially a Doppler correction for the world velocity of the source. The last two terms, which are the most interesting, are similar to gravitational redshift terms on account of B behaving analogously to a Newtonian potential: one can consider $\delta\rho$ at the present time to be the source of this potential. An heuristic approximation can then be made for the magnitude of these last two terms, the result being:

$$\frac{\delta T_R}{T_R} \simeq 0.005 \quad , \tag{5.153}$$

which, as earlier stated, is a *lower* order estimate. This temperature anisotropy constitutes the Sachs-Wolfe effect. It must be emphasised though that Cosmology, especially Observational Cosmology, has come a long way since this modest estimate. Indeed, with the observations of COBE [18] and subsequently BOOMERanG over the past decade, the above estimate, as well as the analysis, have been considerably refined: with the latest BOOMERanG results ([23],[24]) an observed anisotropy of $\Delta T_{200} = (69 \pm 8)\mu K$ peaking at the Legendre multipole of $I_{peak} = 197 \pm 6$ was reported, the measurements taken over a sixty degree portion of the sky. As mentioned in the introduction, the developments of the last decade indeed merit an extensive treatment of their own, as they have revolutionised modern Cosmology, providing ever more credible evidence for the Big Bang theory and ever increasing support for the inflationary scenario. In the context of this thesis, is must also be stressed that the CMB anisotropy is the ultimate evidence for the whole theory of structure formation, and thus vindication of Perturbation Theory.

5.7 Gauge-Invariant Formulation

Having explored thus the merits and disadvantages of the gauge-related formalism, one can consider the natural obviation of the associated gauge difficulties, namely the formulation of gauge-invariant variables. This problem was tackled exhaustively by Bardeen in his seminal 1980 paper [10]. His approach is mathematically elegant, consistent and rigorous, and will be expounded upon later. A criticism which has been levelled at this formulation, though, is the aspect of the cumbersome technical mathematical complexity thereof; and as a consequence, the difficulty encountered in interpreting the physicality of his variables which take on a rather contrived appearance. Upon studying the gauge problem though, one is compelled to concede the inevitability of such a technically demanding formalism.

Indeed, credit should be given for the boldness in tackling the gauge problem head-on, and the consequent emergence of a formalism which, though technically awkward, has as its primary advantage direct practical applicability. Thus the true beauty and physical relevance of the formalism can only be appreciated through application. As an introduction to the gauge-invariant approach, it is preferable first to formulate a simple theory regarding the formulation of gauge-invariant equations; for the sake of simplicity, this will be done for scalar perturbations only, and using the gauge-invariant scalar variables A, B defined earlier. The ideas behind the scalar formulation are then readily extendable to vector and tensor perturbations.

5.7.1 Gauge-Invariant Formulation of the Field equations

Having established and clarified the nature of gauge transformations, one is now in a position to derive a gauge-invariant form of the field equations. This is done by initially calculating the *first order* variation of the field equations in their various components, and proceeding thence to the construction of an analogous gauge-invariant form by using the gauge-invariant variables of the previous section, for the case of scalar perturbations only. The ensuing analysis follows [17] closely. As before, the background model is chosen to be FRW around which one has the linearised equations:

$$\begin{aligned} 8\pi G\delta T^a_{\ b} &= \delta G^a_{\ b} \\ &= \delta R^a_b - \frac{1}{2}\delta R\delta^a_b \ . \end{aligned} \tag{5.154}$$

Now, for metric perturbations the Einstein tensor can be Taylor expanded as:

$$G^a_{\ b} = G^{(0)a}_{\ \ \ b} + \delta G^a_{\ b} + \dots \tag{5.155}$$

where $G^{(0)a}_{\ \ \ b}$ is the background quantity. Note that the variation entailed in the above expression is calculated by *linearising about the background metric*; that is, by disregarding quantities which are quadratic or of higher order in the perturbation variables. Hence, using equations 4.103, 4.104 and 4.105 one performs the linearised variation on them in terms of the perturbed metric, obtaining the following variations (non gauge-invariant) of the Einstein tensor in terms of the previously derived gauge-invariant variables.

$$\delta G^0_{\ 0} = \frac{2}{S^2}\left[-3\mathcal{H}(\mathcal{H}A + B') + \nabla^2 B + 3KB + 3\mathcal{H}(-\mathcal{H}' + \mathcal{H}^2 + K)(\psi - \kappa')\right] \tag{5.156}$$

$$\delta G^0_{\ \alpha} = \frac{2}{S^2}\left[\mathcal{H}A + B' + (\mathcal{H}' - \mathcal{H}^2 - K)(\psi - \kappa')\right]_{,\alpha} \tag{5.157}$$

$$\begin{aligned}\delta G^\alpha_\beta = & -\frac{2}{S^2}\left\{\left[(2\mathcal{H}' + \mathcal{H}^2)A + \mathcal{H}A' + B'' + 2\mathcal{H}B' - KB + \frac{1}{2}\nabla^2(A - B)\right]\delta^\alpha_\beta\right. \\ & \left. +(\mathcal{H}'' - \mathcal{H}\mathcal{H}' - \mathcal{H}^3 - K\mathcal{H})(\psi - \kappa')\delta^\alpha_\beta - \frac{1}{2}o^{\alpha\gamma}(A - B)_{;\gamma\beta}\right\} \quad ,\end{aligned} \tag{5.158}$$

Where $\mathcal{H} = \frac{\dot{S}}{S}$ is the Hubble parameter; the prime sign indicating differentiation with respect to τ, a 'dot' superscript indicating differentiation with respect to co-ordinate time t, as in the preceding chapter. These equations are readily seen to be non-gauge-invariant; indeed, under the infinitesimal co-ordinate transformation 5.40 they transform as:

$$\delta G^0_0 \rightarrow \delta G^0_0 - (G^0_0)'\zeta^0 \tag{5.159}$$

$$\delta G^0_\alpha \rightarrow \delta G^0_\alpha - (G^0_0 - \frac{1}{3}G^\beta_\beta)\zeta^0_{;\alpha} \tag{5.160}$$

$$\delta G^\alpha_\beta \rightarrow \delta G^\alpha_\beta - (G^\alpha_\beta)'\zeta^0 \quad . \tag{5.161}$$

This suggests that one define the following gauge-invariant quantities:

$$\delta G^{(gi)0}_0 = \delta G^0_0 + (G^0_0)'(\psi - \kappa') \tag{5.162}$$

$$\delta G^{(gi)0}_\alpha = \delta G^0_\alpha + (G^0_0 - \frac{1}{3}G^\beta_\beta)(\psi - \kappa')_{;\alpha} \tag{5.163}$$

$$\delta G^{(gi)\alpha}_\beta = = \delta G^\alpha_\beta + (G^\alpha_\beta)'(\psi - \kappa') \quad . \tag{5.164}$$

Note that this form is purely arbitrary - it is quite evident that there is no unique gauge-invariant form; indeed, as will be seen later, Bardeen [10] makes a different choice. For the energy-momentum equations the matter is assumed to be hydrodynamical. More particularly, a *perfect* fluid will be assumed, as before:

$$T^a_b = (\rho + p)u^a u_b + p\delta^a_b \quad , \tag{5.165}$$

the variation of which yields:

$$\delta T^0_{\ 0} = -\delta\rho \tag{5.166}$$

$$\delta T^0_{\ \alpha} = -(\rho + p)S^{-1}\delta u_\alpha \tag{5.167}$$

$$\delta T^\alpha_\beta = \delta p \delta^\alpha_\beta \quad . \tag{5.168}$$

One should note here that, in general, p will be a function of both the energy density ρ *and* the entropy $\mathcal{E}$; hence:

$$\begin{aligned} \delta p &= \left(\frac{\partial p}{\partial \rho}\right)|_\sigma \delta\rho + \left(\frac{\partial p}{\partial \sigma}\right)|_\rho \delta\sigma && (5.169)\\ &= c_s^2 \delta\rho + T\delta\sigma \quad , && (5.170)\end{aligned}$$

where c_s can be interpreted as the speed of sound, and T is the temperature. Now, as with the Einstein tensor components, one can construct the following gauge-invariant variables for the energy-momentum tensor:

$$\begin{aligned} \delta\rho^{(gi)} &= \delta\rho + \rho'(\psi - \kappa') && (5.171)\\ \delta p^{(gi)} &= \delta p + p'(\psi - \kappa') && (5.172)\\ \delta u_\alpha^{(gi)} &= \delta u_\alpha + S(\psi - \kappa')_{;\alpha} \quad , && (5.173)\end{aligned}$$

so that:

$$\begin{aligned} \delta T_0^{(gi)0} &= \delta T_0^0 + (T_0^0)'(\psi - \kappa') \\ &= \delta\rho^{(gi)} && (5.174)\\ \delta T_\alpha^{(gi)0} &= \delta T_\alpha^0 + (T_0^0 - \frac{1}{3}T_\beta^\beta)(\psi - \kappa')_{;\alpha} \\ &= (\rho + p)S^{-1}\delta u_\alpha^{(gi)} && (5.175)\\ \delta T_\beta^{(gi)\alpha} &= \delta T_\beta^\alpha + (T_\beta^\alpha)'(\psi - \kappa') \\ &= -\delta p^{(gi)}\delta_\beta^\alpha \quad . && (5.176)\end{aligned}$$

Combined, the above sets of equations thus yield a fully gauge-invariant analogous form for the Einstein field equations:

$$8\pi G\delta T_b^{(gi)a} = \delta G_b^{(gi)a} \quad . \qquad (5.177)$$

Writing out both sides of these equations in full, one obtains:

$$\begin{aligned} 4\pi G S^2 \delta\rho^{(gi)} &= 3\mathcal{H}(\mathcal{H}A + B') - \nabla^2 B + 3KB && (5.178)\\ 4\pi G S(\rho + p)\delta u_\alpha^{(gi)} &= -(\mathcal{H}A + B')_{,\alpha} && (5.179)\\ 4\pi G S^2 \delta p^{(gi)}\delta_\beta^\alpha &= -\Big[(2\mathcal{H}' + \mathcal{H}^2)A + \mathcal{H}A' + B'' + 2\mathcal{H}B' \\ &\quad -KB + \frac{1}{2}\nabla^2(A - B)\Big]\delta_\beta^\alpha + \frac{1}{2}o^{\alpha\gamma}(A - B)_{;\gamma\beta} \quad . && (5.180)\end{aligned}$$

In the last equation, the absence of non-diagonal terms leads to $A = B$. This leads the following simplification of the above equations:

$$
\begin{aligned}
\nabla^2 A - 3\mathcal{H}A' - 3(\mathcal{H}^2 - K)A &= -4\pi G S^2 \delta\rho^{(gi)} && (5.181)\\
(SA)'_{,\alpha} &= -4\pi G S^2 (\rho + p)\delta u_\alpha^{(gi)} && (5.182)\\
A'' + 3\mathcal{H}A' + (2\mathcal{H}' + \mathcal{H}'' - K)A &= -4\pi G S^2 \delta p^{(gi)} \quad . && (5.183)
\end{aligned}
$$

One can now combine equations 5.170 with 5.181 and 5.183 to obtain a single equation:

$$
A'' + 3\mathcal{H}(1+{c_s}^2)A' - {c_s}^2\nabla^2 A + \left[2\mathcal{H}' + (1+3{c_s}^2)(\mathcal{H}^2 - K)\right] A = -4\pi G S^2 T\delta\sigma \quad . \quad (5.184)
$$

Similarly, one may obtain an evolution equation for the gauge-invariant variable B; or as done by Bardeen, formulate a gauge-invariant variable for the matter, and obtain an evolution equation for this, as will be seen later.

From these equations, the gauge-invariant evolution equations of any desired perturbed quantity can be obtained. However, the above is fully general: in application, one needs to expand the perturbation variables in terms of the spherical harmonics, and manipulate the form of the equations to yield equations which are technically amenable to clear interpretation of, and insight into that particular situation. This will be seen in the Bardeen approach of the following section.

5.7.2 The Bardeen Formalism

The *modus operandi* adopted in the Bardeen formalism is essentially that of the preceding sections, in that the standard metric perturbation form is assumed. The aim of the Bardeen approach is to circumvent the gauge problem. In the preceding sections various gauge-invariant quantities and equations were developed within specialised, simplistic cases devised for the sake of clarity. In this section the gauge-problem is extended to full generality, encompassing vector and tensor perturbations in addition to scalar perturbations. The consequent formulations, particularly that of the gauge-invariant quantities, are practically motivated within a physical context, admitting then a fairly straight-forward application to any desired physical context.

One may motivate the ensuing formalism by addressing the following flaws and problems in the preceding direct metric perturbation approaches.

- Physical interpretation of density perturbations which are larger than the particle horizon (at early times in the universe); which is relevant in explaining perturbations which give rise to galaxy formation.
- The principal problem of gauge freedom, as already explained.
- Choice of pertinent, physically meaningful, and unambiguous hypersurface 'slicing'.

These will be elucidated upon in detail in due course.

One thus desires a formalism which is gauge-invariant *and* solves the above problems, constructing in the process a framework for studying the evolution of physical perturbations. This, as before, entails the formulation of geometric quantities from the metric in terms of separate scalar, vector and tensor perturbations. For scalar perturbations in general, which transpire as the most useful, the relevant variables are the following gauge-invariantly formulated quantities:

- the velocity amplitude
- the entropy amplitude
- the anisotropic stress amplitude
- the fractional energy density perturbation in the comoving frame

Physical perturbations relative to any well-defined set of hypersurfaces, such as comoving, zero shear or uniform-Hubble-constant hypersurfaces, can then be represented by appropriate combinations of these gauge-invariant amplitudes.

In the course of the study the following standard assumptions will be made:

- Perfect fluid matter description
- Robertson-Walker background metric
- A non-zero cosmological constant

The assumption of a homogeneous and isotropic background permits the separation of time from spatial dependence of perturbations of various quantities, with spatial solutions related to solutions of the generalised Helmholtz equation, resulting in scalar perturbations being represented in terms of four-dimensional (scalar) harmonics, as in the Lifshitz-Khalatnikov approach. In the previous section, this was implicitly assumed and incorporated into the perturbation quantities contained in δf_{ab}. The reason for expanding in terms of the harmonics is that they form a complete basis for the space-time; consequently, any quantity defined in the space-time can be expanded in terms of these quantities. Such bases are however, not unique, so one is at liberty to choose a basis which has a structure amenable to the context, namely that of quantities derived in terms of the Einstein Field Equations - this turns out to be the Helmholtz condition stipulated above. All vector and tensor quantities associated with such a scalar perturbation can then, as with the Lifshitz-Khalatnikov [4] approach, be formulated in terms of covariant derivatives (with respect to the spatial part of the metric) of the scalar harmonics. An analogous result follows for vector and tensor perturbations in terms of the vector and tensor harmonics respectively. Hence, in the following theory, it will be assumed that any given quantity defined on the manifold can be written as a linear combination of the appropriately harmonic-constructed quantities, with co-efficients as functions of the conformal time. Consequently, the scalar harmonics Q have to obey equation B.1 in Appendix B; where k is the wave number. This wave number sets the spatial scale of the perturbation relative to the comoving background co-ordinates. Note that this is a consideration which has to feature in practical calculations, but which is absent from the more generic approach of the previous section. Vector and tensorial quantities associated with these scalar harmonics Q can then be defined in a consistent way, as developed in appendix B; likewise for the vector $\mathcal{S}_\alpha$ and tensor harmonics $\mathcal{G}_{\alpha\beta}$. Then by equations 5.6 and 5.15 one can decompose any spatial vector V_α and tensor $H_{\alpha\beta}$ as follows:

$$\begin{aligned} V_\alpha &= \phi_{;\alpha} + B_\alpha \\ &= V^{(s)}(\tau)\mathcal{Q}_\alpha + V^{(v)}(\tau)\mathcal{S}_\alpha \qquad (5.185) \\ H_{\alpha\beta} &= \Delta_{\alpha\beta}\phi + 2B_{(\alpha;\beta)} + W_{\alpha\beta} \\ &= H^{(s)}(\tau)\mathcal{Q}_{\alpha\beta} + H^{(v)}(\tau)\mathcal{S}_{\alpha\beta} + H^{(t)}(\tau)\mathcal{G}_{\alpha\beta} \quad , \qquad (5.186) \end{aligned}$$

where covariant differentiation is taken with respect to the spatial metric $o_{\alpha\beta}$; S_α is solenoidal, and $W_{\alpha\beta}$ is trace-free and transverse. The superscripts $(s), (v)$ and (t) stand for 'scalar', 'vector' and 'tensor' respectively, referring to the scalar, vector and tensor perturbation contributions. These decompositions are then inserted into the perturbed metric form according to equation 5.25:

$$\begin{aligned} ds^2 &= S^2(\tau)\Big[-(1+2A(\tau)\mathcal{Q})d\tau^2 - 2\left(B^{(s)}(\tau)\mathcal{Q}_\alpha + B^{(v)}(\tau)\mathcal{S}_\alpha\right)dx^\alpha d\tau \\ &\quad + \Big((1+H_{(L)})o_{\alpha\beta} + 2\Big(H^{(s)}_{(S)}(\tau)\mathcal{Q}_{\alpha\beta} + H^{(v)}_{(T)}(\tau)\mathcal{S}_{\alpha\beta} \\ &\quad + H^{(t)}_{(T)}(\tau)\mathcal{G}_{\alpha\beta}\Big)\Big)dx^\alpha dx^\beta\Big] \quad . \end{aligned} \qquad (5.187)$$

Naturally, the matter, via the energy-momentum tensor, has to be perturbed separately; hence, analogous to the gravitational perturbations above, one can define a matter perturbation to the overall background energy density ρ_0 and pressure p_0, assuming a perfect fluid matter description, as:

$$\begin{aligned} \rho &= \rho_0(\tau)\left[1+\delta(\tau)\mathcal{Q}\right] \qquad (5.188) \\ p &= p_0(\tau)\left[1+\pi_L(\tau)\mathcal{Q}\right] \quad , \qquad (5.189) \end{aligned}$$

and secondly, the perturbed four-velocity:

$$\begin{aligned} u^\alpha &= u^{0(s)}v^\alpha \\ &= S^{-1}\left[v^{;\alpha} + S^\alpha\right] \\ &= S^{-1}\left[v^{(s)}(\tau)\mathcal{Q}^\alpha + v^{(v)}(\tau)\mathcal{S}^\alpha\right] \qquad (5.190) \\ u^0 &= S^{-1}(1-A\mathcal{Q}) \quad , \qquad (5.191) \end{aligned}$$

where $u^{0(s)}$ is the background velocity time component, while the vectorial decomposition 5.6 has been used. One can then define the *entropy perturbation* $\eta(\tau)\mathcal{Q}$ as the difference between the fractional pressure perturbation and the background pressure-energy density relation:

$$\begin{aligned}\eta(\tau)Q &= \left(\pi_L - \frac{E_0}{P_0}\frac{dP_0}{dE_0}\delta\right)Q \\ &= \frac{1}{W}(W\pi_L - c_s^2\delta)Q \quad . \end{aligned} \tag{5.192}$$

Note that if a generic matter description were assumed, one would require further perturbation variables corresponding to the anisotropic stress terms $\pi_{\alpha\beta}$ in the energy-momentum tensor:

$$\begin{aligned}\pi_{\alpha\beta} &= p_0\left[\pi_{;\alpha\beta} + S_{\alpha;\beta} + T_{\alpha\beta}\right] + P_0 o_{\alpha\beta} \\ &= p_0\left[\pi_T^{(s)}Q_{\alpha\beta} + \pi_L Q o_{\alpha\beta} + \pi_T^{(v)}S_{\alpha\beta} + \pi_T^{(t)}G_{\alpha\beta}\right] + p_0 o_{\alpha\beta} \quad , \end{aligned} \tag{5.193}$$

where the tensorial decomposition 5.15 has been used. Similarly, for the energy flux q_α:

$$\begin{aligned} q_\alpha &= q_{;\alpha} + S_\alpha \\ &= P_0\left[f^{(s)}(\tau)Q_\alpha + f^{(v)}(\tau)S_\alpha\right] \end{aligned} \tag{5.194}$$

using the vector decomposition 5.6 again. Note that the imposition $v^{(s)}(\tau) = 0$ corresponds to what is termed the *comoving gauge* as described earlier; termed as such, because this stipulation implies zero energy flux, consequently implying a gauge in the rest frame of the matter.

Scalar Perturbations

Analogous to the Brandenberger *et al* approach of section 5.2, and making the following transformations from those perturbation variables:

$$\phi = -AQ \tag{5.195}$$

$$\psi = -\frac{B^{(s)}}{k}Q \tag{5.196}$$

$$\gamma = \left(H_{(L)} + \frac{1}{3}H_{(T)}\right)Q \tag{5.197}$$

$$\kappa = -\frac{H_{(T)}^{(s)}}{k^2}Q \quad , \tag{5.198}$$

one thus obtains, for the perturbed metric in conformal time the following scalar-related contributions to the gravitational perturbations:

$$g_{00} = -S^2[1 + 2AQ] \tag{5.199}$$

$$g_{0\alpha} = \frac{S^2}{k}B^{(s)}Q_{,\alpha}$$

$$= -S^2 B^{(s)} \mathcal{Q}_\alpha \quad (5.200)$$

$$g_{\alpha\beta} = S^2 \left[\left(1 + 2[H_{(L)} + \frac{1}{3} H_{(T)}] \mathcal{Q} \right) o_{\alpha\beta} + 2 \frac{H_{(T)}}{k} \mathcal{Q}_{;\alpha\beta} \right]$$

$$= S^2 \left[(1 + 2H_{(L)} \mathcal{Q}) o_{\alpha\beta} + 2H_{(T)}^{(s)} \mathcal{Q}_{\alpha\beta} \right] \quad . \quad (5.201)$$

The associated matter perturbations are:

$$T^0_0 = -\rho \quad (5.202)$$

$$T^\alpha_\beta = p_0[1 + \pi_L \mathcal{Q}]\delta^\alpha_\beta + p_0 \pi_T^{(s)} \mathcal{Q}^\alpha_\beta \quad (5.203)$$

$$T^\alpha_0 = -(\rho_0 + p_0) v^{(s)} \mathcal{Q}^\alpha \quad (5.204)$$

$$T^0_\alpha = (\rho_0 + p_0)(v^{(s)} - B^{(s)}) \mathcal{Q}_\alpha \quad . \quad (5.205)$$

These, together with the velocity perturbation equation yield the set of unknown quantities A, $B^{(s)}$, $H_{(L)}$, $H_{(T)}^{(s)}$, $v^{(s)}(\tau)$ and $\delta(\tau)$ which are the desired arbitrary functions of conformal time. In the above and henceforth, the superscript (s) indicates quantities associated strictly with *scalar* perturbations.

The subscripts 'L' and 'T' stand for 'longitudinal' and 'transverse', corresponding to the associated gauge-choices obtained by setting these respectively equal to zero. A noteworthy point here is that, in the alternative ADM formalism of section 4.5.2, the quantity $S^2 A \mathcal{Q}$ corresponds to the lapse function N, while the quantity $S^2 B \mathcal{Q}_\alpha$ corresponds to the shift vector N_α. This fact can be used to obtain a more geometrical interpretation of the subsequently derived results. Note that in terms of the above, the quantities defined through equations 5.20 and 5.21 take on the following form:

$$L = \left(\dot{H}_{(T)}^{(s)} - \frac{1}{S} B \right) \mathcal{Q} \quad (5.206)$$

$$K = -\left(3 \left(\frac{H_{(L)}{}' + \frac{1}{3} H_{(T)}{}'}{S^2} + 2S\mathcal{H}A \right) + L^{;\alpha}_{;\alpha} \right) \mathcal{Q} \quad . \quad (5.207)$$

As mentioned earlier, these are the perturbation generator of the traceless part, and perturbation to the trace respectively of the extrinsic curvature; that is, the shear in the normal world lines. Now, in order to introduce the gauge-problem, and consequently develop gauge-invariant variables, one has to consider an arbitrary co-ordinate transformation, as done in the preceding sections:

$$\overline{\tau} = \tau T(\tau) \mathcal{Q} \quad (5.208)$$

$$\overline{x}^\alpha = x^\alpha + L^{(s)}(\tau) \mathcal{Q}^\alpha \quad , \quad (5.209)$$

this being motivated by the interpretation of a gauge transformation as a co-ordinate transformation induced by a change in correspondence between the background and physically perturbed space-times. It is important to note here that, owing to the homogeneity and isotropy of the background spatial part of the space-time, perturbations of all physical quantities must be gauge-invariant under *purely spatial* gauge transformations. However, as the background also evolves in time, perturbations to physical quantities, even space-time scalars in the physical space-time, will in general not be gauge-invariant under *time* gauge transformations.

The above co-ordinate transformation now enables one to study the transformation properties of the above perturbation variables. For scalar perturbations the variables then transform in the following way:

$$\overline{A} = A - T' - T\frac{S'}{S} \tag{5.210}$$

$$\overline{B}^{(s)} = B^{(0)} + L^{(s)\prime} + kT \tag{5.211}$$

$$\overline{H}_L = H_L - L^{(s)}\frac{k}{3} - T\frac{S'}{S} \tag{5.212}$$

$$\overline{H}_T^{(s)} = H_T^{(s)} + kL^{(s)} \quad , \tag{5.213}$$

as well as:

$$\overline{v}^{(s)}Q^\alpha \simeq \frac{dx^\alpha}{d\tau} + L^{(s)}Q^\alpha \tag{5.214}$$

$$\rho(\overline{\tau}) \simeq \rho_0(\tau)[1 + (\overline{\delta} + T\frac{\dot{E}_0}{E_0})Q] \tag{5.215}$$

$$\overline{\delta} = \delta + 3(1+W)T\frac{\dot{S}}{S} \tag{5.216}$$

$$\overline{\pi}_L = \pi_L + 3(1+W)\frac{{c_s}^2}{W}\frac{\dot{S}}{S}T \quad , \tag{5.217}$$

where:

$$W = \frac{p_0}{\rho_0}, \quad {c_s}^2 = \frac{dp_0}{d\rho_0} \quad . \tag{5.218}$$

Arguing then for the necessity through physicality of gauge-invariant quantities, one can formulate the following gauge-invariant quantities:

$$\Phi_A = A + \frac{1}{k}\dot{B}^{(s)} + \frac{1}{k}\frac{\dot{S}}{S}B^{(s)} - \frac{1}{k^2}\left(\ddot{H}_{(T)}^{(s)} + \frac{\dot{S}}{S}\dot{H}_T^{(s)}\right) \tag{5.219}$$

$$\Phi_H = H_{(L)} + \frac{1}{3}H_{(T)}^{(s)} + \frac{1}{k}\frac{\dot{S}}{S}B^{(s)} - \frac{1}{k^2}\frac{\dot{S}}{S}\dot{H}_{(T)}^{(s)} \quad . \tag{5.220}$$

These equate to the gauge-invariant variables defined in section 5.2 as equations 5.50 and 5.51 via the transformations 5.195 - 5.198, and can be interpreted as a *lapse function perturbation* and *curvature perturbation* respectively in the ADM formalism provided that the hypersurfaces are chosen so as to make $L = 0$, where L is defined through equation 5.206. One can also define a gauge-invariant velocity variable:

$$v^{(s)}_{(s)} = v^{(s)} - \frac{1}{k}\dot{H}^{(s)}_{(T)} \quad , \tag{5.221}$$

a variable which is then physically interpreted in terms of the shear of the matter field. For the energy density perturbation amplitude δ, one then requires a gauge-invariant quantity which will reduce to δ as soon as the perturbation comes within the particle horizon: $\frac{1}{k}\frac{\dot{S}}{S} \ll 1$. Two possibilities can separately be considered:

$$\varepsilon_m = \delta + 3(1+W)\frac{1}{k}\frac{\dot{S}}{S}(v^{(s)} - B^{(s)}) \tag{5.222}$$

$$\varepsilon_g = \delta - 3(1+W)\frac{1}{k}\frac{\dot{S}}{S}\left(B^{(s)} - \frac{1}{k}\dot{H}^{(s)}_T\right) \quad , \tag{5.223}$$

where $B^{(s)}$ is interpreted as the three-velocity amplitude of world lines orthogonal to the hypersurfaces $\tau = cst$. The expression ε_m is also equal to δ in any gauge where $v^{(s)} = B^{(s)}$ and thus when the matter world lines are orthogonal to $\tau = cst$ hypersurfaces; it is consequently regarded as the natural choice of energy density gauge-invariant amplitude from a matter perspective.

One can then show that the expression ε_g measures energy density perturbations relative to hypersurfaces whose normal unit vectors have zero shear, which is thus a geometrically selected hypersurface as close as possible to the Newtonian "time-slicing". One can then also show that once the perturbation is well within the particle horizon, the difference between the above two energy density expressions is relatively small.

Vector Perturbations

Similarly, for vector perturbations one can perform the following transformation from 5.23:

$$A_\alpha = B^{(v)}\mathcal{S}_\alpha \tag{5.224}$$

$$B_\alpha = -\frac{1}{k}H^{(v)}_{(T)}\mathcal{S}_\alpha \quad , \tag{5.225}$$

yielding the following form of the perturbed metric:

$$g_{0\alpha} = -S^2 B^{(v)}\mathcal{S}_\alpha \tag{5.226}$$

$$g_{\alpha\beta} = S^2\left[o_{\alpha\beta} + 2H^{(v)}_T\mathcal{S}_{\alpha\beta}\right] \quad . \tag{5.227}$$

The matter components are thus:

$$T^0_\alpha = (\rho_0 + p_0)(v^{(v)} - B^{(v)})\mathcal{S}_\alpha \tag{5.228}$$

$$T^\alpha_\beta = p_0(\delta^\alpha_\beta + \pi^{(v)}\mathcal{S}^\alpha_\beta) \quad , \tag{5.229}$$

where the superscript (v) is used here and henceforth to indicate quantities associated with vector perturbations. One then, as for scalar perturbations, considers the arbitrary co-ordinate transformation:

$$\overline{x}^{\alpha} = x^{\alpha} + L^{(1)}(\tau)\mathcal{S}^{\alpha} \quad , \tag{5.230}$$

yielding the transformations:

$$\begin{aligned} \overline{B}^{(v)} &= B^{(v)} + \dot{L}^{(v)} & (5.231)\\ \overline{H}_T^{(v)} &= H_T^{(v)} + kL^{(v)} & (5.232)\\ \overline{v}^{(v)} &= v^{(v)} + \dot{L}^{(v)} \quad , & (5.233)\end{aligned}$$

with the following gauge-invariant metric tensor amplitude:

$$\Psi = B^{(v)} - \frac{1}{k}\dot{H}_T^{(v)} \quad , \tag{5.234}$$

which, when multiplied by $\frac{k}{S}$, gives the amplitude of the shear of the normals to $\tau = cst$ hypersurfaces. As for the scalar perturbation case, one has the following two matter gauge-invariant velocity perturbations:

$$\begin{aligned} v_{(s)}^{(v)} &= v^{(v)} - \frac{1}{k}\dot{H}_T^{(v)} & (5.235)\\ v_{(c)}^{(v)} &= v^{(v)} - B^{(v)} & \\ &= v_s^{(v)} - \Psi \quad , & (5.236)\end{aligned}$$

where the first velocity is related, as before, to the shear; while the second one is relative to the normals of $\tau = cst$ hypersurfaces.

Tensor Perturbations

Finally, for tensor perturbations one has quantities which affect only the traceless parts of the spatial metric and stress tensors. Using the following transformation from the variables of 5.24:

$$h_{\alpha\beta} = -2H_{(T)}^{(t)}\mathcal{G}_{\alpha\beta} \quad , \tag{5.237}$$

one has the perturbed metric as:

$$\begin{aligned} g_{\alpha\beta} &= S^2\left[o_{\alpha\beta} + 2H_T^{(t)}(\tau)\mathcal{G}_{\alpha\beta}\right] & (5.238)\\ T_{\beta}^{\alpha} &= p_0\left[\delta_{\beta}^{\alpha} + \pi_T^{(t)}(\tau)\mathcal{G}_{\beta}^{\alpha}\right] \quad , & (5.239)\end{aligned}$$

such that P_0 is the background pressure; the superscript (t) is used here and henceforth to indicate quantities associated with tensorial perturbations. Since no three-vector or scalar can be formed from the tensor harmonics, the amplitudes $H_T^{(t)}$, $\pi_T^{(t)}$ are automatically gauge-invariant.

The Kinematic and Curvature Quantities in Terms of the Bardeen Variables

In terms of the above-defined perturbation variables, one can derive expressions for the various kinematic quantities; these are calculated from the definition of these quantities and direct substitution of the fully generalised perturbed metric into the field equations. The results are:

$$\begin{aligned}
\Theta &= \Theta_0 + \delta\Theta \\
&= \Theta_0 + \frac{3}{S}\left(H_L{}' - \frac{S'}{S}A + \frac{k}{3}v^{(s)}\right)\mathcal{Q} && (5.240)\\
a_\alpha &= \left(\Phi_{A;\alpha} + \frac{S'}{S}v^{(s)}_{(s)} - k\Phi_A\right)\mathcal{Q}_\alpha + \left(v^{(v)\prime}_{(c)} + \frac{S'}{S}v^{(v)}_{(c)}\right)\mathcal{S}_\alpha && (5.241)\\
\sigma_{\alpha\beta} &= S(-kv^{(s)}_{(s)}\mathcal{Q}_{\alpha\beta} - kv^{(v)}_{(s)}\mathcal{S}_{\alpha\beta} + H^{(t)}_T\mathcal{G}_{\alpha\beta}) && (5.242)\\
\omega_{\alpha\beta} &= Sv_{(c)}\mathcal{S}_{[\alpha;\beta]} \quad . && (5.243)
\end{aligned}$$

Similarly, for the curvature variables one has:

$$\begin{aligned}
{}^{(3)}R &= {}^{(3)}R_0 + \delta{}^{(3)}R \\
&= {}^{(3)}R_0 + \frac{1}{S^2}\left[4(k^2 - 3K)\left(\Phi_H - \frac{S'}{kS}v^{(s)}_{(s)}\right) - 12K\frac{S'}{kS}(v^{(s)} - B^{(s)})\right]\mathcal{Q} && (5.244)
\end{aligned}$$

$$\begin{aligned}
{}^{(3)}R_{\alpha\beta} &= \left[k\frac{S'}{S}v^{(s)}_{(s)} + \frac{k^2}{2}(\Phi_A - \Phi_B) + \frac{S^2}{S}p\pi^{(s)}_T\right]\mathcal{Q}_{\alpha\beta} + (k^2 + 2K)H^{(t)}_T\mathcal{G}_{\alpha\beta} \\
&\quad + \left(k\frac{S'}{S}v^{(v)}_{(s)} + \frac{1}{2}k\Psi' + \frac{S^2}{2}p\pi^{(v)}_T\right)\mathcal{S}_{\alpha\beta} - \frac{S'}{S}v_{(c)}\mathcal{S}_{[\alpha;\beta]} \quad , && (5.245)
\end{aligned}$$

while the electric and magnetic parts of the Weyl tensor are computed as:

$$\begin{aligned}
E_{\alpha\beta} &= \frac{1}{2}\left(k^2(\Phi_A - \Phi_H)\mathcal{Q}_{\alpha\beta} + k\Psi'\mathcal{S}_{\alpha\beta} - [H^{(t)\prime\prime}_T - (k^2 + 2K)H^{(t)}_T]\mathcal{G}_{\alpha\beta}\right) && (5.246)\\
H_{\alpha\beta} &= -S^{-2}\left(\frac{1}{2}\Psi\mathcal{S}^{\gamma\ ;\delta}_{;(\alpha}\eta_{\beta)0\gamma\delta} + H^{(t)}_T\mathcal{G}_{(\alpha}{}^{\gamma;\delta}\eta_{\beta)0\gamma\delta}\right) \quad . && (5.247)
\end{aligned}$$

For a general matter fluid, one can also naturally express the anisotropic terms in terms of the Bardeen variables, directly from the field equations:

$$\begin{aligned}
\kappa q_a &= -S\left[\kappa(p+\rho)v^{(s)}_{(s)} + \frac{2k}{S^2}\left(\Phi_H{}' - \frac{S'}{S}\Phi_A\right)\right]\mathcal{Q}_\alpha \\
&\quad + S\left[\kappa(p+\rho)v_{(c)} - \frac{1}{2S^2}(k^2 - 2K)\Psi\right]\mathcal{S}_\alpha && (5.248)\\
\kappa\pi_{\alpha\beta} &= -\left[\frac{k^2}{S^2}(\Phi_H + \Phi_A)\right]\mathcal{Q}_{\alpha\beta} + \left[\Psi' + 2\frac{S'}{S}\Psi\right]\mathcal{S}_{\alpha;\beta} \\
&\quad + \left[H^{(t)\prime\prime}_T + 2\frac{S'}{S}H^{(t)\prime}_T + (k^2 + 2K)H_T\right]\mathcal{G}_{\alpha\beta} \quad . && (5.249)
\end{aligned}$$

In all of the above, the covariant derivative has naturally been taken with respect to the spatial metric $o_{\alpha\beta}$.

The Perturbation Equations

Having thus formulated the basic gauge-invariant variables, one can proceed to derive the perturbation equations (to first order) in terms of these. The method is directly analogous to the Mukhanov *et al.* [17] approach, performing the first order variation of the Einstein Field Equations and then reformulating them gauge-invariantly. As mentioned earlier, there is no unique gauge-invariant form of the field equations, thus admitting a natural arbitrariness - indeed, any gauge-invariant choice will suffice. Bardeen chooses the following gauge-invariant forms:

$$\delta G^0_0 - \frac{3}{k^2}\mathcal{H}(\delta G^0_\alpha)^{;\alpha} \tag{5.250}$$

$$\delta G^\alpha_\beta - \frac{1}{3}\delta^\alpha_\beta \delta G^\gamma_\gamma \tag{5.251}$$

$$, \tag{5.252}$$

upon substitution into the perturbed field equations, the resulting equations (see [10]) for scalar perturbations, in terms of the above formulated gauge-invariant variables then transpire as:

$$\rho_0 \varepsilon_m = 2\frac{(k^2 - 3K)}{S^2}\Phi_H \tag{5.253}$$

$$p_0 \pi_T^{(s)} = -\frac{k^2}{S^2}(\Phi_A + \Phi_H) \tag{5.254}$$

$$\dot{v}^{(s)}_{(s)} + \frac{\dot{S}}{S}v^{(s)}_{(s)} = k\Phi_A + \frac{k(c_s^2\epsilon_m + W\eta)}{1+W} - \frac{\frac{2}{3}k(1-3\frac{K}{k^2})}{1+W}W\pi_T^{(s)} \tag{5.255}$$

$$\left[\rho_0 S^3 \varepsilon_m\right]^{\cdot} = -(1-3\frac{K}{k^2})(\rho_0 + p_0)S^3 k v^{(s)}_{(s)} - 2(1 - \frac{3K}{k^2})P_0 S^2 \dot{S}\pi_T^{(s)} \quad . \tag{5.256}$$

These can be combined to form:

$$\begin{aligned}(\rho_0 S^3\varepsilon_m)^{\cdot\cdot} &+ (1+3c_s^2)\tfrac{\dot{S}}{S}(\rho_0 S^3\varepsilon_m)^{\cdot} + [(k^2-3K)c_s^2 - \tfrac{1}{2}(\rho_0+p_0)S^2](\rho_0 S^3\varepsilon_m) \\ = &\ (1-3\tfrac{K}{k^2})\left\{-k^2(p_0 S^3\eta) + \tfrac{2}{3}[k^2 + 3(1+3c_s^2)K](p_0 S^3\pi_T^{(s)}) \right. \\ &\left. + 2(w - c_s^2)(\rho_0 S^2)(p_0 S^3 \pi_T^{(s)}) - 2\dot{S}(p_0 S^2\pi_T^{(s)})^{\cdot}\right\} \quad ,\end{aligned} \tag{5.257}$$

which is the central equation of this formalism. Alternatively, but using the same equations, one could have derived equations for either Φ_A or Φ_H. Indeed, on the previous section an evolution equation was derived for the variable $A = -\Phi_A \mathcal{Q}$ (equation 5.184).

Solutions to the Vector Perturbation Equations

Here one has:

$$\frac{1}{2}\frac{(k^2-2K)}{S^2}\Psi S_\alpha = (\rho_0+p_0)v_c S_\alpha \quad (5.258)$$

$$\dot{v}_c = \frac{\dot{S}}{S}(3c_s^2-1)v_c - k\frac{W}{1+W}\pi_T^{(v)} \quad , \quad (5.259)$$

for the *frame-dragging potential* Ψ defined through:

$$\Psi = 2(k^2-2K)^{-1}S^2(\rho_0+p_0)v_c \quad , \quad (5.260)$$

while the matter evolution equation is:

$$\dot{v}_c - \frac{\dot{S}}{S}(3c_s^2-1)v_c = -k\frac{W}{1+W}\pi_T^{(v)} \quad , \quad (5.261)$$

for the *vortical velocity amplitude* v_c.

Solutions to the Tensor Perturbation Equations

For tensor perturbations one has:

$$\frac{1}{S^2}\left(\ddot{H}_T^{(t)} + 2\frac{\dot{S}}{S}\dot{H}_T^{(t)} + (k^2+2K)H_T^{(t)}\right) = p_0\pi_T^{(t)} \quad , \quad (5.262)$$

where $H_T^{(t)}$ is interpreted as a gravitational wave amplitude.

Solutions to the Perturbation Equations

The standard $K=0$, $W=cst$ perfect fluid solutions are used for the background, as with the Lifshitz and Khalatnikov [4] approach, and substituted in for S, ρ_0 and p_0. The equations are here extended to allow for entropy and anisotropic pressure perturbations with arbitrary time dependence.

Scalar Perturbations Using then the background solutions together with standard differential equation techniques (*cf.* [10]), the solution to equation 5.257 transpires as:

$$\begin{aligned}\varepsilon_m \simeq \tfrac{(k\tau)^{2-\beta}}{2\beta+1}\Big\{&(k\tau)^\beta \textstyle\int_0^{k\tau} z^{-1}\left[\tfrac{2}{3}\tfrac{\beta+1}{2\beta-1}W\pi_T^{(s)} - W\eta\right]dz \\ &+(k\tau)^{-\beta-1}\textstyle\int_0^{k\tau} z^{2\beta}[-2\beta(2\beta+1)z^{-2}W\pi_T^{(s)} + W\eta]dz\Big\}\end{aligned} \quad , \quad (5.263)$$

such that:

$$\beta = \frac{2}{3W+1} \quad . \quad (5.264)$$

The form of this solution makes analysis of the contribution of the various quantities to the energy perturbation rather convenient; the following features are now readily evident:

- The entropy perturbation η contributes comparable amounts of *both* growing *and* decaying modes to ε_m.
- Even if the entropy perturbation turns off at some small 'time' $k\tau \ll 1$, by the time the perturbation comes within the particle horizon at $k\tau \sim 1$ the decaying mode becomes insignificant when compared with the growing mode.
- The anisotropic source term $\pi_T^{(s)}$, which is of order $(k\tau)^{-2}$ relative to the entropy perturbation source term, contributes *only* to the decaying mode at $k\tau \ll 1$; in fact, if this term then disappears at some later time $k\tau < 1$, then by the time $k\tau = 1$ the decaying mode contribution to ε_m becomes relatively small compared with the growing mode contribution.
- The perturbation in ε_m is always small if $W\pi_T^{(s)} \ll 1$; however, the zero-shear hypersurface distortion amplitude measure Φ_B is of order $(k\tau)^{-2}\epsilon_m$ and can thus be larger than one in magnitude while the anisotropic stress is present, even if this itself is small compared with the background energy density, as long as the anisotropic stress is present at every earlier time.

One can then argue heuristically that for $\Phi_B > 1$ one could have nonlinear perturbations. However, the question arises as to whether these nonlinear perturbations are physically significant, and could couple the large amplitude decaying mode to the growing mode, thus giving a large value for ε_m at the particle horizon compared with the previous maximum value of $W\pi_T^{(s)}$. This question is answered largely by considering the complete perturbation on different hypersurfaces while the anisotropic stress is present, or just after it is turned off. One can then consider, separately, the three hypersurface conditions of zero-shear, comoving, and uniform-Hubble-constant. This is done by transforming to the relevant gauge associated with each of these conditions, facilitated by the previously developed gauge-invariant variable formalism.

As previously mentioned, for zero-shear hypersurfaces the distortion amplitude Φ_B becomes a measure of the amplitude of metric perturbations, while Φ_A measures the fractional perturbation in the lapse function. Both these together with ϵ_m are of order $(k\tau)^{-2}W\pi_T^{(s)}$. Here the matter and geometry descriptions are independent of the particle horizon.

For comoving hypersurfaces all relative perturbation amplitudes are s mall up until the time the perturbation enters the particle horizon, provided the perturbation initially vanishes and subsequent stress perturbations are small: $W\eta \ll 1$, $W\pi_T^{(s)} \ll 1$.

However, for uniform-Hubble-constant hypersurfaces the description of both matter and geometry changes character depending upon whether the perturbation is larger or smaller than the particle horizon.

Due to the dependence on the particle horizon in the zero-shear and constant-Hubble-hypersurface cases, but independence for the co-moving hypersurface case, one needs to formulate a perturbation amplitude variable which is independent of hypersurface choice, but still carries physical significance. One such choice is the ratio $\mathcal{R}$ of matter shear rate to expansion rate:

$$\mathcal{R} = \frac{kS}{\dot{S}} v_{(s)}^{(s)} \quad . \tag{5.265}$$

The various amplitudes are then compared to this quantity, yielding the following observations:

- For the growing mode, the amplitude of the distortion in the intrinsic geometry is independent of time, and roughly the same in all three hypersurfaces.
- For the decaying mode, the uniform-Hubble-constant hypersurfaces are close to being synchronous when $(k\tau)^2 \ll 1$.

Vector Perturbations

The solution to the evolution equation 5.261 can be written as:

$$v_c = -\frac{1}{S^4(p_0 + \rho_0)} \int_0^{k\tau} S^4 p_0 \pi_T^{(v)} dz \tag{5.266}$$

Due to the fact that $S^4 p_0 \propto (k\tau)^{2(\beta-1)}$, this solution simplifies to:

$$v_c \sim -\frac{k\tau W}{(2\beta - 1)(W+1)} \pi_T^{(v)} \quad , \tag{5.267}$$

after the anisotropic stress has been on for a few expansion times; and:

$$v_c \propto \left[S^4(\rho_0 + p_0)\right]^{-1} \sim (k\tau)^{2(1-\beta)} \quad , \tag{5.268}$$

after it is turned off; while at the particle horizon v_c is comparable to the maximum previous value of $W\pi_T^{(v)}$. From this it can be seen that vector perturbations essentially exhibit no inherent instabilities and decay in expanding universe scenarios: this makes them of little interest for practical applications and considerations, as mentioned earlier in the Lifshitz and Khalatnikov analysis.

Tensor Perturbations

For tensor perturbations the solution to the evolution equation, namely the gravitational wave amplitude $H_T^{(t)}$, 5.262 is fairly standard, with the constraint that this cannot ever exceed in order of magnitude the maximum previous value of $W\pi_T^{(t)}$ if $H_T^{(t)}$ vanishes at the initial singularity, with $\frac{\rho_0 S^2}{k^2} \ll 1$.

Hence it is evident that tensorial perturbations amount to nothing more than gravitational waves which have a weak coupling to matter, and are quite often red-shifted away.

In order to gain further insight as to the behaviour and physicality of the gauge-invariantly defined variables of this approach, as well as to study the contribution of non-linear effects to the development of the perturbations, one can reformulate the above equations in terms of the ADM formalism. This assists particularly in the geometrical interpretation of the variables. However, the details which are quite involved and not relevant to the theme of this approach, need not be examined in full here. One can summarise the results thereof as follows:

- Scalar Perturbations: through the ADM formalism it becomes evident that the evolution equations for linear perturbations are the same as before. The nonlinearities are unable to alter the linear theory prediction for the spatial Ricci scalar appreciably, provided $W\pi_T^{(s)} \ll 1$; while the dynamical origin of the nonlinear corrections to the fractional energy density perturbation is the work done by the anisotropic stress on shearing the volume element.
- Vector Perturbations: The second order perturbation in the lapse function is small on uniform-Hubble-constant hypersurfaces; while the dynamically significant nonlinearities arise in the same manner as for the scalar perturbation case.
- Tensor Perturbations: The potentially largest nonlinear contribution to the fractional energy density perturbation is of the order of $\left(\frac{\pi_T^{(t)}}{\rho_0}\right)^2$, both when the anisotropic stress is present and at the particle horizon, regardless of whether the anisotropic stress is associated with any combination of scalar, vector and tensor harmonics.

Relevance of the Formalism

Having elucidated upon the details of Bardeen's formalism, one is now in a position for evaluation thereof. This is best summarised in terms of advantages and disadvantages:

Advantages:

- The elimination of gauge-dependence, rather than the awkward specification and ambiguous understanding thereof;
- The unambiguous derivation and formulation of fully gauge-independent variables;
- An approach which is conceptually straightforward and mathematically elegant.

Disadvantages:

- As can be seen from the above, the method is technically and computationally horrendous;
- The reverse connection back from the gauge-independent form to a specific gauge (co-ordinate system) wherein other physical processes may be calculated is not trivial.
- Bardeen's variables, although elegantly formulated, have a contrived aspect about them which often makes direct physical interpretation thereof somewhat convoluted and non-trivial.

Even with the formulation of gauge-invariant variables one realises from the preceding analysis that the interpretation of the resultant formalism is non-trivial. In the following chapter this problem is obviated by an alternative approach which develops a formalism yielding variables which are naturally gauge-invariant *ab initio*, thus rendering a clearer and less ambiguous interpretation.

Chapter 6

Covariant Perturbation Theory

"Many physical systems are computationally irreducible, so that their own evolution is effectively the most efficient procedure for determining their future."
S. Wolfram

6.1 Motivation: the Gauge Problem

To comprehend fully the motivation for the covariant formulation, one requires a closer look at the the gauge problem, the most noticeable drawback in the preceding theories. As considered in section 5.3, the metric perturbation approach depended extensively on a specifically chosen gauge, and as with Bardeen, erudite combinations of the relevant variables were chosen so as to be gauge-invariant. In this section a more natural approach will be developed which, by its form, is both automatically gauge-invariant, and more transparent in its physical interpretation.

One thus returns to the notion of a gauge transformation. In the context of section 5.3 it is appropriate to apply the *active* definition of a gauge transformation, whereby the correspondence between a point in the background space-time and the physical space-time is changed.

Hence, even if a quantity is a scalar under a co-ordinate transformation, the value of its associated perturbation will *not* be invariant under gauge transformations if the quantity is non-zero and co-ordinate dependent in the background space-time; this follows from the *Stewart-Walker lemma*, which will be derived later in the chapter. Consequently, any derived perturbation equations with residual gauge freedom will have solutions which can be transformed to zero by an appropriate gauge transformation.

It thus follows that, as with Bardeen [10], in order to avoid such gauge related problems arising, one should conceive of a formalism which describes the relevant perturbations in a gauge-*independent* way.

However, as a precursor to the *gauge-invariant* covariant formalism, both conceptually and historically, it is worthwhile to consider first the covariant approach on its own, without gauge considerations. In essence this amounts to a study of perturbations directly to the curvature of the space-time via the *linearised* field equations. In doing so, as Hawking [6] argues, one is not considering a mathematically abstract and physically ambiguous quantity such as the metric, but rather, physically observable quantities themselves: the dynamical variables.

Hence, in the ensuing sections the material will closely approximate the historical development of the field, commencing with the curvature perturbations of Hawking, proceeding thence to the co-ordinate free formalism of Olson, and finally culminating in the fully covariant formulations of Ellis *et al.* ([13], [14], [15], [16]) In addition to being historically correct, this progression is also logically consistent and self-contained.

6.2 Curvature Perturbations

This section closely follows Hawking [6], studying the generalised Field Equations linearised about an FRW, or conformally flat, Background.

6.2.1 Perturbation Equations

By the assumption 6.87 of conformally flat universes, it follows that:

$$\sigma^{ab} = \omega^{ab} = 0 \tag{6.1}$$
$$h_a^b \rho_{;b} = h_a^b \Theta_{;b} = 0 \quad . \tag{6.2}$$

Assuming also an equation of state $p = p(\mu)$, one has:

$$h_a^b p_{;b} = 0 = \dot{u}_a \quad . \tag{6.3}$$

This, together with the conformal assumption, yields the class of FRW universes. In the *perturbed* model these quantities will be assumed then to be *small*, so that all products of them can be ignored: this is essentially a linearisation procedure. It gives, to first order, the following perturbation evolution equations (see section 4.3; in particular, section 4.3.3):

$$E_{ab}{}^{;b} = \frac{1}{3} h_a^b \rho_{;b} \tag{6.4}$$
$$H_{ab}{}^{;b} = (\rho + p)\omega_a \tag{6.5}$$
$$-\frac{1}{2}(\rho + p)\sigma_{ab} = \dot{E}_{ab} + \Theta E_{ab} + h^f_{(a}\eta_{b)cde} u^c H_f{}^{d;e} \tag{6.6}$$
$$0 = \dot{H}_{ab} + \Theta H_{ab} - h^f_{(a}\eta_{b)cde} u^c E_f{}^{d;e} \tag{6.7}$$
$$\dot{\Theta} = -\frac{1}{3}\Theta^2 + \dot{u}_a{}^{;a} - \frac{1}{2}(\rho + p) \tag{6.8}$$
$$\dot{\omega}_{ab} = -\frac{2}{3}\omega_{ab}\Theta + \dot{u}_{[p;q]} h_a^p h_b^q \tag{6.9}$$
$$\dot{\sigma}_{ab} = E_{ab} - \frac{2}{3}\sigma_{ab}\Theta - \frac{1}{3}h_{ab}\dot{u}_c{}^{;c} + \dot{u}_{(p;q)} h_a^p h_b^q \quad . \tag{6.10}$$

6.2.2 The background Universe

For the equations of state $p = 0$ and $p = \frac{\rho}{3}$ the Raychaudhuri and Friedmann equations can easily be integrated, as done in section 4.6.3. Referring to this section, note that for the $p = 0$ case one has for the background:

$$\rho' + \rho\Theta = 0 \tag{6.11}$$

$$\Theta' + \frac{1}{3}\Theta^2 + \frac{1}{2}\rho = 0 \ . \tag{6.12}$$

Hence, on account of there being no spatial derivatives, each region of the universe evolves independently. Consequently, perturbations will exist as varying values of the total energy E in different regions, although the universal average will still be E. From the above results it then transpires that for a universe having $E > 0$, any small perturbation will continue to expand indefinitely, not contracting to form a galaxy; for $E < 0$ one can have a small perturbation contracting to form a galaxy. However, the case $E = 0$ is unstable: one can study this by considering a small region within such a universe having a small energy perturbation $\delta E < 0$. Applying then the $E < 0$ scenario above to this region, Taylor expanding 4.132 and using the normalisation $M = \frac{3}{|E|}$, one obtains:

$$S = \frac{1}{4\delta E}\left(t^2 - \frac{t^3}{12} + \ldots\right) \ , \tag{6.13}$$

and similarly:

$$\tau = \frac{1}{12\delta E}\left(t^3 - \frac{t^5}{20} + \ldots\right) \ . \tag{6.14}$$

Consequently:

$$\rho = \frac{4}{3\tau^2}\left[1 + \left(\frac{12\delta E}{5}\right)^{\frac{2}{3}}\tau^{\frac{2}{3}} + \ldots\right] \ . \tag{6.15}$$

Hence the perturbation grows only as $\tau^{\frac{2}{3}}$ which is not fast enough for galaxy formation. A similar analysis can be performed for the case $p = \frac{\rho}{3}$.

6.2.3 Rotational and Density Perturbations

Using the momentum conservation equation with the vorticity evolution equation 4.36, one obtains a rotational perturbation equation:

$$\dot{\omega} = -\omega\left(\frac{2}{3}\Theta + \frac{\dot{p}}{\rho + p}\right) \ . \tag{6.16}$$

Hence, considering again the cases $p = \frac{\rho}{3}$ and $p = 0$ one obtains the vorticity magnitude ω as being inversely proportional to the scale factor, and scale factor squared respectively, as can readily be obtained from equation 6.16 after elementary integration. This shows that rotation, via a vorticity perturbation ω, dies away as the universe expands, and is thus an expression of the conservation of angular momentum in an expanding universe. For the energy density perturbations one has firstly, for the $p = 0$ case:

$$\dot{\rho} = -\rho\Theta \tag{6.17}$$

$$\dot{\Theta} = -\frac{1}{3}\Theta^2 - \frac{1}{2}\rho \quad . \tag{6.18}$$

As a result of these equations containing no spatial derivatives, the density evolutions in each particular region will be independent. The above solutions for S and t then show that, for $E > 0$, one will have $\delta\rho > 0$ which will continue to expand, not contracting to form a galaxy; for $E < 0$ one *can* have $\delta\rho$ contracting. This contraction will begin just before the whole universe starts contracting. The only real instability occurs when $E = 0$: In this case one has $\delta\rho \propto t^{\frac{2}{3}}$ which is not fast enough to produce galaxies even from statistical fluctuations.

For $p = \frac{\rho}{3}$ one also has the result that a perturbation can only contract if E is negative. In this case the same results as for Lifshitz and Khalatnikov are obtained.

The Steady-State Universe

For this scenario, following Hoyle and Narlikar [5], one has to add extra terms to the energy-momentum tensor:

$$T_{ab} = \rho u_a u_b + p h_{ab} + C_a C_b + \frac{1}{2} g_{ab} C_d C^d \quad , \tag{6.19}$$

where the quantities satisfy:

$$C_a = C_{,a} \tag{6.20}$$

$$C^a_{\ ;a} = -\mathcal{J}^a_{\ ;a} \tag{6.21}$$

$$\mathcal{J}_a = (\rho + p) u_a \quad , \tag{6.22}$$

$\mathcal{J}_a$ being the energy current. Using the corresponding energy and momentum conservation equations together with the kinematic evolution equations, one obtains the following:

$$C^a_{\ ;a} = -\Theta\left[1 - \frac{p}{\rho' + p' + \Theta(\rho + p)}\right] + \left[\frac{p'}{\rho' + p' + \Theta(\rho + p)}\right]' \quad . \tag{6.23}$$

From which it follows that for $\rho \gg p$ one has

$$\rho' + p' = \Theta[1 - (\rho + p)] \quad , \tag{6.24}$$

which results in $\rho + p \rightarrow 1$. Hence small density perturbations die away. From the kinematic evolution equations this in turn results in the rotational perturbations also dying away, as well as $\Theta \rightarrow \sqrt{3(\frac{1}{2} - p)}$. This naturally means that galaxies cannot form from the growth of small perturbations in a steady-state universe.

Gravitational Waves

One now considers perturbations to the Weyl tensor that do not arise from density or rotational perturbations; this should give rise to perturbations of the "free" gravitational field, and hence gravitational waves. From the kinematic evolution equations, one thus has:

$$E^{ab}_{\ ;b} = H^{ab}_{\ ;b} = 0 \quad . \tag{6.25}$$

This, together with the evolution equations for E_{ab} and H^{ab} yields:

$$\ddot{E}_{ab} - {}^{(3)}\nabla^2 E_{ab} + \frac{7}{3}\dot{E}_{ab}\Theta + E_{ab}[\Theta' + \frac{4}{3}\Theta^2 + \frac{1}{3}(\rho - 3p)] + \sigma_{ab}[\frac{1}{3}\Theta(\rho + p) + \frac{1}{2}(\rho' + p')] = 0 \quad , \tag{6.26}$$

where ${}^{(3)}\nabla^2$ is the spatial Laplacian in the hypersurface $t = cst$, defined as before:

$${}^{(3)}\nabla^2 E_{ab} \equiv h^{ki} h^f_a h^g_b (h^c_f h^d_g h^e_k E_{cd;e})_{;i} \quad . \tag{6.27}$$

Because in the linearised theory one has ${}^{(3)}\nabla^2 E_{ab} = 0$, and which is retracted from the above equation in empty space when $\Theta = \sigma_{ab} = 0$; one can attempt to solve 6.26 by writing E_{ab} as the sum of the eigenfunctions of $\Delta^{(3)}$:

$$E_{ab} = \sum_n A^{(n)} V^{(n)}_{ab} \quad , \tag{6.28}$$

such that:

$$\dot{V}^{(n)}_{ab} = 0 \tag{6.29}$$

$${}^{(3)}\nabla^2 V^{(n)}_{ab} = -\frac{n^2}{S^2} V^{(n)}_{ab} \quad , \tag{6.30}$$

which thus yields:

$$\ddot{E}_{ab} = \frac{1}{S^2} \sum_n V^{(n)}_{ab} \left[\frac{d^2 A^{(n)}}{d\tau^2} - \frac{1}{S}\frac{dS}{d\tau}\frac{dA^{(n)}}{d\tau} \right] \quad . \tag{6.31}$$

Similarly:

$$\sigma_{ab} = \sum_n V^{(n)}_{ab} D^{(n)} \quad , \tag{6.32}$$

which, from the evolution equation for the shear, results in:

$$\frac{dD^{(n)}}{d\tau} = SA^{(n)} - 2D^{(n)}\frac{1}{S}\frac{dS}{d\tau} \quad . \tag{6.33}$$

Substituting these into 6.26 one then obtains:

$$\frac{d^2A^{(n)}}{d\tau^2} + \frac{6}{S}\frac{dS}{d\tau}\frac{dA^{(n)}}{d\tau} + A^{(n)}\left[n^2 + \frac{3}{S}\frac{d^2S}{d\tau^2} + \frac{6}{S^2}\left(\frac{dS}{d\tau}\right)^2 + \frac{\rho+3p}{3}S^2\right]$$
$$+D^{(n)}\left[(\rho+p)\frac{dS}{d\tau} + \frac{1}{2}(\rho'+p')S^2\right] = 0 \quad , \tag{6.34}$$
$$\tag{6.35}$$

which can be differentiated again and substituted into for $\frac{dD^{(n)}}{d\tau}$; upon which, for $n \gg 1$ and $S \gg \frac{1}{n^2}$, the following solution obtains:

$$A^{(n)} \simeq \frac{1}{S^3}e^{in\tau} \quad . \tag{6.36}$$

Interpreting this, one sees that the gravitational field E_{ab} decreases as $\frac{1}{S}$, while the gravitational energy $\frac{1}{2}(E_{ab}E^{ab} + H_{ab}H^{ab})$ decreases as $\frac{1}{S^{-6}}$.

This results in the interaction between the "Weyl field" and gravitational radiation being equal and opposite to that of matter; resulting in there being no net reaction with matter. Hence gravitational waves will not be absorbed by a perfect fluid.

Absorption of Gravitational Waves

Assuming there to be a small amount of viscosity in the perfect fluid model, one may represent this through the introduction of an additional term $\lambda\sigma_{ab}$ to the energy-momentum tensor such that λ is the co-efficient of viscosity. Combined with the amended conservation equations (which contain extra terms of $-2\lambda\sigma^2$ and $\lambda\sigma_{cb}^{\ \ ;b}h_a^c$ respectively) that thus result, the evolution equations for E_{ab} and H_{ab} obtain as:

$$\begin{aligned} \dot{E}_{ab} + E_{ab}\Theta + h^f_{(a}\eta_{b)cde}u^cH_f^{\ d;e} &= -\frac{1}{2}(\rho+p)\sigma_{ab} \\ &\quad -\frac{1}{2}\lambda(E_{ab} - \frac{1}{3}\sigma_{ab}\Theta) \\ \dot{H}_{ab} + H_{ab}\Theta - h^f_{(a}\eta_{b)cde}u^cE_f^{\ d;e} &= -\frac{1}{2}\lambda H_{ab} \quad . \end{aligned} \tag{6.37}$$

The extra terms on the right hand side of these equations can then be interpreted as being analogous to conduction terms in Maxwell's equations, causing the wave to decrease by a factor of $e^{-\frac{\lambda}{2}t}$. If one now ignores the expansion Θ and postulates a form for the gravitational wave as:

$$E_{ab} = E_{ab}^{(0)}e^{i\nu t} \quad , \tag{6.38}$$

then the wave will be absorbed in a characteristic time of $\frac{2}{\lambda}$ independent of the frequency. Comparing this with the energy conservation equation for the above scenario, one obtains that the rate of gain of rest-mass energy of the matter will be $2\lambda\sigma^2$ which, by the evolution equation for the shear, is equal to $2\frac{\lambda E_0{}^2}{\nu^2}$ where ν is the gravitational frequency; thus resulting in the available energy in the wave being $\frac{4E_0{}^2}{\nu^2}$. This means that the density of the available energy of gravitational radiation will decrease as $\frac{1}{S^4}$ in an expanding universe; thus it is apparent that gravitational radiation behaves very much as that of standard radiation fields. Now, by the expansion evolution equation, with $\sigma^2 = \frac{E_0{}^2}{\nu^2}$ and a radiation energy density of $4\frac{E_0{}^2}{\nu^2}$, one obtains:

$$\Theta' = -\frac{1}{3}\Theta^2 - \frac{1}{2}\rho_G - \frac{1}{2}(\rho + 3p) \quad , \tag{6.39}$$

where ρ_G is the gravitational energy density. This means that gravitational radiation has an *active attractive gravitational effect*. Such gravitational radiation emitted in an expanding universe will eventually be absorbed by other matter if $\int \lambda dt$ diverges; which will certainly be true for the steady-state universe where λ is constant. Generally, in evolutionary universes λ will be a function of time.

Advantages and Disadvantages

The above approach of Hawking is seen to be more physically intuitive in that the operative dynamical physical variables which one would expect to have an explicit influence on structure formation in the universe are perturbed directly through a perturbation to the curvature, without recourse to the underlying metric. This is in antithesis to the metric perturbation approach where the metric is the pivotal mathematical quantity, the perturbation to which indirectly determines the perturbations of all physical quantities via the field equations; in the curvature perturbation approach one has no need to consider perturbations to the metric at all.

The outstanding problem though, is still that of gauge-invariance; although the above formalism is covariant, it still contains implicitly the arbitrariness of gauge choice. This problem will be obviated by application of the Stewart-Walker lemma later in the chapter. In addition, several errors have been detected in Hawking's derivations, particularly in the gravitational analysis: notably those pointed out by Olson [9], which will be mentioned shortly, and that of Dunsby *et al.* [22]. In leading up to the gauge problem, one can consider a synthesis of the above Hawking approach and that of Lifshitz and Khalatnikov through the following co-ordinate-free approach of Olson [9].

6.3 A Co-ordinate-Free Formulation

Using the continuity and Raychaudhuri equations one can derive a co-ordinate independent evolution equation for linearised energy-density perturbations in a fluid co-moving frame in a flat FRW universe; an equation which is in terms of *proper* time. The approach here, due to Olson [9], initially follows Hawking [6]. Assuming, as before, a perfect fluid scenario for a FRLW background model universe one has the standard continuity and linearised Raychaudhuri equations:

$$0 = \dot{\rho} + (\rho + p)\Theta \tag{6.40}$$

$$\dot{\Theta} = -\frac{1}{3}\Theta^2 - \frac{1}{2}k(\rho + 3p) + \dot{u}^a{}_{;a} \quad , \tag{6.41}$$

and from the momentum conservation:

$$\dot{u}^a = -\frac{^{(3)}\nabla^a p}{\rho + p} \quad . \tag{6.42}$$

Similarly, one has the following linearised expression for the hypersurface-projected Ricci scalar $^{(3)}R$:

$$^{(3)}R = 2(\rho - \frac{1}{3}\Theta^2) \quad . \tag{6.43}$$

In the background model, as before, one has:

$$\rho = 3\left(\frac{\dot{S}}{S}\right)^2 \tag{6.44}$$

$$\Theta = 3\frac{\dot{S}}{S} \tag{6.45}$$

$$0 = \omega, \sigma, \dot{u}^a \quad , \tag{6.46}$$

where the first equation follows from 4.122 with $K = \Lambda = 0$. Note that the linearisation is such that products and powers of ω, σ and $\dot{u}^a$ are dropped, as these quantities are assumed to be very small in the real universe which is *almost*-Robertson-Walker.

6.3.1 The linearised equations

One can express the perturbed 'real universe' energy density ρ in terms of the background energy density ρ_0 and a perturbation $\delta\rho$ as follows:

$$\rho \equiv \rho_0(t)(1 + \frac{\delta\rho}{\rho}) \quad , \tag{6.47}$$

where ρ_0 is that function of co-moving proper time which corresponds to the density in the flat background universe at the proper time t after the big bang, and defined through equations 6.44 and 6.45 as $\rho_0 = \frac{1}{3}\Theta_0^2$, the 'zero' subscript indicating a background quantity. Bearing this latter expression in mind, one can define a related small quantity ϵ, which can be thought of in the linearised theory as a 'spatial curvature perturbation', as follows:

$$\rho \equiv \frac{1}{3}\theta^2(1 + \epsilon) \quad , \tag{6.48}$$

this definition being made in the *real*, perturbed universe. In irrotational models it becomes, to linear order $\epsilon = \frac{^{(3)}R}{2\rho}$ using equations 6.43 and 6.48. In all subsequent work, The evolution equations for the dynamical quantities are linearised; hence, differentiating 6.48 with respect to proper time and substituting into the result for $\dot{\rho}, \dot{\Theta}, \rho$ and $\dot{u}^a_{\ a}$ from equations 6.40, 6.41, 6.48 and 6.42 respectively, and linearising, one obtains:

$$\Theta\dot{\epsilon} = (\rho + p)\epsilon + 2\frac{v_s^2}{\rho + p}X^a_{\ ;a} \quad , \tag{6.49}$$

while taking first and second spatial derivatives of 6.40 and linearising, one obtains the evolution equation:

$$\Theta\dot{X}^a_{;a} = -X^a_{;a}[5\rho + \frac{3}{2}(\rho + p)] - \frac{3}{2}\rho(\rho + p)\frac{n^2}{S^2}\epsilon \quad , \tag{6.50}$$

such that $X^a = h^{ab}\rho_{,b}$, n^2 is constant, and a plane wave expansion has been made of ϵ. As $X^a_{;a}$ can be eliminated from the above, these two equations are sufficient to determine ϵ. One can also then relate $\frac{\delta\rho}{\rho}$ to ϵ by the continuity equation, depending on the equation of state. Hence, as with Hawking, one can consider pertinent equations of state, thus enabling integration of the above equations. For an equation of state $p = \alpha\rho$ one has $S(t) =\sim t^{\frac{2}{3+3\alpha}}$ and $\rho_0(t) = \frac{4}{3}\frac{1}{t^2(1+\alpha)}$, and consequently:

$$\epsilon = \left(t\frac{\delta\rho}{\rho}\right)^{\cdot} \quad . \tag{6.51}$$

Integrating along the world line of a fundamental observer, one obtains:

$$\frac{\delta\rho}{\rho} = \frac{1}{t}\left(C + \int \epsilon dt\right) , \tag{6.52}$$

such that C is an arbitrary constant of integration, which thus gives rise to an inherent arbitrariness in gauge; hence the gauge problem is still not fully eliminated in this formalism.

6.3.2 Consequences

Upon integrating the above linearised equations, having assumed an equation of state $p = \alpha\rho$, one discovers a non-oscillatory decaying mode $\frac{\delta\rho}{\rho} \sim \frac{1}{t}$ for the cases: $p = 0$, $p = \rho$, $p = \frac{\rho}{3}$. It is readily evident from equation 6.51 that only for this decaying mode does the quantity ϵ vanish; however, the interpretation of this mode is complicated by the additional degree of freedom inherent in the proper, co-moving time t in the definition of the density perturbation. This follows from the fact that the quantity:

$$t^\dagger = t + t_0 \quad , \tag{6.53}$$

where t_0 satisfies $t_{,a}u^a = 0$, and *also* represents co-moving proper time. Consequently, for the equation of state $p = \alpha\rho$, this equivalent proper time results in:

$$\frac{\delta\rho}{\rho} \rightarrow \frac{\delta\rho}{\rho}^\dagger = \frac{\delta\rho}{\rho} + 2\frac{t_0}{t} \quad . \tag{6.54}$$

Hence it follows that, by a suitable choice of t_0, one can have the density perturbation vanishing, while the quantity ϵ remains unchanged. Hence one can view the decaying mode as being 'real' if it can be removed from $\frac{\delta\rho}{\rho}$ only by a suitably chosen t_0, but *not* from the other quantities; and 'fictitious' if it can be removed entirely from all quantities.

Advantages and Disadvantages

The above analysis of Olson [9] is in keeping with that of Hawking [6] in so far as a non-metric perturbation is concerned, but focuses specifically on density perturbations. Naturally one would desire a formalism which incorporates perturbations to *all* the dynamical variables: this will be generalised later. Also, although fully covariant in the direct treatment of curvature perturbations as opposed to metric perturbations it is, nonetheless, still not gauge-invariant, as the constant of integration in equation 6.52 demonstrates . Consequently, the same problems related to gauge choice occur, as with metric perturbations. In the following section a covariant approach which is in addition gauge-invariant will be pursued: the underpinning principle of this formalism resides on the following lemma, due to Stewart and Walker [8]; a result which is particularly powerful due to its generality and co-ordinate independent formulation.

6.4 The Stewart-Walker Lemma

Consider a single-parameter family of four-dimensional manifolds $\mathcal{M}_\epsilon$ embedded within a five-dimensional manifold $\mathcal{N}$. Each of the $\mathcal{M}_\epsilon$ represents a space-time, the *base* or *unperturbed* manifold $\mathcal{M}_0$ while ϵ is a smallness parameter.

Secondly, consider a point identification map $P_\epsilon : \mathcal{M}_0 \rightarrow \mathcal{M}_\epsilon$ which identifies each point in the perturbed manifold with a given point in the base manifold. This is achieved through the additional specification of a vector field $\mathbf{X}$ defined on $\mathcal{N}$ such that $\mathbf{X}$ is everywhere transverse (i.e. orthogonal) to the embeddings of the $\mathcal{M}_\epsilon$; while points lying on the same integral curve γ of $\mathbf{X}$ are regarded as the same point. Conveniently parametrising γ as $\gamma(\epsilon)$, one can define:

$$X^A = \frac{dx^A}{d\epsilon}, \quad A = 0, \ldots 4 \quad . \tag{6.55}$$

Letting Q_ϵ be some arbitrary geometrical field defined on each $\mathcal{M}_\epsilon$, one can expand Q_ϵ for small ϵ as a Taylor series along γ :

$$\Pi^{\dagger}_{\epsilon}(Q_\epsilon) = Q_0 + \epsilon(\mathcal{L}_{\mathbf{X}}Q)_0 + \mathcal{O}(\epsilon^2) \tag{6.56}$$
$$= Q_0 + \delta Q + \mathcal{O}(\epsilon^2) \quad , \tag{6.57}$$

where $\mathcal{L}_{\mathbf{X}}$ is the Lie derivative in the direction of $\mathbf{X}$, while $\Pi^{\dagger}_{\epsilon}$ is the *pull-back* of $\mathcal{M}_\epsilon$ to $\mathcal{M}_0$. The quantity $\delta Q = \epsilon(\mathcal{L}_{\mathbf{X}}Q)_0$ is now readily seen to be the *linear* perturbation of Q_0 to Q, and depends on the choice of $\mathbf{X}$, which is arbitrary. This naturally implies that the *choice of* $\mathbf{X}$ *corresponds to a choice of gauge.* The difference between two such choices $\mathbf{X}, \mathbf{Y}$ is then:

$$\Delta\delta Q = \epsilon(\mathcal{L}_{\mathbf{X}-\mathbf{Y}}Q)_0 \quad , \tag{6.58}$$

where the quantity $\xi = \epsilon(\mathbf{X} - \mathbf{Y})$ is a vector field in each $\mathcal{M}_\epsilon$ (which is trivial to show), yielding in turn:

$$\Delta\delta Q = \mathcal{L}_{\epsilon\xi_0}Q_0 \quad . \tag{6.59}$$

It now follows naturally that δQ will be gauge-independent ($\Delta \delta Q = 0$) only if $\mathcal{L}_{\epsilon \xi_0} Q_0 = 0$; that is, only if Q_0 is either zero, a constant scalar field, or a linear combination of Krönecker deltas with constant co-efficients. This is the Stewart-Walker Lemma [8].

6.5 Obviation of the Gauge Problem

By the above, one has the following simple way of circumventing the gauge problem: one chooses an idealised background model of the universe, and considers some covariant formulation (such as a covariant derivative), of a relevant quantity, *which vanishes therein*; this same covariantly formulated quantity will then be fully gauge-invariant in the real, physical universe by the Stewart-Walker lemma as formulated above. One would then naturally find the requisite evolution equation for this gauge-invariant and *covariant* quantity in order to analyse its behaviour. This is the fundamental precept of the covariant approach. For reasons discussed in the previous chapter, a physically suitable background choice would be that of FRW; this assumption will be maintained throughout the remainder of the analysis, although the full generality of the Stewart-Walker Lemma should be appreciated.

In application of the above lemma, one could consider some scalar density $\phi(t)$ which is a *function of time only* in the background universe; the reason for the singular time functional dependence is that this would result in the spatial gradient of ϕ vanishing in the background universe, thus yielding the realistic counterpart in the real, 'lumpy' universe as being fully gauge-invariant by the Stewart-Walker lemma.

6.6 The Covariant Approach

The approach of this section follows the work of Ellis *et al.* ([13], [14], [15], [16]), and essentially generalises the previous non-metric approaches by formulating general covariant quantities which vanish in the background, and are thus additionally gauge-invariant by the Stewart-Walker Lemma. In the process, evolution equations for these quantities are obtained, and subsequently linearised to yield propagation equations appropriate to the scenari o of an *almost* Robertson-Walker universe. Solutions to these are obtained and compared with the standard gauge-related theory, and gauge-invariant metric theory of Bardeen; solutions which recover the usual growing and decaying modes.

The formalism will then be extended to incorporate the more physically meaningful and useful multi-component matter case, in the process deriving equations which provide insight into the interactions between *different* matter types within an overall matter description of the universe.

6.6.1 Gauge-Invariant Variables

By definition, in the class of FRW background models the following hold true:

$$\sigma_{ab} = \omega_{ab} = \dot{u}_a = 0 \quad . \tag{6.60}$$

Consequently, any scalar field ϕ defined in the background universe must be a function of time only; if it were a function of the spatial variables, both the isotropy and homogeneity prerequisites would be violated:

$$\phi_{,a} \equiv 0, \Rightarrow \phi = \phi(t) \quad . \tag{6.61}$$

The principal such scalar fields which determine the matter description are:

$$\rho = \rho(t) \tag{6.62}$$
$$p = p(t) \tag{6.63}$$
$$\Theta = \Theta(t) \quad , \tag{6.64}$$

which thus have vanishing spatial derivatives:

$$X_a \equiv {}^{(3)}\nabla_a \rho = 0 \tag{6.65}$$
$$Y_a \equiv {}^{(3)}\nabla_a p = 0 \tag{6.66}$$
$$Z_a \equiv {}^{(3)}\nabla_a \Theta = 0 \tag{6.67}$$
$$S_a \equiv {}^{(3)}\nabla_a s = 0 \quad , \tag{6.68}$$

where s is the entropy density. The following useful gauge-invariant quantities then transpire:

$$\mathcal{D}_a = S\frac{X_a}{\rho} \tag{6.69}$$
$$\Phi_a = S^3 \kappa^{(3)}\nabla_a \rho \tag{6.70}$$
$$\mathcal{Y}_a = S\frac{Y_a}{p} \tag{6.71}$$
$$\mathcal{Z}_a = SZ_a \tag{6.72}$$
$$\mathcal{E}_a = \frac{S}{p}\left(\frac{\partial p}{\partial s}\right)_{(\rho)} S_a \quad , \tag{6.73}$$

where the variable $\mathcal{E}_a$ is defined assuming an equation of state $p = p(\rho, s)$. These are naturally gauge-invariant by the Stewart-Walker Lemma, and are by form dimensionless, except for the Bardeen-like variable Φ_a. These variables encapsulate the physical matter properties of the universe, and are variables which are most easily physically verifiable and will thus feature prominently in the subsequent analysis.

Naturally, the Ricci scalar R must, by the selfsame argument, have a vanishing spatial gradient in the background FRW space-time; hence this generates a covariant, gauge-invariant variable:

$$\mathcal{C}_a = S^{3(3)}\nabla_a R \quad , \tag{6.74}$$

with the following associated quantity which is conserved in certain cases:

$$\tilde{\mathcal{C}}_a = \mathcal{C}_a - \frac{4K}{1+W}\mathcal{D}_a \quad , \tag{6.75}$$

such that $K = 0, \pm 1$ depending on whether the universe is open, closed or flat. Associated with these vectorial quantities one can define the following scalar gauge-invariant quantities:

$$\mathcal{D} \equiv S^{(3)}\nabla^a \mathcal{D}_a \tag{6.76}$$
$$\mathcal{Z} \equiv S^{(3)}\nabla^a \mathcal{Z}_a \tag{6.77}$$
$$\mathcal{C} \equiv S^{(3)}\nabla^a \mathcal{C}_a \tag{6.78}$$
$$\tilde{\mathcal{C}} \equiv S^{(3)}\nabla^a \tilde{\mathcal{C}}_a \tag{6.79}$$
$$\mathcal{E} \equiv S^{(3)}\nabla^a \mathcal{E}_a \quad . \tag{6.80}$$
$$\tag{6.81}$$

In addition to the above quantities, others which vanish in the FRW background and are thus gauge-invariant in the real universe are:

$$\omega_{ab} \equiv h_a^c h_b^d u_{[c;d]} \tag{6.82}$$
$$\sigma_{ab} \equiv h_a^c h_b^d u_{(c;d)} - \frac{1}{3} h_{ab} u^c{}_{;c} \tag{6.83}$$
$$\dot{u}^a \equiv u^a{}_{;b} u^b \tag{6.84}$$
$$q_a \equiv -h_a^c u^d T_{cd} \tag{6.85}$$
$$\pi_{ab} \equiv h_a^c h_b^d T_{cd} - \frac{1}{3} h^{cd} T_{cd} h_{ab} \quad . \tag{6.86}$$

Another property of FRW universes is that they are conformally flat; that is, the Weyl conformal tensor vanishes:

$$C_{abcd} = 0 \quad . \tag{6.87}$$

Consequently, the conformal tensor and any decomposition thereof will be gauge-invariant *in any reference frame*, as opposed to the shear and vorticity and spatial scalar gradients which are only gauge-invariant in the matter rest frame. Consequently one has that the electric and magnetic parts of the Weyl tensor must be gauge-invariant:

$$E_{ab} \equiv C_{acbd} u^c u^d \tag{6.88}$$
$$H_{ab} \equiv \frac{1}{2} C_{acst} u^c u^d \eta^{st}{}_{bd} \quad . \tag{6.89}$$

6.6.2 The Linearised Field Equations

The gauge-invariant variables defined above will be the operative perturbation quantities in the following analysis, and are consequently regarded as being suitably 'small' deviations from the background. This then naturally defines a linearisation procedure whereby terms quadratic, or of higher order, in the gauge-invariant variables appearing in the field equations are discarded.

The Conservation Equations

The energy and momentum conservation equations are linearised to yield:

$$\dot{\rho} + 3h\mathcal{H} + h^{(3)}\nabla^a \Psi_a = 0 \quad (6.90)$$

$$h\dot{u}_a + {}^{(3)}\nabla_a p + h(F_a + \Pi_a) = 0 \ , \quad (6.91)$$

$$(6.92)$$

such that

$$h = \rho + p \quad (6.93)$$

$$\Psi_a = \frac{q_a}{h} \quad (6.94)$$

$$F_a = \Psi_a - (3c_s^2 - 1)\mathcal{H}\Psi_a \quad (6.95)$$

$$\Pi_a = \frac{1}{h}{}^{(3)}\nabla^b \pi_{ab} \ , \quad (6.96)$$

with the associated quantities:

$$F \equiv S^{(3)}\nabla^a F_a \quad (6.97)$$

$$\Pi \equiv S^{(3)}\nabla^a \Pi_a \ , \quad (6.98)$$

$$(6.99)$$

while the Raychaudhuri equation linearises to:

$$3\dot{\mathcal{H}} + 3\mathcal{H}^2 - D + \frac{1}{2}\kappa(\rho + 3p) - \Lambda = 0 \ , \quad (6.100)$$

such that $D \equiv \dot{u}^a_{;a}$. The linearised evolution equations for the shear and vorticity transpire as:

$$\dot{\sigma_{ab}} + E_{ab} + 2\mathcal{H}\sigma_{ab} + \frac{c_s^2}{S(1+W)}\left({}^{(3)}\nabla_{(a}\mathcal{D}_{b)} - \frac{1}{3}h_{ab}{}^{(3)}\nabla^c\mathcal{D}_c\right)$$
$$= -\frac{W}{S(1+W)}\left({}^{(3)}\nabla_{(a}\mathcal{E}_{b)} - \frac{1}{3}h_{ab}{}^{(3)}\nabla^c\nabla_c\right)$$
$$-\left({}^{(3)}\nabla_{(a}(F_{b)} + \Pi_{b)}) - \frac{1}{3}h_{ab}{}^{(3)}\nabla^c(F_c + \Pi_c)\right)$$
$$(6.101)$$

$$\dot{\omega_{ab}} + 2\mathcal{H}\omega_{ab} = \left[3c_s^2\mathcal{H} - \left(\frac{\partial p}{\partial s}\right)_\rho \frac{\dot{s}}{\rho + p}\right]\omega_{ab} - {}^{(3)}\nabla_{[b}(F_{a]} + \Pi_{a]}) \ ,$$
$$(6.102)$$

while the evolution equations for the electric and magnetic parts of the Weyl tensor transpire as:

$$\begin{aligned} \dot{E}_{ab} + 3\mathcal{H}E_{ab} + h^{f}_{(a}\eta_{b)cde}u^{c}H_{f}^{\ d;e} + \frac{1}{2}\kappa h\sigma_{ab} &= -\frac{1}{2}\kappa({}^{(3)}\nabla_{(a}q_{b)} - \frac{1}{3}h_{ab}{}^{(3)}\nabla^{c}q_{c}) \\ &\quad -\frac{1}{2}\kappa(\mathcal{H}\pi_{ab} + \dot{\pi}_{ab}) \qquad (6.103) \\ \dot{H}_{ab} + 3\mathcal{H}H_{ab} - h^{f}_{(a}\eta_{b)cde}u^{c}E_{f}^{\ d;e} &= \frac{1}{2}\kappa h_{c(a}\eta_{b)def}u^{d}\pi^{ce;f} \quad . \qquad (6.104) \end{aligned}$$

Constraint Equations

From the Friedmann equation one obtains the linearised energy constraint:

$$\frac{1}{6}{}^{(3)}R = -\mathcal{H}^{2} + \frac{1}{3}\rho + \frac{1}{3}\Lambda \quad . \qquad (6.105)$$

The $(0,\nu)$ equations are linearised to yield:

$$\begin{aligned} \frac{2}{3}\mathcal{Z}_{a} + S{}^{(3)}\nabla^{b}\omega_{ab} - S{}^{(3)}\nabla^{b}\sigma_{ab} &= \kappa S q_{a} \qquad (6.106) \\ {}^{(3)}\nabla^{a}\omega_{a} &= 0 \qquad (6.107) \\ -h^{c}_{a}h^{d}_{b}(\omega^{e;f}_{(c} + \sigma^{e;f}_{(c})\eta_{d)gef}u^{g} &= H_{ab} \quad , \qquad (6.108) \\ & \qquad (6.109) \end{aligned}$$

while the linearised constraint equations on the electric and magnetic parts of the Weyl tensor transpire as:

$$\begin{aligned} S{}^{(3)}\nabla^{b}E_{ab} &= \frac{1}{3}\kappa\rho\mathcal{D}_{a} - \frac{1}{2}\kappa S\Pi_{a} - \kappa S\mathcal{H}q_{a} \qquad (6.110) \\ {}^{(3)}\nabla^{b}H_{ab} &= \frac{1}{2}\kappa\eta_{abcd}(hu^{b}\omega^{cd} + q^{[c;d]}) \quad . \qquad (6.111) \end{aligned}$$

6.6.3 Gauge-Invariant Evolution Equations

The aim now is to derive evolution equations for the operative perturbation variables; these equations will then naturally be derived from the linearised equations above. It should be noted, however, that in deriving requisite linearised equations from linearised equations, further linearisation is required in the derivation. This is necessary, as equations derived from substitution and differentiation of linearised equations will not necessarily be linear.

Propagation Equations for the Vectorial Quantities

The linearised propagation equations for (an imperfect fluid) the perturbation variables discussed above are naturally obtained from the linearised field equations of the previous section; they are:

$$\dot{\mathcal{D}}_a - 3\mathcal{H}W\mathcal{D}_a + (1+W)\mathcal{Z}_a = 3S(1+W)\mathcal{H}(F_a + \Pi_a) - S(1+W)h_a^b h_b^c \Psi_c^{;b} \quad (6.112)$$

$$\begin{aligned} &\dot{\mathcal{Z}}_a + 2\mathcal{H}\mathcal{Z}_a + \frac{1}{2}\kappa\rho\mathcal{D}_a + \frac{c_s^2}{1+W}\left(\frac{K}{S^2} + {}^{(3)}\nabla^2\right)\mathcal{D}_a + \frac{W}{1+W}\left(\frac{K}{S^2} + {}^{(3)}\nabla^2\right)\mathcal{E}_a \\ &= -S^{(3)}\nabla_a{}^{(3)}\nabla^b(F_b + \Pi_b) + S\left(\frac{3h}{2} - \frac{3K}{S^2}\right)(F_a + \Pi_a) - 6S\mathcal{H}c_s^{2(3)}\nabla^b\omega_{ab} \quad , \end{aligned} \quad (6.113)$$

such that:

$$W = \frac{p}{\rho} \quad , \quad (6.114)$$

and $K = 0, \pm 1$ as before.

To solve for $\mathcal{D}_a$, one can then substitute the second of these equations into the time derivative of the first, thereby obtaining a second order equation in $\mathcal{D}_a$.

It should be noted here that, should an investigation of the coupling between the above operative gauge-invariant variables and other dynamical quantities such as the shear be sought, a suitable substitution into the above two equations using the remaining linearised evolution equations of the previous section can readily be made.

For a perfect fluid the above simplify and decompose into:

$$\begin{aligned} h_a^b\dot{\mathcal{D}}_c &= \Theta W\mathcal{D}_a - \Theta\mathcal{Z}_a\left(1 + \frac{p}{\rho}\right) - \mathcal{D}_b\left(\sigma_b^a + \omega_b^a\right) && (6.115) \\ h_a^b\dot{\mathcal{Z}}_b &= \frac{\dot{\Theta}}{\Theta}\mathcal{Z}_a - \mathcal{Z}_b\left(\sigma_a^b + \omega_a^b\right) + \frac{\dot{\Theta}p}{\dot{\rho}}\mathcal{Y}_a + Sh_a^b\frac{\dot{\Theta}_{,b}}{\Theta} && (6.116) \\ h_a^b\dot{\mathcal{Y}}_b &= \frac{\dot{p}}{p}\left[\frac{\rho + p(1-S)}{\rho + p}\right]\mathcal{Y}_a - \mathcal{Y}_b\left(\sigma_a^b + \omega_a^b\right) + Sh_a^b\frac{\dot{p}_{,b}}{p} \quad . && (6.117) \end{aligned}$$

6.6.4 Propagation Equations for the Scalar Quantities

The gauge-invariant scalar quantities associated with the above vectorial quantities have the following propagation equations:

$$\begin{aligned} \dot{\mathcal{D}} - 3\mathcal{H}W\mathcal{D} &= (1+W)\left(3S\mathcal{H}(F+\Pi) - \mathcal{Z} - S^{(3)}\nabla^2\Psi\right) && (6.118) \\ \dot{\mathcal{Z}} + 2\mathcal{H}\mathcal{Z} &= \frac{3}{2}h(F+\Pi) - \frac{1}{2}\kappa\rho\mathcal{D} - \frac{c_s^2}{1+W}\left({}^{(3)}\nabla^2 + \frac{3K}{S^2}\right)\mathcal{D} \\ &\quad -\frac{W}{1+W}\left({}^{(3)}\nabla^2 + \frac{3K}{S^2}\right)\mathcal{E} - S\left({}^{(3)}\nabla^2 + \frac{3K}{S^2}\right)(F+\Pi) \quad . && (6.119) \end{aligned}$$

Taking the time derivative of the first equation, and then substituting therein from the second, one obtains a second-order equation for $\mathcal{D}$:

$$
\begin{array}{rl}
& \ddot{\mathcal{D}} + (2 + 3c_s^2 - 6W)\mathcal{H}\dot{\mathcal{D}} - c_s^{2(3)}\nabla^2\mathcal{D} \\
& - \left[\left(\frac{1}{2} + 4W - \frac{3}{2}W^2 - 3c_s^2\right)\kappa\rho + (5W - 3c_s^2)\Lambda + (c_s^2 - W)\frac{12K}{S^2}\right]\mathcal{D} \\
= & W\left({}^{(3)}\nabla^2 + \frac{3K}{S^2}\right)\mathcal{E} + S(1+W)\left[-3W\rho + 3\Lambda + \left({}^{(3)}\nabla^2 - \frac{3K}{S^3}\right)\right](F + \Pi) \\
& + S(1+W)\left[(3\mathcal{H}(\dot{F} + \dot{\Pi}) - (2\mathcal{H}^{(3)}\nabla^2\Psi + {}^{(3)}\nabla^2\dot{\Psi})\right] \quad .
\end{array} \tag{6.120}
$$

6.6.5 Comparison with the Bardeen Formalism

As the above covariant variables are exact as well as fully covariant, it follows naturally that they can, if desired, be expanded in terms of gauge-dependent quantities; moreover, due to this *de facto* gauge-independence, to first order they should then appear as linear combinations of the Bardeen gauge-invariant variables.

In deriving thus the relationship between the covariant and Bardeen formalisms, one first needs to note the distinction between the arbitrary space-time slicing in the Bardeen formalism, defined through $o_{\alpha\beta}$, and that defined by the fluid four-velocity via h_{ab} in the covariant formalism. This is described by taking the covariant derivative of a scalar function $f = \overline{f} + \delta f$ defined in the almost Robertson-Walker space-time, and with non-constant background value $\overline{f}$:

$$
\begin{aligned}
{}^{(3)}\nabla_\alpha f &= (\delta f)_{,\alpha} + \overline{u}^0\partial_0\overline{f}\delta u_\alpha \\
&= [\delta f + \overline{f}'(v - B)]_{,\alpha} + \overline{f}'v_{(c)\alpha} \\
&= -k\left[\delta f(\tau) - \frac{\overline{f}'}{k}(v^{(s)} - B^{(s)})\right]\mathcal{Q}_\alpha + \overline{f}'v_{(c)}\mathcal{S}_\alpha \quad .
\end{aligned} \tag{6.121}
$$

Using the above, one can express the operative covariant variables $\mathcal{D}_a$, $\mathcal{Z}_a$ and $\mathcal{C}_a$ in terms of the Bardeen variables defined in the previous chapter:

$$
\mathcal{D}_\alpha = -kS\varepsilon_m(\tau)\mathcal{Q}_\alpha - 3S'(1+W)v_{(c)}(\tau)\mathcal{S}_\alpha \tag{6.122}
$$

$$
\begin{aligned}
\mathcal{Z}_\alpha = {} & \left[-3k\left({\Phi_H}' - \frac{S'}{S}\Phi_A\right) + \left((3K - k^2) - \frac{3}{2}\kappa h S^2\right)v_{(s)}^{(s)}\right]\mathcal{Q}_\alpha \\
& + \left[3K - \frac{3}{2}\kappa h S^2\right]v_{(c)}(\tau)\mathcal{S}_\alpha
\end{aligned} \tag{6.123}
$$

$$
\mathcal{C}_\alpha = S\left[-k\left(4(k^2 - 3K)\left(\Phi_H - \frac{S'}{kS}v_{(s)}^{(s)}\right)\right)\mathcal{Q}_\alpha - 12K\frac{S'}{S}v_{(c)}\mathcal{S}_\alpha\right] \quad , \tag{6.124}
$$

such that $h = \mu + p$ as before, and k refers to the k^{th} harmonic component, as explained in the Appendix B.

6.6.6 Extension to Multi-Component Matter

For the sake of simplicity and clarity, the foregoing analysis was developed for a single-type matter description. However, the theory can be extended in a straight-forward fashion to encompass multi-component matter descriptions. The added complication that this introduces, apart from an extended set of equations, is interactions between the various matter types.

The above notion is initiated by assuming the total energy-momentum tensor as comprising the sum of the separate energy-momentum tensors of the constituent matter forms:

$$T_{ab} = \sum_{(i)} T^{(i)}_{ab} \tag{6.125}$$

$$= \sum_{(i)} \left(\rho_{(i)} u^{(i)}_a u^{(i)}_b + p_{(i)} h^{(i)}_{ab} + q^{(i)}_a u^{(i)}_b + q^{(i)}_b u^{(i)}_a + \pi^{(i)}_{ab} \right) \quad , \tag{6.126}$$

where the index (i) labels the energy-momentum tensor of the i^{th} matter type in the total matter content; and where the general energy-momentum tensor expansion form has been assumed. In order to consider interactions between the various matter components, one defines the following quantity:

$$J^a_{(i)} \equiv T^{ab}_{(i);b} \quad , \tag{6.127}$$

which, by the conservation of the total energy-momentum tensor, must therefore yield:

$$\sum_{(i)} J^a_{(i)} = 0 \quad . \tag{6.128}$$

For later considerations, it will be convenient to decompose $J^a_{(i)}$ into components perpendicular and parallel to the total matter fluid four-velocity u_a:

$$J^{(i)}_a = \epsilon_{(i)} u_a + f^{(i)}_a \tag{6.129}$$

$$u_a f^a_{(i)} = 0 \tag{6.130}$$

$$h^a_b u_a \epsilon_{(i)} = 0 \quad . \tag{6.131}$$

The following relative velocity will also be used subsequently:

$$V^a_{(i)} \equiv u^a_{(i)} - u^a \quad . \tag{6.132}$$

Using the above with the formalism of the foregoing sections, one has the following generalised form of the energy and momentum conservation equations:

$$\dot{\rho}_{(i)} + 3h_{(i)}\mathcal{H} + h_{(i)}{}^{(3)}\nabla_a \Psi^a_{(i)} = \epsilon_{(i)} \tag{6.133}$$

$$h_{(i)} a_a + Y^{(i)}_a + h_{(i)}(F^{(i)}_a + \Pi^{(i)}_a) + (1 + c^2_{s(i)})\epsilon_{(i)} \Psi^{(i)}_a = f^{(i)}_a \quad , \tag{6.134}$$

such that, analogous to the quantities defined earlier in the chapter, one has:

$$
\begin{aligned}
h_{(i)} &\equiv p_{(i)} + \rho_{(i)} && (6.135)\\
\Psi_a^{(i)} &\equiv \frac{q_a^{(i)}}{h_{(i)}} + V_a^{(i)} && (6.136)\\
\Pi_a^{(i)} &\equiv \frac{1}{h_{(i)}} {}^{(3)}\nabla^b \pi_{ab} && (6.137)\\
F_a^{(i)} &\equiv \dot{\Psi}_a^{(i)} - (3c_{s(i)}^2 - 1)\mathcal{H}\Psi_a^{(i)} \quad . && (6.138)
\end{aligned}
$$

Hence one has a set of covariant perturbation inhomogeneity variables defined, as the previous sections, for each matter component:

$$
\begin{aligned}
\mathcal{D}_a^{(i)} &= S\frac{{}^{(3)}\nabla_a \rho_{(i)}}{\rho_{(i)}} && (6.139)\\
\mathcal{Z}_a^{(i)} &= S^{(3)}\nabla_a \Theta_{(i)} && (6.140)\\
Y_a &= {}^{(3)}\nabla_a p_{(i)} && (6.141)\\
\mathcal{E}_a^{(i)} &= \frac{S}{p_{(i)}}\left(\frac{\partial p_{(i)}}{\partial s_{(i)}}\right)_{\rho_{(i)}} {}^{(3)}\nabla_a s_{(i)} \quad . && (6.142)
\end{aligned}
$$

Associated with these variables are scalar quantities defined as the projected (matter frame) divergence of each variable. For these vectorial and associated scalar quantities one can derive evolution equations as in the previous sections, the details of which are not necessary here (see [15] and [16] for an extensive treatment).

A crucial consequence of, if not indeed a motivation for the multi-fluid approach, is the explicit treatment of the *interactions* between the various matter components, a scenario which is of self-evident physical significance. One can analyse such interactions by deriving a set of evolution equations in terms of *relative* variables which define the *difference* between the perturbation variables of an arbitrary pair of matter components, and which have the advantage in permitting one to distinguish clearly between adiabatic and isothermal perturbations; for example, one can define the following relative variables:

$$
\begin{aligned}
S_a^{(ij)} &\equiv \frac{\rho_{(i)}}{h_{(i)}}\mathcal{D}_a^{(i)} - \frac{\rho_{(j)}}{h_{(j)}}\mathcal{D}_a^{(j)} && (6.143)\\
\mathcal{E}_a^{(ij)} &\equiv \frac{p_{(i)}}{h_{(i)}}\mathcal{E}_a^{(i)} - \frac{p_{(j)}}{h_{(j)}}\mathcal{E}_a^{(j)} && (6.144)\\
\epsilon_{(ij)} &\equiv \frac{\epsilon_{(i)}}{h_{(i)}} - \frac{\epsilon_{(j)}}{h_{(j)}} && (6.145)\\
\Psi_a^{(ij)} &\equiv \Psi_a^{(i)} - \Psi_b^{(j)} && (6.146)\\
f_a^{(ij)} &\equiv \frac{f^{(i)}}{h_{(i)}} - \frac{f^{(j)}}{h_{(j)}} && (6.147)\\
\Pi_a^{(ij)} &\equiv \Pi_a^{(i)} - \Pi_a^{(j)} && (6.148)\\
V_a^{(ij)} &\equiv V_a^{(i)} - V_b^{(j)} \quad , && (6.149)
\end{aligned}
$$

and, as before, evolution equations for these can be (tediously) derived (once again, for details see [15] and [16]). This formalism is of particular utility in the study of two-component matter descriptions, permitting a clear, unambiguous exposition on the interactions between any two given matter types. Numerous such studies on two-component matter have been performed by Ellis, Bruni, Dunsby *et al.* over the past decade. This inherent hydrodynamical approach has also been instrumental in recent applications [25] of kinetic theory ([26], [27]) in Cosmology.

6.7 The Perfect Fluid Case

As noted previously, the perfect fluid assumption forms the simplest meaningful matter assumption, and consequently, by substantially simplifying the evolution equations, clarifies the functioning of the covariant approach, highlighting its strengths and weaknesses. This consequently motivates a closer inspection of the perfect fluid case.

The relevant evolution equations for the primary perturbation variables $\mathcal{D}_a$, $\mathcal{Z}_a$, $\mathcal{Y}_a$ are then equations 6.115, 6.116 and 6.117, as before. One will be able to see a parallel between the Bardeen scalar variable Φ_H and the analogous vector gauge-invariant variable previously formulated as Φ_a. Using the linearisation procedure previously outlined, these evolution equations linearise to:

$$h^b_a \dot{\mathcal{D}}_b = W\Theta\mathcal{D}_a - (1+W)\mathcal{Z}_a \qquad (6.150)$$

$$h^b_a \dot{\mathcal{Z}}_b = -\frac{2}{3}\Theta\mathcal{Z}_a - \frac{1}{2}\kappa\rho\mathcal{D}_a + S(\dot{u}_a\mathcal{R} + D_a) \qquad (6.151)$$

$$h^b_a \dot{\mathcal{Y}}_b = \frac{\dot{p}}{\rho+p}\mathcal{Y}_a - S\kappa h^b_a \dot{p}_{,b} \quad , \qquad (6.152)$$

and for the Bardeen-type variable Φ_a:

$$h^b_a\dot{\Phi}_b + \frac{1}{3}\Theta\Phi_a + \frac{3}{2}(W+1)\frac{\kappa\rho}{\Theta}\Phi_a = \frac{3}{4}(1+W)\frac{\kappa\rho}{\Theta}S^3\mathcal{K}_a \quad , \qquad (6.153)$$

where, to linear order:

$$D \equiv \dot{u}^c_{;c} \qquad (6.154)$$

$$D_a \equiv h^b_a D_{,b} \qquad (6.155)$$

$$\mathcal{R} \equiv -\frac{1}{3}\Theta^2 + D + \kappa\rho + \Lambda \qquad (6.156)$$

$$\mathcal{K} \equiv 2(-\frac{1}{3}\Theta^2 + \kappa\rho + \Lambda) \qquad (6.157)$$

$$\mathcal{K}_a \equiv h^b_a \nabla_b \mathcal{K} \quad , \qquad (6.158)$$

such that when there is no vorticity, $\mathcal{K}$ equals the spatial Ricci scalar ${}^{(3)}R$, orthogonal to the fluid flow. One also has the following propagation equation for $\mathcal{K}$:

$$\dot{\mathcal{K}} = -\frac{2}{3}\Theta(\mathcal{K} + 2D) \quad , \qquad (6.159)$$

where the term ΘD can be linearised away, yielding upon integration:

$$\mathcal{K} = \frac{6K}{S^2} \quad , \tag{6.160}$$

such that $K = 0, \pm 1$ as before, corresponding to a flat, open or closed space-time. A useful equation related to the above, and here to linear order, which will be used later in decoupling the vectorial evolution equations is the following acceleration equation:

$$S(\frac{\mathcal{K}}{2}\dot{u}_a + D_a) = -\frac{c_s^2}{1+W}\left(\frac{K}{S^2}\mathcal{D}_a + {}^{(3)}\nabla^2\mathcal{D}_a\right) \tag{6.161}$$

Note that for the Bardeen analogue Φ_a evolution equation, in addition to it being vectorial, no harmonic decomposition has been necessary. For the relevant propagation equations, one has:

$$-\frac{2}{3}h_a^b\Theta_{,b} = h_a^b(\omega_b^c + \sigma_b^c)_{;c} \tag{6.162}$$

$$E^{ab}_{\ ;b} = \frac{1}{3}h^{ab}\kappa\rho_{,b} \quad . \tag{6.163}$$

One can then proceed to apply these equations to specific matter descriptions, for example pressure-free matter and the false vacuum. For the pressure-free case the relevant equations simplify to:

$$\mathcal{D}_a{}^{\cdot} = -\mathcal{Z}_a \tag{6.164}$$

$$\mathcal{Z}_a{}^{\cdot} = -\frac{2}{3}\mathcal{Z}_a - \frac{1}{2}\kappa\rho\mathcal{D}_a \quad . \tag{6.165}$$

Which can thus simply be separated to obtain:

$$0 = \mathcal{D}_a{}^{\cdot\cdot} + \frac{2}{3}\Theta\mathcal{D}_a{}^{\cdot} - \frac{1}{2}\kappa\rho\mathcal{D}_a = 0 \tag{6.166}$$

$$0 = \mathcal{Z}_a{}^{\cdot\cdot} + \frac{5}{3}\Theta\mathcal{Z}_a{}^{\cdot} + \frac{2}{3}\mathcal{Z}_a(D + \Lambda + \frac{2}{3}\Theta^2 - \frac{5}{4}\kappa\rho) \quad . \tag{6.167}$$

An interesting point to note here is that if $\mathcal{D}_a$ and $\mathcal{Z}_a$ are parallel at any point on some world line γ, then they are parallel everywhere along that world line; similarly, if either vanishes at any point on γ, they will be parallel at every point where they do not vanish. For these scenarios the vector equations reduce to scalar equations; for example:

$$\mathcal{D}^{\cdot\cdot} + \frac{2}{3}\Theta\mathcal{D}^{\cdot} - \frac{1}{2}\kappa\rho\mathcal{D} = 0 \quad , \tag{6.168}$$

such that:

$$\mathcal{D}_a \equiv \mathcal{D}e_a \tag{6.169}$$

$$e^a e_a \equiv 1 \tag{6.170}$$

$$e_a u^a = 0 \tag{6.171}$$

$$\mathcal{D} \equiv (\mathcal{D}^a\mathcal{D}_a)^{\frac{1}{2}} \quad . \tag{6.172}$$

The (scalar) solutions to these will indicate the extreme behaviour of vector solutions, as the magnitude of the vector solutions cannot grow faster than those of the scalar solutions. Hence one may use the scalar solutions to investigate how fast density inhomogeneities can grow. The solution also contains an inherent gravitational instability in terms of a growing mode - this will be seen shortly. Equation 6.168 was also obtained by Lifshitz [1], although in a different form, and can also be obtained from Newtonian theory, as derived by Bonnor [2], as will be seen later.

One can now consider specific universe models, namely the Einstein static and Einstein-de Sitter universes. The subsequent equations are then integrated to obtain standard well-defined and physical growing and decaying modes for a density inhomogeneity. To illustrate this, consider the Einstein-de Sitter universe ($K = \Lambda = 0$); this presents a solution ([13]) to equation 6.168 of the form:

$$\mathcal{D}^2 = \alpha(t - t_0)^{\frac{4}{3}} + \beta(t - t_0)^{\frac{-1}{3}} + \gamma(t - t_0)^{-2} \tag{6.173}$$

such that α, β, γ are constants. A gravitational instability is thus clearly exhibited by the *growing* mode nature of the first term of this solution, a crucial condition for structure formation in the universe. This will be compared with the Newtonian Jeans instability later (section 6.8.2).

It is worth noting that if the usual variables are used in the above, the results are less clear due to the gauge freedom inherent in the choice of initial surface from which to measure proper time. This approach also prevents the inconvenience in eliminating the decaying mode as illustrated by Olson [9].

Furthermore, for the above vanishing pressure case, the evolution of the variable $\mathcal{D}_a$ will be unaffected by the wavelength of the density fluctuation because the evolution along each world line is independent. In addition to this, and for the same reason, the evolutions are unaffected by particle horizons, a central issue in Bardeen's analysis.

6.8 Comparison with Newtonian Results

One can now make comparisons of the above with standard analogous results in Newtonian physics; the General Relativistic results having to meet the criterion of simplifying to the classical results in the Newtonian limit. This comparison assists in the interpretation of the results obtained from the relativistic equations by analogy with the well-established and understood Newtonian results.

6.8.1 Covariant Newtonian Gravity

The following, due to Jackson [19], expounds on earlier work by Jeans, Bonnor [2] *et al.* on the Newtonian regime of gravitational instability, but in a fully relativistic and covariant manner; deriving in the process general, exact second order acoustic evolution equations for the co-moving fractional spatial energy gradient. The results are compared with the standard Newtonian results, as well as the special cases of the Einstein static universe, and expanding homogeneous isotropic cosmological models.

One thus commences by comparing the Newtonian equations of motion with those of their Relativistic counterparts. Looking at the Newtonian framework, the analogues of the continuity and momentum conservation equations for a perfect fluid are:

$$\begin{aligned} \frac{\dot{\rho}}{\rho} + \Theta &= 0 & (6.174) \\ \dot{u}_\alpha + \frac{1}{\rho}\nabla_\alpha p + \nabla_\alpha \Phi &= 0 \quad , & (6.175) \end{aligned}$$

such that Φ is the gravitational potential determined by Poisson's equation:

$$\nabla^2 \Phi = 4\pi G\rho + \Lambda \quad . \tag{6.176}$$

Where the latter equation is referred to as *Euler's equation of motion*. Now, analogous to the formulation within General Relativity, in Newtonian fluid mechanics one can decompose the gradient of the fluid flow vector u_α as follows:

$$u_{\alpha;\beta} = \frac{1}{3}\Theta g_{\alpha\beta} + \sigma_{\alpha\beta} + \omega_{\alpha\beta} \quad . \tag{6.177}$$

This is naturally defined within Galilean spatial co-ordinates. The dynamical quantities contained in the above are defined exactly analogously to their relativistic counterparts in terms of the fluid flow vector. Taking the divergence of Euler's equation, and utilising Poisson's equation together with 6.177, one obtains a direct analogue of the Relativistic Raychaudhuri equation:

$$\dot{\Theta} + \frac{1}{3}\Theta^2 + 2\sigma^2 - 2\omega^2 + 4\pi G\rho - \Lambda + \nabla^\alpha \left(\frac{1}{\rho}\nabla_\alpha p\right) = 0 \quad . \tag{6.178}$$

This equation, together with the continuity equation and a suitable equation of state $p = p(\rho)$, form a closed system in the variables ρ, Θ and p. One now wishes to derive a second order equation for the density evolution. This is done by taking the total time derivative of the continuity equation and using the above equations to obtain:

$$\frac{\ddot{\rho}}{\rho} - \frac{4}{3}\Theta^2 - 2\sigma^2 + 2\omega^2 - 4\pi G\rho + \Lambda - \nabla^\alpha \left(\frac{1}{\rho}\nabla_\alpha p\right) = 0 \quad . \tag{6.179}$$

Analogously to the Bruni-Ellis covariant variable $\mathcal{D}_a$, one can define here:

$$\mathcal{D}_\alpha = \frac{1}{\rho + p}\nabla_\alpha \rho = \nabla_\alpha \int \frac{d\rho}{\rho + p} \quad . \tag{6.180}$$

This definition will be motivated later. Hence one can derive the Newtonian evolution equation thereof as:

$$\begin{aligned} 0 \;=\; & \ddot{\mathcal{D}}_\alpha + \tfrac{2}{3}\Theta\dot{\mathcal{D}}_\alpha + 2\left(\omega^\beta_\alpha + \sigma^\beta_\alpha\right)\left(\dot{\mathcal{D}}_\beta + \tfrac{1}{3}\Theta\mathcal{D}_\beta\right) - S\nabla_\alpha\nabla^\beta\left(\tfrac{1}{\rho}\nabla_\beta p\right) \\ & -4\pi G\rho\mathcal{D}_\alpha + \left(\omega^\beta_\alpha + \sigma^\beta_\alpha\right)^{\cdot}\mathcal{D}_\beta + \left(\omega^\beta_\alpha + \sigma^\beta_\alpha\right)\left(\omega^\gamma_\beta + \sigma^\gamma_\beta\right)\mathcal{D}_\gamma - 2S\nabla_\alpha\left(\sigma^2 - \omega^2\right) \quad . \end{aligned} \tag{6.181}$$

One can compare this with the spatially projected linearised (as defined earlier) covariant equivalent:

$$
\begin{aligned}
&h_a^b\ddot{\mathcal{D}}_b + \left(\tfrac{2}{3} - \tfrac{dp}{d\rho}\right)\Theta h_a^b\dot{\mathcal{D}}_b - \left[\left(\tfrac{dp}{d\rho}\right)^{\cdot}\Theta + \tfrac{dp}{d\rho}\left(4\pi G(\rho - 3p) + 2\Lambda\right) + 4\pi G(\rho + p)\right]\mathcal{D}_a \\
&-\nabla^c\nabla_c\left(S\mathcal{P}_a\right) - 2\tfrac{dp}{d\rho}\Theta S\nabla^c\omega_{ac} - \tfrac{2}{3}\left[\tfrac{1}{3}\Theta^2 - (\kappa\rho + \Lambda)\right]S\mathcal{P}_a \;=\; 0 \quad ,
\end{aligned}
\tag{6.182}
$$

such that:

$$
\mathcal{P}_a \equiv h_a^b\frac{p_{,b}}{\rho + p} \tag{6.183}
$$

$$
K_{ab} \equiv h_a^c h_b^d u_{c;d} \quad . \tag{6.184}
$$

The above equation is somewhat simpler than the previously derived equation in the Ellis-Bruni approach, the simplification being due to the alternative definition of the quantity $\mathcal{D}_a$. This, along with the relevant equation of state, does not form a closed system due to the appearance of the vorticity gradient term. The vorticity is, however, governed by the exact evolution equation:

$$
h_a^c h_b^d\dot{\omega}_{cd} = -\omega_{ab}\left(\frac{2}{3} - \frac{dp}{d\rho}\right)\Theta + 2\sigma_{c[a}\omega_{b]}^c \quad . \tag{6.185}
$$

However, this means that the vorticity dies away rapidly as the universe expands, given a reasonable equation of state. Hence one can consider the extra vorticity gradient term in equation 6.182 as insignificant, except at early times in the universe.

6.8.2 Gravitational Stability and the Jeans Criterion

The Jeans instability in essence provides a condition for the occurrence of a growing mode in the solution to a matter perturbation evolution equation, and thus a necessary criterion for possible structure formation. By using the standard Navier-Stokes and continuity equations in Newtonian gravity, one can perturb the Newtonian matter density ρ_0 as follows:

$$
\rho \equiv \rho_0 + \delta\rho \quad , \tag{6.186}
$$

such that

$$
\delta\rho = \delta\rho(r, t) \quad . \tag{6.187}
$$

Then, using Poisson's equation together with the Navier-Stokes and continuity equations (see Bonnor [2]) one can derive the following linearised evolution equation:

$$
\frac{\partial^2 s}{\partial t^2} - 4\pi G\rho_0 s - \nabla^2\left(s\frac{dp}{s\rho}\right) = 0 \quad , \tag{6.188}
$$

where $s \equiv \frac{\delta\rho}{\rho_0}$ is the *condensation* and the analogue to the covariant relativistic variable $\mathcal{D}$. Equation 6.188 is thus the Newtonian analogue of equation 6.168. To derive then Jeans' criterion, one considers an ansatz solution to this equation in the form of a spherical wave:

$$s = \frac{h(t)}{r}\cos(\frac{2\pi r}{\lambda}) \quad , \tag{6.189}$$

such that r is the radial distance, and λ is the wavelength. Upon substitution into 6.188 this yields:

$$\frac{d^2h}{dt^2} = \left(4\pi G\rho_0 - \frac{4\pi^2}{\lambda^2}\frac{dp}{d\rho}\right)h \quad . \tag{6.190}$$

Hence, for a growing mode, one would require the solution to increase exponentially thus making the term in brackets positive, or:

$$\lambda > \left(\frac{\pi}{G\rho_0}\frac{dp}{d\rho}\right)^{\frac{1}{2}} \quad . \tag{6.191}$$

A matter content with linear dimensions greater than that in the above will therefore be unstable. This is the Jeans instability, and as explained provides a criterion for structure formation *in theory*; however, in practice it transpires that the condensation process is too slow for the formation of nebulae. Bonnor ([2]) has shown that this criterion holds for all Newtonian Cosmologies except the Newtonian steady-state universe, where small perturbations decay away. Hence one could extend this argument to conclude that Newtonian Gravitation provides little, if any, evidence for structure formation in the universe.

6.9 Alternative Perturbation Variables

In relation to the covariant perturbation theory, one can generalise the previous approaches to include a whole plethora of various covariantly defined perturbation variables. One can proceed by assuming a general functional form for an arbitrary perturbation variable, bearing in mind the constraints of the Stewart-Walker lemma, derive an evolution equation for this variable, and then impose the functional form as desired. In the following analysis, for the sake of simplicity, a perfect fluid matter form will be assumed. The ensuing formulation will then facilitate the covariant perturbative Lagrangian formulation developed in Part III.

6.9.1 Generalised Equations

The standard exact (i.e. 'non-linearised') fractional energy density, pressure and expansion parameter gradient evolution equations for a perfect fluid are, as before:

$$\dot{\mathcal{D}}_a = \Theta\frac{p}{\mu}\mathcal{D}_a - \mathcal{Z}_a\Theta\left(1+\frac{p}{\mu}\right) - \mathcal{D}_b\left(\sigma^a_b + \omega^a_b - u_a\dot{u}^b\right) \tag{6.192}$$

$$\dot{\mathcal{Y}}_a = -\frac{\dot{p}}{p}\left[\frac{\mu+2p}{\mu+p}\right]\mathcal{Y}_a - \mathcal{Y}_b\left(\sigma^b_a + \omega^b_a - u_a\dot{u}^b\right) + Sh^b_a\frac{\dot{p}_{,b}}{p} \tag{6.193}$$

$$\dot{\mathcal{Z}}_a = -\frac{\dot{\Theta}}{\Theta}\mathcal{Z}_a - \mathcal{Z}_b\left(\sigma^b_a + \omega^b_a - u_a\dot{u}^b\right) + \frac{\dot{\Theta}p}{\dot{\mu}}\mathcal{Y}_a + Sh^b_a\frac{\dot{\Theta}_{,b}}{\Theta} \quad , \tag{6.194}$$

such that:

$$\mathcal{D}_a = Sh^b_a \frac{\mu_{,b}}{\mu} \tag{6.195}$$

$$\mathcal{Z}_a = Sh^b_a \frac{\Theta_{,b}}{\Theta} \tag{6.196}$$

$$\mathcal{Y}_a = Sh^b_a \frac{p_{,b}}{p} \quad . \tag{6.197}$$

If one now wishes to establish a basis from which to derive alternative and/or additional perturbation variables incorporating the energy density, and which are functionally similar to those in the above equations, one can proceed by positing a general functional form for the desired variable; this must be a function which incorporates a spatial gradient of a scalar, as this would vanish identically in an FRLW background model, hence admitting the Stewart-Walker lemma.

6.9.2 Energy-Density Variables

For variables analogous to the above fractional energy-density gradient, one postulates the following functional form.

$$\mathcal{D}_a = Sh^b_a \frac{(f(\mu))_{,b}}{g(\mu)} \quad , \tag{6.198}$$

such that f, g are general functions of the energy density μ which are to be determined; preferably in such a way so as to obtain evolution equations which have as simple a form as possible, while still maintaining a physical realism. One thus continues by deriving an evolution equation for the above generalised form in accordance with the standard fractional energy density gradient. Proceeding thus, one obtains the following form of the exact evolution equation:

$$\dot{\mathcal{D}}_a + \mathcal{D}_a\Theta(\mu+p)\left(\frac{f''}{f'} - \frac{g'}{g} + \frac{1}{\mu+p}\right) + \mathcal{D}_b\left(\sigma^b_a + \omega^b_a - u_a\dot{u}^b\right) = -\frac{Sf'\Theta}{g}(\mu+p)h^b_a\Theta_{,b} \quad , \tag{6.199}$$

such that:

$$f' = \frac{df}{d\mu} \tag{6.200}$$

$$g' = \frac{dg}{d\mu} \tag{6.201}$$

$$\dot{f} = f'\dot{\mu} \quad . \tag{6.202}$$

When linearised, the above evolution equation simplifies to the following:

$$h^b_a\left(\mathcal{D}_b\right)^{\cdot} = -\mathcal{D}_a\left(\ln g\right)^{\cdot} + \frac{S}{g}h^b_a\left(\dot{u}_b\dot{f} + \dot{f}_{,b}\right) \quad . \tag{6.203}$$

An interesting point to note here is that the linearised equation is also completely projected onto the spatial part of the comoving frame. This point does not seem to be clarified in the Ellis-Bruni paper [13], where the projection tensor h_{ab} is still redundantly appended to the $\dot{\mathcal{D}}_a$ term. It is easily seen that this is not necessary, as the linearisation eliminates the entire term $\mathcal{D}_b\left(\sigma^a_b + \omega^a_b - u_a\dot{u}^b\right)$ which itself incorporates the only temporal component $u_a\dot{u}^b$ in the entire equation. Having established this general functional form, one is now in a position to try a number of ansätze for f and g.

Different Functional Forms for the Energy Density

Ansatz I

Bearing simplification in mind, it would be convenient if the last term of equation 6.203 vanished. To this end one sets this last term to zero to obtain the following constraint:

$$\dot{u}_a = -h^b_a \frac{\dot{f}_{,b}}{\dot{f}} \quad , \tag{6.204}$$

which is essentially an acceleration potential equation:

$$\dot{u}_a = -h^b_a (\ln r)_{,b} \quad , \tag{6.205}$$

where r is the acceleration potential, and the 'dot' superscript indicates differentiation with respect to the proper time t as before. Whence one can equate $\dot{f}$ with the acceleration potential function r and in turn integrate the latter with respect to the proper time t to obtain the desired form for f. For a perfect fluid with a γ-law equation of state, this yields:

$$f = \int_{t_0}^{t} \left(\frac{\mu}{\mu_0}\right)^{\frac{\gamma-1}{\gamma}} dt' \quad . \tag{6.206}$$

If one now continues by setting:

$$g = \left(\frac{\mu}{\mu_0}\right)^{\frac{\gamma-1}{\gamma}} \quad , \tag{6.207}$$

then the evolution equation simplifies to the following:

$$h^b_a \left(h^c_b t_{,c}\right)^{\cdot} = (\gamma - 1)\,\Theta h^c_a t_{,c} - h^c_b t_{,c}\left(\sigma^b_a + \omega^b_a\right) \quad , \tag{6.208}$$

that is, an evolution equation for the spatial gradient of the proper time.

Ansatz II

If one tries the ansatz:

$$f = \mu, g = \mu + p \quad , \tag{6.209}$$

then the evolution equation becomes:

$$h^c_a \dot{\mathcal{D}}_c = -\frac{\dot{p}}{\mu + p}\mathcal{D}_a - S h^b_a \Theta_{,b} - \mathcal{D}_c\left(\sigma^c_a + \omega^c_a\right) \quad . \tag{6.210}$$

This definition of $\mathcal{D}_a$ can be considered meaningful in the sense that one is comparing the energy density gradient with the *total* energy density $\mu + p$.

Ansatz III

Define:

$$f = g = \mu + p \quad . \tag{6.211}$$

One then has the following evolution equation:

$$h_a^b \dot{\mathcal{D}}_b = -\frac{\dot{p}}{\mu + p}\mathcal{D}_a - S h_a^b \Theta_{,b} - \mathcal{D}_c \left(\sigma_a^c + \omega_a^c\right) + \frac{S}{\mu + p} h_a^b \left(\dot{p}_{,b} + \dot{u}_b \left(\dot{\mu} + \dot{p}\right)\right) \quad . \tag{6.212}$$

Here one is merely considering the fractional gradient of the total energy density.

Ansatz IV

Define:

$$f = g = \mu + 3p \quad . \tag{6.213}$$

One then obtains the evolution equation as:

$$\begin{aligned} h_a^b \dot{\mathcal{D}}_b &= -\frac{\dot{\mu} + 3\dot{p}}{\mu + 3p}\mathcal{D}_a - \mathcal{D}_b \left(\sigma_a^b + \omega_a^b\right) - \frac{(\mu + p)}{\mu + 3p} S h_a^b \Theta_{,b} \\ &+ \frac{3\Theta}{\kappa\left(\mu + 3p\right)} S h_a^b p_{,b} + \frac{3S}{\mu + 3p} h_a^b \left(\dot{p} \dot{u}_b + \dot{p}_{,b}\right) \quad . \end{aligned} \tag{6.214}$$

Ansatz V

If one now considers as a specific case a γ-law perfect fluid equation of state:

$$p = (\gamma - 1)\mu \quad , \tag{6.215}$$

then equation 6.203 obtains the following form:

$$\begin{aligned} h_a^b \dot{\mathcal{D}}_b &= \mathcal{D}_a \Theta \left(\frac{g'}{g}\gamma\mu - 1\right) - \mathcal{D}_b(\omega_a^b + \sigma_a^b) \\ &- \frac{\gamma\mu}{g} S h_a^b (f'\Theta_{,b} + \Theta f'_{,b}) \quad , \end{aligned} \tag{6.216}$$

using the energy conservation equation. Whence one obtains the following special cases:

i) $f = g = \mu + p = \gamma\mu$:

$$h_a^b \dot{\mathcal{D}}_b = -\left[-\Theta\left(\gamma - 1\right)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + \gamma S h_a^b \Theta_{,b}\right] \quad . \tag{6.217}$$

ii) $f = \mu, g = \mu + p = \gamma\mu$

$$h_a^b \dot{\mathcal{D}}_b = -\left[-(\gamma - 1)\Theta\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + S h_a^b \Theta_{,b}\right] \quad . \tag{6.218}$$

iii) $f = g = \mu$

$$h_a^b \dot{\mathcal{D}}_b = -\left[-\Theta(\gamma-1)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + \gamma S h_a^b \Theta_{,b}\right] \quad . \tag{6.219}$$

iv) $f = g = \mu + 3p = \mu\,(3\gamma - 2)$

$$h_a^b \dot{\mathcal{D}}_b = -\left[-\Theta(\gamma-1)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + \gamma S h_a^b \Theta_{,b}\right] \quad . \tag{6.220}$$

Similarly one has the following additional ansatz forms:

v) $f = \mu, g = \mu + 3p = \mu\,(3\gamma - 2)$

$$h_a^b \dot{\mathcal{D}}_b = -\left[-\Theta(\gamma-1)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + \frac{\gamma}{3\gamma-2} S h_a^b \Theta_{,b}\right] \quad . \tag{6.221}$$

vi) $f = \mu + p = \gamma\mu, g = \mu + 3p = \mu\,(3\gamma - 2)$

$$h_a^b \dot{\mathcal{D}}_b = -\left[-\Theta(\gamma-1)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + \frac{\gamma^2}{3\gamma-2} S h_a^b \Theta_{,b}\right] \quad . \tag{6.222}$$

It is evident that the above equations obtain their simplest form when $\gamma = 0$ or $\gamma = 1$. Note that in the above analysis only the case $\gamma = constant$ was considered. If one now releases this restriction and postulates $\gamma = \gamma(t)$, then equation 6.203 transforms into the following:

$$\begin{aligned} h_a^b \dot{\mathcal{D}}_b &= -\frac{1}{g}\mathcal{D}_a\left(g\Theta + g_\gamma\dot{\gamma} - g_\mu\mu\Theta\gamma\right) - \mathcal{D}_b(\omega_a^b + \sigma_a^b) \\ &\quad -\frac{S}{g}h_a^b\left[\dot{\gamma}\left(h_b^c\mu_{,c}f_\gamma\frac{\gamma-1}{\gamma} - f_{\gamma,b}\right) + \gamma\mu(\Theta f_{\mu,b} + f_\mu\Theta_{,b})\right] \quad , \end{aligned} \tag{6.223}$$

whence one has for the above special cases:

i) $f = g = \mu + p = \gamma\mu$:

$$h_a^b \dot{\mathcal{D}}_b = -\left[\left(\Theta\,(2\gamma-1) - 2\frac{\dot{\gamma}}{\gamma}\right)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + \gamma\Theta\mathcal{Z}_a + \frac{\dot{\gamma}(\gamma-1)}{\gamma^2}p\kappa\mathcal{Y}_a\right] \quad . \tag{6.224}$$

ii) $f = \mu, g = \mu + p = \gamma\mu$

$$h_a^b \dot{\mathcal{D}}_b = -\left[\left(\gamma\Theta - \frac{\dot{\gamma}}{\gamma}\right)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b + \omega_a^b\right) + \Theta\mathcal{Z}_a\right] \quad . \tag{6.225}$$

iii) $f = g = \mu$

This equation remains unchanged.

iv) $f = g = \mu + 3p = \mu\,(3\gamma - 2)$

$$h_a^b\dot{\mathcal{D}}_b = -\left[\left(\Theta\gamma\frac{9\gamma^2-9\gamma+1}{3\gamma-2} - 2\frac{3\dot{\gamma}}{3\gamma-2}\right)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b+\omega_a^b\right) + \gamma\left(3\gamma-2\right)\Theta\mathcal{Z}_a + \frac{3\dot{\gamma}(\gamma-1)}{\gamma(3\gamma-2)}p\kappa\mathcal{Y}_a\right] \quad . \tag{6.226}$$

v) $f=\mu, g=\mu+3p=\mu\left(3\gamma-2\right)$

$$h_a^b\dot{\mathcal{D}}_b = -\left[\Theta\frac{3\gamma^2-4\gamma+2}{3\gamma-2}\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b+\omega_a^b\right) + \frac{\gamma}{3\gamma-2}\Theta\mathcal{Z}_a + \frac{3\dot{\gamma}(\gamma-1)}{\gamma(3\gamma-2)}p\kappa\mathcal{Y}_a\right] \quad . \tag{6.227}$$

vi) $f=\mu+p, g=\mu+3p=\mu\left(3\gamma-2\right)$

$$h_a^b\dot{\mathcal{D}}_b = -\left[\left(\Theta\frac{4\gamma^2-5\gamma-2}{3\gamma-2} - \frac{2\dot{\gamma}(3\gamma-1)}{\gamma(3\gamma-2)}\right)\mathcal{D}_a + \mathcal{D}_b\left(\sigma_a^b+\omega_a^b\right) + \frac{\gamma^2}{3\gamma-2}\Theta\mathcal{Z}_a + \frac{\dot{\gamma}(\gamma-1)}{\gamma(3\gamma-2)}p\kappa\mathcal{Y}_a\right] \quad . \tag{6.228}$$

6.9.3 Functional Forms Involving the Pressure p

Proceeding analogously as for the energy density variable, and assuming an equation of state $p=p(\mu)$, one postulates here the following functional form for a generalised pressure variable:

$$\mathcal{Y}_a := h_a^b S\frac{f(p)_{,b}}{g(p)} \quad . \tag{6.229}$$

Whence one has an arbitrariness, as before, in the functions f and g which are now functions of the pressure p. The associated exact evolution equation for the above variable is the following:

$$\dot{\mathcal{Y}}_a + \mathcal{Y}_a c_s^2\Theta\left((\mu+p)\tfrac{f''}{f'} - 1 - (\mu+p)\tfrac{g'}{g}\right) + \mathcal{Y}_b\left(\sigma_a^b+\omega_a^b-u_a\dot{u}^b\right) = -\tfrac{Sf'}{g}h_a^b\left(\Theta(\mu+p)c_s^2\right)_{,b} \quad , \tag{6.230}$$

such that

$$f' = \frac{df}{dp} \tag{6.231}$$

$$g' = \frac{dg}{dp} \tag{6.232}$$

$$c_s^2 = \frac{dp}{d\mu} \quad , \tag{6.233}$$

the latter variable being the speed of sound within the fluid. The above evolution equation can then be linearised to yield:

$$\dot{\mathcal{Y}}_a + \mathcal{Y}_a c_s^2 \Theta \left((\mu+p)\frac{f''}{f'} - 1 - (\mu+p)\frac{g'}{g} \right) = -\frac{Sf'}{g} h_a^b \Big(\Theta(\mu+p)c_s^2 \Big)_{,b} \quad , \tag{6.234}$$

where, as with the energy-density evolution equation, the linearisation has eliminated the only time-like component, yielding a fully spatially-projected equation.

6.9.4 Functional Forms Involving the Expansion Parameter Θ

Here one postulates the following functional form:

$$\mathcal{Z}_a := h_a^b S \frac{f(\Theta)_{,b}}{g(\Theta)} \quad , \tag{6.235}$$

which admits the exact evolution equation:

$$\begin{aligned} &\dot{\mathcal{Z}}_a + \mathcal{Z}_a \left[\left(\tfrac{g'}{g} - \tfrac{f''}{f} \right) \left(\Lambda + D - \tfrac{1}{3}\Theta^2 - \tfrac{\kappa}{2}(\mu+3p) \right) + \tfrac{2}{3}\Theta \right] + \mathcal{Z}_b(\sigma_a^b + \omega_a^b - u_a \dot{u}^b) \\ &= \tfrac{Sf'}{g} h_a^b \left(\dot{u}_b(\Lambda + D - \tfrac{1}{3}\Theta^2 + \kappa\mu) + D_b - \tfrac{1}{2}\kappa\mu_{,b} \right) \end{aligned} \quad , \tag{6.236}$$

such that:

$$f' = \frac{df}{d\Theta} \tag{6.237}$$

$$g' = \frac{dg}{d\Theta} \tag{6.238}$$

$$D = \dot{u}^a_{;a} \tag{6.239}$$

$$D_b = h_a^b D_{,b} \tag{6.240}$$

$$\dot{\Theta} = \Lambda + D - \frac{1}{3}\Theta^2 - \frac{\kappa}{2}(\mu+3p) \quad , \tag{6.241}$$

the latter equation being the linearised Raychaudhuri equation. The above is linearised to yield:

$$\begin{aligned} &\dot{\mathcal{Z}}_a + \mathcal{Z}_a \left[\left(\tfrac{g'}{g} - \tfrac{f''}{f} \right) \left(\Lambda + D - \tfrac{1}{3}\Theta^2 - \tfrac{\kappa}{2}(\mu+3p) \right) + \tfrac{2}{3}\Theta \right] \\ &= \tfrac{Sf'}{g} h_a^b \left(\dot{u}_b(\Lambda + D - \tfrac{1}{3}\Theta^2 + \kappa\mu) + D_b - \tfrac{1}{2}\kappa\mu_{,b} \right) \end{aligned} \quad , \tag{6.242}$$

where, as before, the linearisation has eliminated the solitary time-like term, yielding a completely projected equation.

6.9.5 Decoupling the Evolution Equations

In general, from the preceding theory one notices that the desired information regarding the evolution of matter perturbations is contained primarily within the variables $\mathcal{D}_a$, $\mathcal{Y}_a$ and $\mathcal{Z}_a$. Looking then at the coupled linearised evolution equations for these, it readily follows that, in decoupling them, one would obtain third order ordinary linear differential equations. However, by using the simplified equation of state assumption $p = p(\mu)$, it transpires that *second* order equations can be extracted from the decoupling. This procedure will be highlightedΦΦΦhighlighted here through use of the following variables:

$$\mathcal{D}_a = Sh_a^b \frac{\mu_{,b}}{\mu} \tag{6.243}$$

$$\mathcal{Z}_a = Sh_a^b \frac{\Theta_{,b}}{\Theta} \quad . \tag{6.244}$$

The evolution equations for these variables, as evidenced before in equations 6.199 and 6.236, are coupled to each other, but *not* to any $\mathcal{Y}_a$ variable; this follows as:

$$p_{,b} = \frac{dp}{d\mu}\mu_{,b} = c_s^2 \mu_{,b} \quad , \tag{6.245}$$

resulting in the $\mathcal{Y}_a$ variable being absorbed into the $\mathcal{D}_a$ variable. Consequently, the evolution equation for $\mathcal{Y}_a$ need not be considered for their decoupling; it is essentially redundant. From equations 6.203 and 6.236 one obtains the following linearised evolution equations for $\mathcal{D}_a$ and $\mathcal{Z}_a$:

$$\dot{\mathcal{D}}_a = \Theta W \mathcal{D}_a - (1+W)\Theta \mathcal{Z}_a \tag{6.246}$$

$$\dot{\mathcal{Z}}_a = -\mathcal{Z}_a\left[\frac{\Theta}{3} + \frac{1}{\Theta}\left(\Lambda + D - \frac{1}{2}\kappa(\mu + 3p)\right)\right] + \frac{S}{\Theta}D_a - \frac{\kappa\mu}{2\Theta}\mathcal{D}_a + \frac{S\mathcal{R}}{\Theta}\dot{u}_a \tag{6.247}$$

Differentiating then equation 6.246 with respect to time, one obtains:

$$\begin{aligned} \ddot{\mathcal{D}}_a = {} & \left(\dot{\Theta}W + \Theta\dot{W}\right)\mathcal{D}_a + W\Theta\dot{\mathcal{D}}_a - \left(\dot{\Theta}(1+W) + \Theta\dot{W}\right)\mathcal{Z}_a \\ & - \Theta(1+W)\dot{\mathcal{Z}}_a \quad . \end{aligned} \tag{6.248}$$

Substituting in then for $\dot{\Theta}$ from 6.241, and $\dot{W}$ from:

$$\dot{W} = \Theta(W - c_s^2)(W+1) \quad , \tag{6.249}$$

and for $\mathcal{Z}_a$ from 6.246, $\dot{\mathcal{Z}}_a$ from 6.247, one obtains, after considerable simplification:

$$\begin{aligned} \ddot{\mathcal{D}}_a + \dot{\mathcal{D}}_a\Theta\left[\tfrac{2}{3} - 2W + c_s^2\right] + \mathcal{D}_a\left[\Theta^2\left(c_s^2 - \tfrac{4}{3}W\right) - W\Lambda - \tfrac{\kappa\mu}{2}(1 - 3W^2)\right] \\ = -S(1+W)\left(D_a + \tfrac{K}{2}\dot{u}_a\right) \end{aligned} \quad , \tag{6.250}$$

Where the term $\mathcal{D}_a W D$ has been linearised away, as it is a second-order term. One can then substitute in for the right hand side of this equation using 6.161, obtaining:

$$\begin{aligned} \ddot{\mathcal{D}}_a + \dot{\mathcal{D}}_a\Theta\left[\tfrac{2}{3} - 2W + c_s^2\right] + \mathcal{D}_a\left[\Theta^2\left(c_s^2 - \tfrac{4}{3}W\right) - W\Lambda - \tfrac{\kappa\mu}{2}(1 - 3W^2)\right] \\ = c_s^2\left(\tfrac{K}{S^2}\mathcal{D}_a + {}^{(3)}\nabla^2\mathcal{D}_a\right) \end{aligned} \tag{6.251}$$

Now, one can solve for this equation by assuming that the time and spatial dependence of the covariant variables are separable, thus permitting a harmonic decomposition as in equation B.2, as with the metric perturbation variables. This decomposition then results in a purely time dependent propagation equation for the k^{th} harmonic amplitude. Hence, using B.4, the right hand side of 6.251 becomes:

$$\frac{c_s^2}{S^2}(K - k^2)\mathcal{D}_a \quad , \tag{6.252}$$

where k indicates the k^{th} harmonic component. Now, as the vectorial components of the variable $\mathcal{D}_a$ enter via the vector harmonics $\mathcal{Q}_a$ through the decomposition:

$$\mathcal{D}_a = \sum_k \mathcal{D}_{(k)} \mathcal{Q}_a \tag{6.253}$$

upon substitution into equation 6.251 one obtains the following homogeneous equation in the k^{th} harmonic amplitude $\mathcal{D}_{(k)}$:

$$\begin{aligned} \ddot{\mathcal{D}}_{(k)} + \dot{\mathcal{D}}_{(k)}\Theta\left[\tfrac{2}{3} - 2W + c_s^2\right] + \mathcal{D}_{(k)}\Big[\Theta^2\left(c_s^2 - \tfrac{4}{3}W\right) - W\Lambda - \tfrac{\kappa\mu}{2}(1 - 3W^2) \\ + \tfrac{c_s^2}{S^2}(k^2 - K)\Big] \;=\; 0 \end{aligned} \tag{6.254}$$

A similar, although far more technically hazardous, method and calculation will then yield the corresponding evolution equation for $\mathcal{Z}_a$.

Part III

The Variational Principle in General Relativity

"There is a school of mathematical physicists which objects to the introduction of ideas which do not relate to things which can actually be observed and measured... I hold that if the introduction of a quantity promotes clearness of thought, then even if at the moment we have no means of determining it with precision, its introduction is not only legitimate but desirable. The immeasurable of today may be the measurable of tomorrow."

J.J. Thomson

In this section the theory of General Relativity will be formulated in terms of the Variational Principle, and this notion then extended to incorporate Cosmological Perturbation Theory.

The former will focus one the principal Lagrangians for General Relativity, the process of variation via different techniques, and a brief look at the tetrad version. The latter will incorporate the standard theory formulated for metric perturbations, and a new formulation for the covariant approach.

Chapter 7

The Variational Approach to Gravitation

"Physics is mathematical not because we know so much about the physical world, but because we know so little: it is only its mathematical properties that we can discover."
B. Russell

As already discussed in chapter 3, justification for the Variational principle *per se* does not seem to be well substantiated physically, its appeal being primarily mathematical and aesthetic. However, recent developments in theoretical physics, specifically quantum gravity, have revealed that a variational formulation is most suitable to canonical quantisation, especially in the case of Ashtekar gravity. However, the latter will not be considered here, as the principal aim is to study 'classical' General Relativity, and is only mentioned in order to provide the necessary context and motivation for the theory.

The following chapter is thus intended to provide a concise yet informative foray into the standard variational formulations of gravitation, covering several different Lagrangians as well incorporating the tetrad formalism.

The results will then be readily pliable to a variational formulation of Cosmological Perturbation Theory in Chapter 7.

7.1 The Standard Variational Formulation

As is well documented, the Hilbert variational formulation of General Relativity was almost contemporaneous with Einstein's original presentation; consequently, the variational approach to gravitation has been around for some eighty-five years and has been extensively analysed and developed. Most of this development has centred on mathematical formalism, and has spawned several Lagrangian forms based on Lagrangian functional dependence, as well as the parallel Hamiltonian formulation briefly mentioned in chapter 3.

The underlying concept for the Lagrangian form in general Relativity is, as expected, the same as that of Einstein's original formulation: the coupling between matter and space-time geometry. This notion leads naturally to the hypothesis of a master

Lagrangian comprising the sum of a matter term and a gravitational term which incorporates the geometry of space and time. Hence one would have:

$$\mathcal{L} = \mathcal{L}_G + \mathcal{L}_M \quad , \tag{7.1}$$

such that $\mathcal{L}_G$ is the gravitational Lagrangian, while $\mathcal{L}_M$ is the matter Lagrangian. The expression $\mathcal{L}_G$ which incorporates the geometry of space-time must then, in the absence of matter, yield the vacuum field equations, and hence constitute a simple starting point for the analysis: variation of the action:

$$\mathcal{A} = \int_{\mathcal{M}} \mathcal{L}_G d^4x \quad , \tag{7.2}$$

where naturally the Lagrangian is integrated over the manifold of space and time$\mathcal{M}$. By the formalism of chapter 3 one desires that the Lagrangian be a scalar density (chapter 1). To incorporate the geometry of space-time, one would then expect the quantity to be a function of the metric and its derivatives. The most suitable candidate transpires as being the Ricci scalar multiplied by the square root of the determinant of the metric, the latter ensuring the scalar density requirement. Hence one has:

$$\mathcal{L}_G = \sqrt{-g}R = \sqrt{-g}g^{ab}R_{ab} \quad . \tag{7.3}$$

One can now proceed with the standard variational formalism.

7.2 The Field Equations Via Different Lagrangian Forms

When applying the variational procedure, one needs to know with respect to what to vary the Lagrangian; hence the exact functional form of the Lagrangian needs to be known. Naturally, one does not have full carte blanche as to this functional form: the same Einstein Field Equations have to result. From the plethora of Lagrangian and Hamiltonian formulations, those selected for consideration here are the Einstein-Hilbert and Hilbert-Palatini formulations, due to their simplicity and utility. The calculations required are usually quite horrendous; consequently, only the major steps and results will be given here. For more details, one is referred to D'Inverno.

The Einstein-Hilbert action

This is the simplest possible Lagrangian form generating the field equations. The functional dependence of the lagrangian here is merely that of the metric and its first and second partial derivatives:

$$\mathcal{L} = \mathcal{L}(g_{ab}, g_{ab,c}, g_{ab,cd}) \quad . \tag{7.4}$$

Expanding then the Ricci scalar in terms of the metric, one has the Lagrangian as being:

$$\begin{aligned} \mathcal{L}_G &= \sqrt{-g}g^{cd}R_{cd} \\ &= \sqrt{-g}g^{cd}\left\{\left[\frac{1}{2}g^{ef}(g_{cf,d} + g_{df,c} - g_{cd,f})\right]_{,e} - \left[\frac{1}{2}g^{ef}(g_{cf,e} + g_{ef,c} - g_{ce,f})\right]_{,d}\right. \end{aligned}$$

$$+\left[\frac{1}{2}g^{fh}(g_{ch,d}+g_{dh,c}-g_{cd,h})\right]\left[\frac{1}{2}g^{ei}(g_{fi,e}+g_{ei,f}-g_{fe,i})\right]$$
$$-\left[\frac{1}{2}g^{fh}(g_{ch,e}+g_{eh,c}-g_{ce,h})\right]\left[\frac{1}{2}g^{ei}(g_{fi,d}+g_{di,f}-g_{fd,i})\right]\Bigg\} \quad . \tag{7.5}$$

The Euler-Lagrange equations then become:

$$\frac{\delta\mathcal{L}_G}{\delta g_{ab}}=\frac{\partial\mathcal{L}_G}{\partial g_{ab}}-\left(\frac{\partial\mathcal{L}_G}{\partial g_{ab,c}}\right)_{,c}+\left(\frac{\partial\mathcal{L}_G}{\partial g_{ab,cd}}\right)_{,cd}=0 \quad , \tag{7.6}$$

which (eventually) yield:

$$-\sqrt{-g}(R^{ab}-\frac{1}{2}Rg^{ab})=0 \quad , \tag{7.7}$$

that is, the vacuum field equations. By the constraint equation 3.117 derived in chapter 3 one has:

$$\left(\frac{\delta\mathcal{L}_G}{\delta g_{ab}}\right)_{;b}=0 \quad , \tag{7.8}$$

which by the above is merely the contracted Bianchi identities $G^{ab}_{\ ;b}=0$. Naturally in the above calculations the assumption of vanishing boundary terms is made. The above is the formal derivation which requires considerable calculation, although this is simplified by the identities derived in section 3.3.3. However, an equivalent yet simpler derivation follows by performing the variation *ab initio* using the results 3.10, 3.11 and 3.14. Following this, one thus obtains:

$$\begin{aligned}\delta\mathcal{A} &= \int_{\mathcal{M}}\delta\sqrt{-g}g^{ab}R_{ab}d^4x\\ &= \int_{\mathcal{M}}(\sqrt{-g}g^{ab}\delta\Gamma^c_{\ ab}-\sqrt{-g}g^{ac}\delta\Gamma^b_{\ ab})_{;c}d^4x+\int_{\mathcal{M}}\sqrt{-g}(\frac{1}{2}Rg^{cd}-R_{ab}g^{ac}g^{bd})\delta g_{cd}d^4x\\ &= \int_{\partial\mathcal{M}}(\sqrt{-g}g^{ab}\delta\Gamma^c_{\ ab}-\sqrt{-g}g^{ac}\delta\Gamma^b_{\ ab})d^3x+\int_{\mathcal{M}}(-\sqrt{-g}G^{ab})\delta g_{ab}d^4x \quad .\end{aligned} \tag{7.9}$$

The surface integral results from the observation that the quantity in the first integral is the covariant derivative of a scalar density which thus simplifies to a partial derivative, yielding the surface integral by Stokes' theorem. This then vanishes as one is assuming zero boundary variation. Hence one regains the vacuum field equations by Bliss' lemma (chapter 3) applied to the remaining integral.

The Hilbert-Palatini Action

This method entails regarding the metric and connections in the Lagrangian as being *independent* fields; that is, the connections are here not being regarded as functions of the metric. One thus has:

$$\mathcal{L}_G=\mathcal{L}_G(g^{ab},\Gamma^a_{\ bc},\Gamma^a_{\ bc,d}) \quad . \tag{7.10}$$

Hence in the variation of the action, one has to perform variations with respect to the metric and connection separately. The variation with respect to the metric thus simplifies to:

$$\delta\mathcal{A} = \int_{\mathcal{M}} R_{ab}\delta\sqrt{-g}g^{ab}d^4x \quad , \tag{7.11}$$

which by Bliss' lemma yields the vacuum equations in the form $R_{ab} = 0$. Performing then the variation with respect to the connection, integrating by parts and neglecting the surface term as per usual, one eventually obtains:

$$\delta\mathcal{A} = \int_{\mathcal{M}} \left(\delta^b_c(\sqrt{-g}g^{ad})_{;d} - (\sqrt{-g}g^{ab})_{;c}\right)\delta\Gamma^c{}_{ab}d^4x \quad . \tag{7.12}$$

Hence:

$$\left(\delta^b_c(\sqrt{-g}g^{ad})_{;d} - (\sqrt{-g}g^{ab})_{;c}\right)\delta\Gamma^c{}_{ab} = 0 \quad . \tag{7.13}$$

As the connection is symmetric, only the symmetric part of the expression in brackets vanishes; manipulating this symmetric part somewhat, one eventually obtains:

$$g_{ab;c} = 0 \quad , \tag{7.14}$$

which thus necessarily imposes that the connection be the metric connection. Naturally for both the Einstein-Hilbert and Hilbert-Palatini actions, the inclusion of a cosmological constant for generality doesn't change the procedure and techniques of the above results. If one does wish to include the cosmological term Λ, the Lagrangian takes on the following form:

$$\mathcal{L}_G = \sqrt{-g}(R + 2\Lambda) \quad . \tag{7.15}$$

For many practical considerations in Relativistic Cosmology, it is useful to convert the above variational formalism into the tetrad notation. In the following section this will be done, closely following work done by Peldan [20].

7.3 Tetrad Formulation

In the tetrad formalism the Lagrangian with cosmological constant can be written in the following convenient form:

$$\mathcal{L}_G = e\left(e^a_A e^b_B R^{AB}_{ab} + 2\Lambda\right) \quad . \tag{7.16}$$

Where the $R_{ab}{}^{AB}$ is here the curvature of the (unique) torsion-free spin-connection Σ^{AB}_a. The Einstein-Hilbert action then becomes equivalent to assuming the Lagrangian to be a function of the tetrad field e^a_A only. This is the fundamental difference between the tetrad and metric approaches. Hence, using the variations for the tetrad field 3.16, its determinant 3.17 and the Ricci tensor 3.19, one obtains:

$$\delta\mathcal{A} = \int_{\mathcal{M}} \left(-2e^a_C e^c_A e^b_B R_{cb}{}^{CB} + (2\Lambda + e^c_C e^b_B R_{cb}{}^{CB})e^a_A\right) e^A_a d^4x$$
$$+ \int_{\mathcal{M}} \left(e(g^{c[a}e^{b]}_B \mathcal{D}_{[b}\delta e^{B]}_{c]} + g^{c[a}g^{b]d}e_{Cb}\mathcal{D}_d\delta e^C_c)\right)_{,a} d^4x \quad . \tag{7.17}$$

As per usual, the second integral yields a boundary term by Stokes' theorem and which vanishes by assumption. The first term gives the vacuum field equations in tetrad form:

$$e^a_C e^c_A e^b_B R_{cb}{}^{CB} - (\Lambda + \frac{1}{2}e^c_C e^b_B R_{cb}{}^{CB})e^a_A = 0 \quad . \tag{7.18}$$

With the Hilbert-Palatini approach, one assumes the spin-connection and tetrad fields to be independent, analogous to the formulation of the previous section. Hence one has to perform variations separately with respect to the tetrad field and the spin-connection. Performing the variation with respect to these fields separately one obtains:

$$\delta\mathcal{A} = \int_{\mathcal{M}} \left(-2e^a_C e^c_A e^b_B R_{cb}{}^{CB} + (2\Lambda + e^c_C e^b_B R_{cb}{}^{CB})e^a_A\right) e^A_a d^4x$$
$$- \int_{\mathcal{M}} \left[\frac{1}{2}\mathcal{D}_a \left(\epsilon^{abcd}\epsilon_{ABCD}e^C_c e^D_d\right)\delta\Sigma^{AB}_{\ \ b} - (ee^a_A e^b_B \delta\Sigma^{AB}_{\ \ b})_{,a}\right] d^4x \quad . \tag{7.19}$$

Neglecting again the surface term one obtains the field equations 7.18 as well as the constraints:

$$\mathcal{D}_{[a}e^C_{c]} = 0 \quad . \tag{7.20}$$

This is, however, just the zero-torsion assumption, which thus yields the unique spin-connection associated with the tetrad field e^A_a.

7.3.1 The 3+1 ADM Action

Recall from section 4.5.2 the metric form in terms of the ADM 3 + 1 tetrad splitting:

$$ds^2 = (N^2 - N_\alpha N^\alpha)d\tau^2 - 2N_\alpha dx^\alpha d\tau - o_{\alpha\beta}dx^\alpha dx^\beta \quad . \tag{7.21}$$

In terms of this formalism the gravitational part of the Lagrangian becomes:

$$\begin{aligned}\mathcal{A} &= \int R\sqrt{-g}d^4x \\ &= \int \left[N\sqrt{o}(K^\alpha_\beta K^\beta_\alpha - K^2 + {}^{(3)}R) - 2(\sqrt{o}K)' + 2(\sqrt{o}KN^\alpha - \sqrt{o}o^{\alpha\beta}N_{,\beta})_{,\alpha}\right] d^4x \\ &= \int \left[N\sqrt{o}(K^\alpha_\beta K^\beta_\alpha - K^2) + \tfrac{1}{2}(\sqrt{o}o^{\alpha\beta}N)_{,\alpha}(\ln 0)_{,\beta} + N_{,\alpha}(\sqrt{o}o^{\alpha\beta})_{,\beta} - \tfrac{1}{2}N\sqrt{o}\,{}^{(3)}\Gamma^\gamma{}_{\alpha\beta}o^{\alpha\beta}{}_{,\gamma}\right] d^4x \\ &\quad - \int \left[2(\sqrt{o}K)' - 2(\sqrt{o}KN^\alpha - \sqrt{o}o^{\alpha\beta}N_{,\beta})_{,\alpha} + [No^{\alpha\beta}\sqrt{o}_{,\beta} + N(\sqrt{o}o^{\alpha\beta})_{,\beta}]_{,\alpha}\right] d^4x \quad ,\end{aligned} \tag{7.22}$$

such that o indicates the determinant of the spatial part of the metric tensor, $o_{\alpha\beta}$; while $K_{\alpha\beta}$ is the extrinsic curvature tensor, defined through equation 4.72, for the spatial hypersurface associated with $\tau = cst$. The covariant derivative is taken with respect to $o_{\alpha\beta}$, and a primed sign indicates differentiation with respect to proper time, as before.

Note that the last line of 7.22 is obtained by expanding the three-curvature scalar ${}^{(3)}R$ in terms of the associated connections, while the second integral in this expression is merely a total derivative term which will thus vanish when the Variational Principle is applied, as this constitutes a boundary term.

7.4 The Matter Lagrangian

In general contexts, the matter Lagrangian $\mathcal{L}_M$ is defined such that the functional derivative (chapter 2) thereof is the requisite energy-momentum tensor; i.e.

$$T_{ab} \equiv \frac{\delta \mathcal{L}_M}{\delta g^{ab}} \quad , \tag{7.23}$$

as, for example, is proposed by Weinberg ('Gravitation and Cosmology: Principles and applications of the general theory of relativity' Chapter 12). Note that this, in essence, presupposes the existence of a matter Lagrangian *from which an energy-momentum tensor is derived*; however, in practice, one presupposes the form of the energy-momentum tensor. Then, in order to find the matter Lagrangian, one has to solve the 'inverse problem' by integrating the Euler-Lagrange equations. As can well be imagined, this procedure is not at all in general trivial; hence, it makes sense to assume a simple matter description, for which the Lagrangian can easily be solved, if one intends looking specifically at perturbation theory in terms of a Lagrangian. This will be seen later.

For a general matter description in a given system one will have several kinds of matter forms present; in this case, the total matter Lagrangian will naturally be the algebraic sum of the separate Lagrangians for each constituent matter component. For most of the subsequent analysis, the matter will be assumed to be that of a perfect fluid; for the pressure-free case this has the simple Lagrangian:

$$\mathcal{L}_M = -\sqrt{-g}\rho \quad , \tag{7.24}$$

where ρ is the total energy density; and such that:

$$T_{ab} = -\rho v_a v_b \quad , \tag{7.25}$$

where v_a is the normalised four-velocity of the matter of density ρ located at x^a:

$$v_a = \frac{dx_a}{ds} \tag{7.26}$$

$$ds^2 = g^{ab} dx_a dx_b \quad , \tag{7.27}$$

from which one can define:

$$p^a = \rho\sqrt{-g} v^a \quad , \tag{7.28}$$

as being the 'energy flux'. Hence:

$$\begin{aligned} \mathcal{A}_M &= -\int \rho\sqrt{-g}d^4x \\ &= -\int (p^a p_a)^{\frac{1}{2}} d^4x \quad . \end{aligned} \tag{7.29}$$

7.5 Bibliography

- Courant and Hilbert (1933) *Methods in Mathematical Physics* Vol. 1, 2 (Interscience)
- D Lovelock and H Rund (1975) *Tensors, Differential Forms and Variational Principles* (Wiley)
- R D'Inverno (1992) *Introducing Einstein's Relativity* (Oxford)
- PAM Dirac *General Relativity* (Cambridge)
- SL Shapiro and SA Teukolsky (1983) *Black Holes, White Dwarfs and Neutron Stars* (Wiley)

Chapter 8

A Variational Approach to Perturbation Theory

"I am inclined to think that scientific discovery is impossible without faith in ideas which are of a purely speculative kind and sometimes quite hazy; a faith which is quite unwarranted from the scientific point of view."
K. Popper

8.1 Perturbation Theory Formulated via the Variational Principle

8.1.1 Motivation and Aims

As discussed earlier, a variational formulation of any particular branch of physics is usually motivated through the consequent access to the powerful mathematical machinery of the Variational Calculus, as evidenced in chapter 3. In addition to using these results, one also has an elegant and simple means of looking at composite matter systems, scenarios involving the inclusion of abstract scalar fields, Electromagnetism and so forth, in that one merely formulates a master Lagrangian as being the algebraic sum of the individual Lagrangians of all these particular cases. This was seen earlier in chapter 3 when Electromagnetism was incorporated into General Relativity via the addition of the Electromagnetic Lagrangian to the Relativistic Lagrangian. The Lagrangian formulation is also readily amenable to quantisation (see [17]) and which, as in Ashtekar gravity, provides a platform for Quantum Gravity - the unification of Quantum Mechanics and General Relativity - not to mention the whole plethora of Grand Unified Theories (GUT). In addition, once a relevant Lagrangian has been formulated, a Legendre transformation can be performed on it to yield the associated Hamiltonian, as derived in Chapter 2, which is fairly useful in the determination of conserved quantities via Noether's theorem.

For the case of Cosmological Perturbation Theory, the aim is therefor e to derive generalised evolution equations for the linearised perturbation equations through the variation of some prescribed Lagrangian.

8.1.2 Linearisation About a Background Model

As expounded in the preceding chapters, the notion of a perturbation theory requires the assumption of an idealised background space-time model which is 'perturbed' to obtain a realistic space-time - the 'lumpy' model. Hence the difference between these two models is in essence the perturbation. Consequently, a particular quantity in the background model is correspondingly perturbed to obtain its observable counterpart in the real universe.

This line of reasoning would imply that, if one were to start from a prescribed action for the real universe, one would have to expand it as a Taylor series about some background model. In such an expansion, the successive terms added to the background term would correspond to perturbations of various orders. One can express this expansion of the Lagrangian $\mathcal{L}$ as:

$$\mathcal{L} = \mathcal{L}_0 + \delta_1\mathcal{L} + \delta_2\mathcal{L}\ldots \quad , \tag{8.1}$$

where $\mathcal{L}_0$ is the background Lagrangian, and δ_n refers to the n^{th} order Taylor expansion term. In this section only first order perturbation theory will be considered; consequently one would need to expand the Lagrangian to second order, as second order terms give rise to first order terms in the variation. Hence, in order to perform the expansion, one needs to express the desired variables in terms of background quantities and some generalised perturbative functions. This decomposition into background and perturbed quantities then has the advantage of yielding two sets of equations: the background equations obtained through variation of the action with respect to the background quantities alone; and the perturbation equations obtained through variation with respect to the perturbation variables.

8.1.3 Metric Perturbation Theory

As a well-defined and complete set of perturbation equations for general perturbation theory is not self-evident, it would seem natural to try to derive such a set unambiguously by stipulating an action *ab initio*. Applying the least action principle to this action should then yield a set of equations containing the established perturbation equations. In the following, such an action will be developed for the classic metric perturbation theory, using the ADM and Bardeen [10] formalisms; the derivation closely follows that of Brandenberger *et al.* ([11],[17]).

The Gravitational Action

In configuring such an action for the metric perturbation case, it is convenient to utilise the ADM formalism, as this simplifies the ensuing calculations substantially. For the same physical reasons as given in the preceding chapters, only scalar perturbations will be considered in this formalism. One thus starts with the ADM Lagrangian 7.22, and substitutes in for the perturbed metric from 5.19. In the process, one expands the Lagrangian to second order in the perturbation variables only, as a second order Lagrangian will yield first order perturbation equations. Firstly, one notes the following identifications after comparing the two metrics:

$$
\begin{aligned}
N_\alpha &= S^2\psi_{;\alpha} && (8.2)\\
N &= S(1+\phi-\frac{1}{2}\phi^2+\frac{1}{2}\psi_{,\alpha}\psi^{,\alpha}) && (8.3)\\
o_{\alpha\beta} &= S^2(1-2\gamma)\delta_{\alpha\beta}+2S^2\kappa_{,\alpha\beta} \quad , && (8.4)
\end{aligned}
$$

where the notation of chapter 3 has been used.

Noting this, one then has the gravitational Lagrangian for the perturbed metric, *keeping terms only up to second order in the perturbation variables:*

$$
\begin{aligned}
\mathcal{A}_{2G} = & \int \Big\{ S^2 \Big[-6\gamma'^2 - 12\mathcal{H}(\phi+\gamma)\gamma' - 9\mathcal{H}^2(\phi+\gamma)^2 - 2\gamma_{,\alpha}(2\phi_{,\alpha}-\gamma_{,\alpha}) \\
& -4\dot{\mathcal{H}}(\phi+\gamma)(\psi-\kappa')_{,\alpha\alpha} - 4\gamma'(\psi-\kappa')_{,\alpha\alpha} - 4\mathcal{H}\gamma_{,\alpha}\psi_{,\alpha} + 6\mathcal{H}^2(\phi+\gamma)\kappa_{,\alpha\beta} \\
& +4\mathcal{H}\gamma'\kappa_{,\alpha\beta} - 4\mathcal{H}\kappa_{,\alpha\alpha}(\psi-\kappa')_{,\beta\beta} + 4\mathcal{H}\kappa_{,\alpha\alpha}\psi_{,\beta\beta} \\
& +3\mathcal{H}^2{\kappa^2}_{,\alpha\alpha} + 3\mathcal{H}^2\psi_{,\alpha}\psi_{,\alpha} \Big] + \Sigma_G \Big\} d^4x \quad ,
\end{aligned}
\tag{8.5}
$$

where summation over repeated lower indices is implicit; the '2' subscript implies expansion to second order in the perturbation variables; and such that the term Σ_G refers to the total derivative terms which vanish when the least action principle is applied to the action. Consequently, these need not be considered in detail here. Variation with respect to the metric perturbation variables above would then yield evolution equations for the desired perturbation variables, as will be seen later.

The Matter Lagrangian

Here, for the sake of simplicity, the matter Lagrangian for a perfect fluid only will be considered, and formulated in terms of the metric perturbation variables. One thus starts with the action for the 'real' perturbed space-time:

$$
\mathcal{A}_M = -\int \rho\sqrt{-g}d^4x \quad , \tag{8.6}
$$

where ρ is the total energy density. In order to expand the action in terms of the perturbation variables, one needs to rewrite ρ in terms of dynamical degrees of freedom which characterise the fluid flow. Hence one expresses the energy density in terms of test particles with rest mass m_0, space-time-dependent number density ϵ and a potential energy $\pi(\epsilon)$ dependent upon the pressure p:

$$
\rho = \epsilon[m_0 + \pi(\epsilon)] \quad , \tag{8.7}
$$

where, as in Fock [3], one can express the potential as:

$$
\pi(\epsilon) = \int^{\epsilon} \frac{dp}{d\epsilon}\frac{d\epsilon}{\epsilon} - \frac{p(\epsilon)}{\epsilon} \quad , \tag{8.8}
$$

and where, naturally, ϵ satisfies the continuity equation. To obtain the perturbed action, one then Taylor expands both ϵ and $\sqrt{-g}$ to second order about their background quantities:

$$
\begin{aligned}
\epsilon &= \epsilon_0 + \delta_1\epsilon + \delta_2\epsilon && (8.9)\\
\sqrt{-g} &= \sqrt{-g}_0 + \delta_1\sqrt{-g} + \delta_2\sqrt{-g} \quad , && (8.10)
\end{aligned}
$$

using the notation of equation 8.1. Inserting these results into 8.6 one obtains the second order perturbed action:

$$
\begin{aligned}
\mathcal{A}_{2M} &= -\int\left[\rho_0\frac{\delta_2\sqrt{-g}}{\sqrt{-g}_0} + (\rho_0 + p_0)\left(\frac{\delta_1\epsilon}{\epsilon_0}\frac{\delta_1\sqrt{-g}}{\sqrt{-g}_0} + \frac{\delta_2\epsilon}{\epsilon_0}\right)\right.\\
&\left.+\frac{1}{2}c_s^2(\rho_0+p_0)\frac{(\delta_1\epsilon)^2}{(\epsilon_0)^2}\right]\sqrt{-g}_0 d^4x \quad . \qquad (8.11)
\end{aligned}
$$

One then needs to evaluate the individual terms separately; this is a tedious though straight-forward exercise, and so the details, which are contained in [17], will not be included here. The final perturbed matter Lagrangian for $K = 0$ then transpires as:

$$
\begin{aligned}
\mathcal{A}_{2M} &= \int\Big(\Big[\tfrac{1}{2}\rho_0\phi^2 + p_0(\tfrac{3}{2}\gamma^2 - 3\phi\gamma + \phi\kappa_{,\alpha\alpha} - \gamma\kappa_{,\alpha\alpha} + \tfrac{1}{2}\kappa_{,\alpha\alpha}\kappa_{,\beta\beta}\\
&-\kappa_{,\alpha\beta}\kappa_{,\beta\alpha} + \tfrac{1}{2}\psi_{,\alpha}\psi_{,\beta}) + (\rho_0+p_0)(\tfrac{1}{2}\xi^{\alpha\prime}\xi^{\alpha\prime} + \Psi_\alpha\xi^{\alpha\prime} + \phi\xi^{\alpha}_{,\alpha}) \quad , \qquad (8.12)\\
&-\tfrac{1}{2}c_s^2(\rho_0+p_0)(3\gamma - \kappa_{,\alpha\alpha} - \xi^{\alpha}_{,\alpha})^2\Big]S^4 + \Sigma_M\Big)d^4x
\end{aligned}
$$

where, as before, summation over repeated lower indices is implicit; such that Σ_M represents the surface terms, and where ξ^α is the spatial part of a shift vector which shifts the position of a test particle from its background position to where it would be in the 'real' universe. Combining the above gravitational and matter Lagrangians, one then obtains:

$$
\begin{aligned}
\mathcal{A}_2 &= \mathcal{A}_{2G} + \mathcal{A}_{2M} && (8.13)\\
&= \int\left[S^2\left(-6\left[\gamma'^2 + 2\mathcal{H}\phi\gamma' + \left(\mathcal{H}^2 - \frac{\beta}{3c_s^2}\right)\phi^2\right] - 4(\gamma' + \mathcal{H}\phi)(\psi - \kappa')_{,\alpha\alpha}\right.\right.\\
&\left.-2\gamma_{,\alpha}(2\phi_{,\alpha} - \gamma_{,\alpha}) + 2\beta(\xi^{\alpha\prime} + \psi_{,\alpha})(\xi^{\alpha\prime} + \kappa_{,\alpha}) - 2\beta c_s^2\left(3\gamma - \kappa_{,\alpha\alpha} - \xi^{\alpha}_{,\alpha} + \frac{1}{c_s^2}\phi\right)^2\right)\\
&\left.+\Sigma_G + \Sigma_M + \Sigma_{GM}\right]d^4x \quad , && (8.14)
\end{aligned}
$$

where the background equations 4.121 and 4.122 have been used for p_0 and ρ_0, and $\beta = K + \mathcal{H}^2 - \mathcal{H}'$; while Σ_{GM} is a further surface term arising from joint contributions from the gravitational and matter actions. The original perturbation equations can now be obtained from the above action by varying with respect to ϕ, γ and κ respectively, whereby the following equivalent set of equations to that of 5.32, 5.33 and 5.34 (with $K = 0$) obtain:

$$
\begin{aligned}
4S^2\pi G\delta\rho &= \nabla^2\gamma - 3\mathcal{H}\gamma' - 3\mathcal{H}^2\phi - \mathcal{H}(\psi - \kappa')_{\alpha\alpha} && (8.15)\\
4S^2\pi G\delta p &= \gamma'' + 2\mathcal{H}\gamma' + \mathcal{H}\phi' + (2\mathcal{H}' + \mathcal{H}^2)\phi && (8.16)\\
4S^2\pi G\delta p &= \gamma'' + 2\mathcal{H}\gamma' + \mathcal{H}\phi' + (2\mathcal{H}' + \mathcal{H}^2)\phi\\
&+\frac{1}{3}\left\{\nabla^2(\phi - \gamma) + S^{-2}\nabla^2[S^2(\psi - \kappa')]'\right\} \quad , && (8.17)
\end{aligned}
$$

such that 8.15 is equation 5.32; 8.17 is the trace of 5.34; 8.16 incorporates the non-diagonal terms of 5.34 as well as 5.33; and where the following have been assumed:

$$\begin{aligned}\delta u^0 &= -\frac{\phi}{S} &(8.18)\\ \delta p &= \frac{dp}{d\rho}\delta\rho \\ &= c_s^2\delta\rho &(8.19)\\ \delta\rho &= (\rho_0 + p_0)(3\gamma - \kappa_{,\alpha\alpha} - \xi^{\alpha}_{;\alpha}) \quad , &(8.20)\end{aligned}$$

where the last expression is obtained from the first order approximation to the Taylor expansion of ρ with respect the the background co-ordinates and shift vector.

The Bardeen Action

The following results, due to Brandenberger *et al.* [11] formulate a specific action which yields the perturbation equations of Bardeen [10]. The action is thus a specialisation of the above. The approach can thus be seen as complimentary to that of Bardeen. It is an approach which in the process produces a natural decomposition of an arbitrary scalar field into spatially homogeneous and inhomogeneous parts; the former couples to the background metric, while the latter is a source term related to the evolution of the gauge-invariant metric perturbation potential. It is particularly useful in the study of the coupling of classical, cosmologically growing modes to quantum fluctuations in various GUT theories of matter at early stages of the universe; these matter fields in general act as source terms. Also, due to the inherent stability of vector and tensorial perturbations, as shown by Bardeen, only scalar perturbations (of a FRW metric) are considered here: these have growing inhomogeneities in the early universe as shown previously, and are the only known class of perturbations which couple to matter and have growing modes.

As motivated above, only scalar perturbations will be considered here. In this vain, as before, one starts by decomposing the space-time metric into the sum of a smooth background metric and a perturbation:

$$\begin{aligned}g_{ab} &= f_{ab} + \delta f_{ab} \\ &\equiv f_{ab} + \epsilon S^2 g^{(1)}_{ab} \quad , &(8.21)\end{aligned}$$

and where the background metric for FRW has the most general form:

$$f_{ab} = \mathrm{diag}[-\sigma^2(t), S^2(t), S^2(t), S^2(t)] \quad , \tag{8.22}$$

for an arbitrary conformal time scalar $\sigma(t)$ and scale factor $S(t)$. In the above relations, the superscript (1) indicates a perturbation of first order smallness, with parameter of smallness ϵ; background quantities such as f_{ab} being regarded as zeroth order. The scalar-perturbed metric has the standard form:

$$g^{(1)}_{ab} = \begin{bmatrix} E & F_{,\alpha} \\ F_{,\alpha} & A\delta_{\alpha\beta} + B_{,\alpha\beta} \end{bmatrix} \quad , \tag{8.23}$$

for arbitrary scalar functions A, B, E, F which are all functions of both co-ordinate time and space. As A, B, E, F are the fundamental perturbation variables, the aim here will be to obtain evolution equations for them. This will entail variation of the action with respect to these variables separately; this is done after writing the Lagrangian in terms of the metric 8.21. Analogously, one can then expand the Einstein and energy momentum tensors to first order as:

$$G_{ab} = G^{(0)}_{ab} + G^{(1)}_{ab} \tag{8.24}$$

$$T_{ab} = T^{(0)}_{ab} + T^{(1)}_{ab} \quad . \tag{8.25}$$

Note, however, that the order in the expansion of G_{ab} will depend on that of the metric, as the Einstein tensor is a function of the metric. Consequently, products such as $G^{(1)}_{ab} g^{(1)ac}$ will be regarded as second order, and thus discarded.

The procedure adopted here will be first to assume the synchronous gauge $E = F = 0$ for simplicity in deriving the equations for A, B; it will be seen that this results in no loss of generality. Secondly, the adoption of the synchronous gauge highlights the fact that E, F are not dynamical degrees of freedom, but instead correspond to the the ADM lapse function and shift vectors respectively as $E, F_{,\alpha}$: these are geometrical quantities pertaining to the space-time and are thus not dynamical degrees of freedom. Here the dynamical quantities are the remaining functions A and B. By looking then at arbitrary co-ordinate transformations, a basic gauge-invariant function of these perturbation variables will then be formulated in such a way as to correspond to a *metric potential* as formulated by Bardeen. An evolution equation will then be derived for this using the evolution equations of the perturbation variables A and B. The procedure will thus essentially amount to formulating evolution equations for the quantities A and B through variation of the following standardised gravitational action:

$$\begin{aligned} \mathcal{I} &= \int \mathcal{L}\sqrt{-g}d^4x \\ &= \int (\mathcal{L}_G + \mathcal{L}_M)\sqrt{-g}d^4x \\ &= \int \left(\frac{R}{16\pi G} + \mathcal{L}_M\right)\sqrt{-g}d^4x \quad , \end{aligned} \tag{8.26}$$

where R is the Ricci curvature scalar, and $\mathcal{L}_M$ is an arbitrary matter Lagrangian density. By substituting 8.21 and 8.22 into 8.26, one obtains a Lagrangian which will be varied with respect to the functions $S(t)$ and $\sigma(t)$, respectively. The reason for this is that the latter two functions depend only upon time; thus variation with respect to them will not explore the spatially varying degrees of freedom, thus producing natural *space-averaged* equations. The beauty of this feature is that, to lowest order in ϵ, the resultant equations will be *exactly* the FRW equations for a smooth, zero-order background; hence one does not need to assume that the matter fields consist of a smooth background plus a small inhomogeneous perturbation. The above property automatically induces a decomposition of an arbitrary matter field into a smooth part that couples to the background metric, plus a spatially varying part that couples to the metric perturbations. This is essentially the primary motivation of the approach.

Thus, performing the variation separately first with respect to $S(t)$ and then $\sigma(t)$, and omitting higher order source terms in ϵ one obtains the two fundamental cosmological equations, namely the Raychaudhuri and Friedman equations. Considering then scalar perturbations in the synchronous gauge, and thus varying the above action with respect to the two scalar perturbation variables A, B respectively, one then obtains the following perturbation evolution equations:

$$-8\pi G(S^2\delta_{\alpha\beta}T^{\alpha\beta(1)}) = (3A + {}^{(3)}\nabla^2 B)_{tt} + 3\frac{\dot{S}}{S}(3A + {}^{(3)}\nabla^2 B)_{,t} - \frac{{}^{(3)}\nabla^2 A}{S^2} \quad (8.27)$$

$$-8\pi G^{(3)}\nabla^{-2}(S^2 T^{\alpha\beta(1)}_{,\alpha\beta}) = \ddot{A} + 3\frac{\dot{S}}{S}\dot{A} \quad , \quad (8.28)$$

where $T^{\alpha\beta(1)}$ is the fist order perturbation to the stress energy tensor; ${}^{(3)}\nabla$ indicates a purely spatial gradient as before; ${}^{(3)}\nabla^{-2}$ denotes the solution to the spatial Laplace equation; the tt subscript indicates a double co-ordinate time derivative, and the 'dot' superscript indicates a single co-ordinate time derivative, as before.

These further simplify to:

$$\ddot{A} + 3\frac{\dot{S}}{S}\dot{A} = -8\pi G\mathcal{P}_1 \quad (8.29)$$

$$\ddot{B} + 3\frac{\dot{S}}{S}\dot{B} - \frac{A}{S^2} = -8\pi G\mathcal{P}_2 \quad (8.30)$$

$$, \quad (8.31)$$

where the stress tensor matter contributions have been grouped together as:

$$\mathcal{P}_1 = {}^{(3)}\nabla^{-2}(S^2 T^{\alpha\beta(1)}_{,\alpha\beta}) \quad (8.32)$$

$$\mathcal{P}_2 = (S^2\delta_{\alpha\beta}T^{\alpha\beta(1)} - 3^{(3)}\nabla^{-2}T^{\alpha\beta(1)}_{,\alpha\beta}) \quad . \quad (8.33)$$

Now, due to the assumption of a synchronous gauge, there are a further two equations required, derived from the remaining two scalar perturbation parameters E and F. These are obtained by performing a similar variational procedure to the above, yielding:

$$8\pi G T^{0(1)}_0 = \frac{{}^{(3)}\nabla^2 A}{S^2} - s\frac{\dot{S}}{S}\dot{A} - \frac{\dot{S}}{S}{}^{(3)}\nabla^2\dot{B} \quad (8.34)$$

$$\begin{aligned} {}^{(3)}\nabla^2\dot{A} &= -8\pi G S^2 T^{0\alpha(1)}_{,\alpha} \\ &\equiv -8\pi G S^2\mathcal{P}_3 \quad . \end{aligned} \quad (8.35)$$

It is now desirable to obtain a fully gauge-invariant evolution equation for the gravitational gauge-invariant metric potential Φ_H arising from scalar perturbations. As outlined in the general approach to metric perturbations, one commences by assuming a co-ordinate transformation of the form 5.40. As only scalar perturbations are being considered; this takes on the scalar gauge-transformation:

$$\bar{t} = t + f^0(t, \mathbf{x}) \tag{8.36}$$
$$\overline{x_\alpha} = x_\alpha + f_{,\alpha}(t, \mathbf{x}) \quad , \tag{8.37}$$

where f, f^0 are, as before, arbitrary functions of space and time. Proceeding then in the fashion of section 5.2, one uses this gauge transformation to calculate the scalar perturbation variables, obtaining:

$$A = 2\frac{\dot{S}}{S} f^0 \tag{8.38}$$
$$B = \frac{2}{S^2} f \tag{8.39}$$
$$E = -\frac{2}{S^2}\dot{f}^0 \tag{8.40}$$
$$F = \frac{f^0}{s^2} + \left(\frac{f}{s^2}\right)^{\cdot} \quad , \tag{8.41}$$

which subsequently yield the following feasible candidate for a gauge-invariant metric potential:

$$\Phi_H = \frac{1}{2}(A - S\dot{S}\dot{B} + 2S\dot{S}F) \quad , \tag{8.42}$$

which in the synchronous gauge simplifies to:

$$\Phi_H = \frac{1}{2}(A - S\dot{S}\dot{B}) \quad . \tag{8.43}$$

One can then proceed to derive an evolution equation for Φ_H in a somewhat unusual way, namely by assuming the following ansatz:

$$\ddot{\Phi}_H = C_1\dot{\Phi}_H + C_2\Phi_H + C_3 \quad , \tag{8.44}$$

the second order assumption being naturally inferred from the second order nature of the field equations. The co-efficients C_1, C_2, C_3 are determined by comparing the co-efficients of $\dot{A}$, $\dot{B}$ on both sides of 8.44 after substituting into 8.43. In addition, equation 8.35 is then used to eliminate the resultan t remaining gauge-dependent terms, replacing them with gauge-invariant functions of Φ_H, $\dot{\Phi}_H$ and gauge-invariant matter terms. After much tedious calculation, the final result is:

$$\begin{aligned} &\ddot{\Phi}_H + (4 + 3c_s^2)H\dot{\Phi}_H + 8\pi G\rho(c_s^2 - \Gamma)\Phi_H \\ &= 4\pi G[-\mathcal{P}_1 - 3c_s^2 H S^{2(3)}\nabla^{-2}\mathcal{P}_3 + S^2 H\dot{\mathcal{P}}_2 + 8\pi G(\tfrac{2}{3}\rho - P + \rho c_s^2)S^2\mathcal{P}_2] \quad , \end{aligned} \tag{8.45}$$

where:

$$
\begin{aligned}
H &\equiv \frac{\dot{S}}{S} && (8.46)\\
P &\equiv \frac{1}{3} < T^{\alpha}_{\alpha} > && (8.47)\\
\rho &\equiv - < T^{0}_{0} > && (8.48)\\
\Gamma &\equiv 3\frac{\ddot{S}}{S} - \left[\frac{\dot{S}}{S}\right]^2 - \frac{S}{S\dot{S}}[3\dot{S}^2 - (S\dot{S})^{\cdot}] && (8.49)
\end{aligned}
$$

and the angled brackets indicate the spatial averaging naturally contained in the formalism as explained earlier. This is now the desired, fully gauge-invariant evolution equation for the metric potential Φ_H; in it, is contained information concerning the desired inhomogeneities of the perturbations of the space-time. The initial conditions for this equation are derived in the form of an equation from the first order Einstein equation for G^{0}_{0} in terms of the perturbations A, B into which is substituted the evolution equations for E, F. This yields the dynamical equation:

$$
\Phi_H = 4\pi G S^{2\,(3)}\nabla^{-2}[T^{0(1)}_{0} - 3S\dot{S}^{(3)}\nabla^{-2}T^{\alpha(1)}_{0,\alpha}] \quad . \tag{8.50}
$$

This equation, along with its first time derivative, can be thought of as providing, for specified energy densities and energy fluxes on an initial temporal surface, a consistent set of initial conditions which are subsequently evolved in time according to 8.45. The reason for the relationship between equations 8.45 and 8.50 is that, inside the particle horizon (the region of primary interest) the matter evolution is dominated by matter self-interactions, making it thus appropriate to solve for the matter evolution equations 8.45 and infer the geometric fluctuations from equation 8.50. However, outside the horizon matter evolution is dominated by gravitational effects, resulting in it being more appropriate to solve for 8.45 in place of the dynamical equations 8.50.

Equation 8.45 is equivalent to, but not algebraicly identical with, Bardeen's equation 5.257; the two equations essentially differ by adding and subtracting a factor of $\frac{1}{S^2}{}^{(3)}\nabla^2 c_s^2 \Phi_H$. Also, the source terms in 8.45 are written in a general form as opposed to a fluid with shear stresses in Bardeen's equation. Adding the term $\frac{1}{S^2}{}^{(3)}\nabla^2 c_s^2 \Phi_H$ to both sides of equation 8.45, and using 8.50 one has a form more comparable with that of Bardeen:

$$
\begin{aligned}
&\ddot{\Phi}_H + (4 + 3c_s^2)H\dot{\Phi}_H + \left\{\left[8\pi G\rho - \tfrac{^{(3)}\nabla^2}{S^2}\right]c_s^2 - 8\pi G P\right\}\Phi_H \\
&= 4\pi G[-\mathcal{P}_1 - c_s^2 T^{0(1)}_{0} + S^2 H\dot{\mathcal{P}}_2 + 8\pi G(\tfrac{2}{3}\rho - P + \rho c_s^2)S^2\mathcal{P}_2]
\end{aligned}
\quad , \tag{8.51}
$$

where the terms on the right-hand-side of the above equation are the matter source terms. The noticeable disadvantage of equation 8.51 is that the source terms depend on the energy fluxes and density ($T_{0\alpha}$, T_{00} respectively) *as well as* the desired spatial stresses ($T_{\alpha\beta}$). Ideally, one would want a sole dependence on the stress terms as this results in physical processes at different cosmological epochs causing source terms to generate perturbations in Φ_H which subsequently grow homogeneously; i.e. cosmological perturbations today would be the sum of homogeneous solutions, each of which could be traced back to a physical 'cause'. The reason for this is that the stresses are freely specifiable at each instant of time, whereas the energy fluxes evolve from the influence of spatial stresses *over time* according to the momentum conservation equations, while the energy density is even less specifiable since it evolves in time according to the conservation equation from the accumulation of fluxes. Hence equation 8.51 contains homogeneous terms (matter stress terms) and inhomogeneous terms (energy density and flux terms); so far these have not successfully been separated.

Applications

A perfect Fluid

For the perfect fluid scenario one has the stress-energy-momentum tensor:

$$T^{ab} = (P + P^0)g^{ab(0)} + (P + P^0 + \rho + \rho^0)u^a u^b \quad . \tag{8.52}$$

Performing the above analysis one obtains a simplified form for the right-hand-side of equation 8.51, namely:

$$4\pi G \left[c_s^2 - \frac{P}{\rho} \right] \rho \quad . \tag{8.53}$$

This, however, is just an entropy perturbation, and thus vanishes if the equations of state for the background and perturbation are identical. This would imply that the gauge-invariant equation of motion for Φ_H is homogeneous, resulting in the perturbations evolving in a simple way.

A fluid with shear stresses

The stress-energy-momentum tensor of a fluid with shear stresses forms the most generic matter scenario, and has form:

$$T^{ab} = \mathrm{diag}\left(\rho, \tfrac{P}{S^2}, \tfrac{P}{S^2}, \tfrac{P}{S^2}\right) + T^{ab(1)} \quad , \tag{8.54}$$

where the perturbation $T^{ab(1)}$ takes on the form:

$$T^{ab(1)} = \begin{bmatrix} \rho\delta & (\rho + P)\frac{v_\alpha}{S^2} \\ (\rho + P)\frac{v_\alpha}{S^2} & \frac{P}{S^2}\left[(\Pi_L - \frac{1}{3}\nabla^2 \Pi_T)\delta^{\alpha\beta} + \Pi_{T,\alpha\beta} \right] \end{bmatrix} \quad , \tag{8.55}$$

such that v_α is the gradient of some scalar velocity potential; δ is the fractional energy density perturbation; Π_L is the fractional pressure perturbation; and Π_T is the fractional anisotropic stress perturbation. The above tensor is of the same form as the matter perturbations used in Bardeen. In this form, the following equivalences hold with the formalism of this paper:

$$\mathcal{P}_1 = P\Pi_L + \frac{2}{3}P\nabla^2 \Pi_T \tag{8.56}$$

$$\mathcal{P}_2 = -2P\Pi_T \quad , \tag{8.57}$$

these identifications result in the the appropriate form of the above central equation 8.51 coinciding exactly with Bardeen's equation 5.257.

At this point it is insightful to make a comparison with previous major works on perturbation theory. In particular, comparisons will be made with the Lifshitz-Khalatnikov and Bardeen approaches.

Lifshitz-khalatnikov

Noting the Lifshitz-Khalatnikov decomposition of the perturbed metric into scalar harmonics via equation 5.83, one can easily show that the variables A, B formulated above are related to their variables a, b and the scalar harmonic $Q(\mathbf{x})$ in the following way:

$$
\begin{aligned}
A &= \frac{1}{3}(a+b)Q && (8.58)\\
B &= \frac{1}{k^2}Q \quad , && (8.59)
\end{aligned}
$$

where k is the wave number. However, there are disadvantages to their approach: it is only applicable to perfect fluids as matter sources; it is only *partially* gauge-invariant; and the matter perturbations do not enter the equations for gravitational perturbations only as source terms.

Bardeen

The difference between the above approach and that of Bardeen is that Bardeen uses the energy-momentum conservation equations to derive explicit gauge-invariant equations of motion for gauge-invariant matter perturbations. Considering Bardeen's variables and those derived in this paper, namely A and B, they are related as follows:

$$
\begin{aligned}
A &= [2H_L + \frac{2}{3}H_T]Q && (8.60)\\
B &= \frac{2H_T}{k^2}Q && (8.61)\\
v_s &= v - \frac{1}{2}S^4(S^{-2}B)_{,t} && (8.62)\\
\epsilon_m &= \delta + 3(1+W)\frac{\dot{S}}{S}(v - S^2F) \quad . && (8.63)
\end{aligned}
$$

Hence one can see that the principal advantage of the approach of this paper is that the same equations as that of Bardeen can be far more simply derived via a variational approach, and that the source terms are expressed in a form that is immediately applicable to completely general matter sources as opposed to Bardeen's emphasis on fluids with shear stresses. Also, the metric potential Φ_H derived in this paper can be thought of as generating the slightly perturbed geometrical 'arena' in which arbitrary matter physics can be evolved according to its co-ordinate-dependent equations rather than forcing the matter field to be formulated in gauge-invariant form.

It is also important to note here that this simpler derivation cannot be considered as a purely independent and alternative approach; rather, it is a less physically sound derivation which owes much of its credibility, and in a more abstract sense its justification, to the more rigorous and soundly based formulation of Bardeen.

8.1.4 Covariant Perturbation Theory

Method

The evolution equations of the covariantly defined variables, when separable, simplify to second-order linear inhomogeneous equations with variable co-efficients. Such equations can be transformed into Klein-Gordon type equations with variable mass and source terms, an equation which itself has a standard Lagrangian. As mentioned earlier, this can then be used to generate the Hamiltonian and associated conserved quantities.

Consider a general second-order linear and homogeneous equation in $\Phi(t)$ with variable co-efficients:

$$\ddot{\Phi} + A\dot{\Phi} + B\Phi = 0 \quad , \tag{8.64}$$

such that A, B are functions of time. Now consider the transformation:

$$\Phi = D(t)\Psi(t) \quad . \tag{8.65}$$

Substituting this into 8.64 yields the following:

$$\ddot{\Psi} + \left[2\frac{\dot{D}}{D} + A\right]\dot{\Psi} + \left[\frac{\ddot{D}}{D} + A\frac{\dot{D}}{D} + B\right]\Psi = 0 \quad . \tag{8.66}$$

As one wishes to transform the above equation into a Klein-Gordon type equation, the co-efficient of the $\dot{\Psi}$ term needs to vanish; hence one imposes:

$$2\frac{\dot{D}}{D} + A = 0 \quad . \tag{8.67}$$

This is simply integrated with respect to t to obtain:

$$D = \exp\left(-\frac{1}{2}\int_0^t A dt'\right) \quad . \tag{8.68}$$

Hence the equation simplifies to:

$$\ddot{\Psi} + \left[B - \frac{1}{2}\dot{A} - \frac{1}{4}A^2\right]\Psi = 0 \quad . \tag{8.69}$$

Which is the required Klein-Gordon equation with (variable) mass term m and zero source term:

$$m^2 \equiv \left[B - \frac{1}{2}\dot{A} - \frac{1}{4}A^2\right] \tag{8.70}$$

Equation 8.69 thus has the standard Lagrangian:

$$\mathcal{L} = \frac{1}{2}\Psi^2\left[B - \frac{1}{2}\dot{A} - \frac{1}{4}A^2\right] - \frac{1}{2}\dot{\Psi}^2 \quad . \tag{8.71}$$

This can thus be transformed back via $\Psi = \Phi \exp\left(\frac{1}{2}\int_0^t A dt'\right)$ to obtain a direct Lagrangian for the original equation 8.64, yielding:

$$\mathcal{L} = \exp\left(\int_0^t A dt'\right)\left\{\frac{1}{2}\Phi^2\left[B - \frac{1}{2}\dot{A} - \frac{1}{2}A^2\right] - \frac{1}{2}\left[\dot{\Phi}^2 + A\Phi\dot{\Phi}\right]\right\} \quad , \tag{8.72}$$

for which one has the action:

$$\mathcal{A} = \int_{\mathcal{D}} \mathcal{L} dt \quad . \tag{8.73}$$

This method can then be applied to the second order evolution equations for the covariantly defined variables derived in Chapter 5, as this is a second order time-dependent homogeneous equation by the harmonic decomposition. Considering then equation 6.250 one has the associated Lagrangian:

$$\begin{aligned} \mathcal{L} \quad = \quad & \exp\left(\int_0^t \Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right) dt'\right)\left\{\tfrac{1}{2}\mathcal{D}_{(k)}{}^2\left[\Theta^2(c_s^2 - \tfrac{4}{3}W) - W\Lambda\right.\right. \\ & -\tfrac{\kappa\mu}{2}(1 - 3W^2) + \tfrac{c_s^2}{S^2}(k^2 - K) - \tfrac{1}{2}\Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right) \\ & \left.\left. -\tfrac{1}{2}\left(\Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right)\right)^{\cdot}\right] - \tfrac{1}{2}\left[\dot{\mathcal{D}}_{(k)}{}^2 + \Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right)\mathcal{D}_{(k)}\dot{\mathcal{D}}_{(k)}\right]\right\} \quad . \end{aligned} \tag{8.74}$$

It is important to note here that the above Lagrangian is formulated in terms of the separate harmonics of the evolution variable, and therefore yields a separate action for each harmonic.

Having calculated the Lagrangian, it is now a simple task to write down the associated Hamiltonian. Once the Hamiltonian has been calculated, application of Noether's theorem should then yield a set of conserved quantities. Using equation 3.44 one thus has the Hamiltonian for the above Klein-Gordon equation:

$$\mathcal{H} = -\frac{1}{2}\dot{\Psi}^2 - \frac{1}{2}\left[B - \frac{1}{2}\dot{A} - \frac{1}{4}A^2\right]\Psi^2 \tag{8.75}$$

where, from the definition 3.42, the canonical momentum p here is $p \equiv -\dot{\Psi}$. Hence, for equation 6.250, one has the Hamiltonian:

$$\begin{aligned} \mathcal{H} \quad = \quad & -\tfrac{1}{2}\left[\left(\mathcal{D}_{(k)}\exp\left(\tfrac{1}{2}\int_0^t \Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right) dt'\right)\right)^{\cdot}\right]^2 - \tfrac{1}{2}\left[\Theta^2(c_s^2 - \tfrac{4}{3}W)\right. \\ & -W\Lambda - \tfrac{\kappa\mu}{2}(1 - 3W^2) + \tfrac{c_s^2}{S^2}(k^2 - K) - \tfrac{1}{2}\Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right) \\ & \left. -\tfrac{1}{2}\left(\Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right)\right)^{\cdot}\right]\mathcal{D}_{(k)}{}^2\exp\left(\tfrac{1}{2}\int_0^t \Theta\left(\tfrac{2}{3} + c_s^2 - 2W\right) dt'\right) \end{aligned} \tag{8.76}$$

One can then write down an expression for the Noether current N_b using the definition 3.145 which will then be a conserved quantity by Noether's theorem 3.146.

Other Approaches

The above can be termed a 'bottom up' approach; as mentioned, this is technically the more difficult approach as it entails solving the inverse problem in the calculus of variations. A generally more practical approach would be to attempt a 'top down' strategy by starting with a standard Lagrangian, performing some expansion around a suitable background, and then seeing whether this yields the desired evolution equations after performing the variation.

Chapter 9

Conclusions

"We are all inclined to direct our inquiry not by the matter itself, but by the views of our opponents; and, even when interrogating oneself, one pushes the enquiry only to the point at which one can no longer find objections"
Aristotle

When one contemplates the major advances in physics over the past century, one invariably envisions grand and profound concepts codified in succinct mathematical formalism, and represented by readily identifiable and familiar equations, or sets of laws or models, the latter 'models' being commensurate with more contemporary paradigms. More than often, these theories contain evocative terminologies and labels which to the uninitiated person usually appear either as mystical and nebulous or idiosyncratically eccentric and peculiar. Amusingly, this latter attitude combined with general scientific ignorance has spawned a whole 'new age' pseudoscience literature which delves into the supposed deeper significance of modern scientific theories and their hypothetical relation to religion, mysticism and such. However, upon investigating the various provenances of these much celebrated scientific theories more closely, one encounters a whole plethora of long forgotten alternative theories and concepts which ultimately reached dead-ends, lost popular support from the scientific community, or were just replaced by superior ideas; indeed, together with the ultimately triumphant theories these create a general picture infused with a pervasive untidiness and confusion. This is in stark contrast to what one reads in standard texts which represent only the correct theories in the much revised, cleaned-up formulation which can only be constructed from hindsight - such a picture is misleading as firstly it portrays Science as perfectly ordered and consistent; secondly it omits much , if not most, of the toil which took many scientists years to develop their theories; and thirdly, as mentioned, it frequently declines to mention the countless failed attempts.

In light of the above, the preceding chapters have served to present such a much vaunted theory in the form of General Relativity. Much of the confusion, and perhaps mystique, surrounding this complex subject was hopefully removed by firstly elucidating upon the mathematical technicalities of its formulation *ab initio*, an attribute which is regarded by many as an insurmountable impediment to understanding such a theory; and secondly, thorough physical application via Cosmological Perturbation Theory. In focusing on the latter, an historical approach incorporating much of the technical detail of the chronological developments was adopted, in keeping with the sentiments expressed above. This was also intended, as expressed in the introduction, to create a working knowledge of the subject. In the process, the various theoretical models of Cosmology were encountered - it transpired that the closest approximation to the physically observable and quantifiable universe was the class of Friedmann-Robertson-Walker models. However, in keeping with the ultimate objective in physics, namely the understanding and interpretation of physical reality, these models had to be 'perturbed' to that which is physically observed - whence the motivation for Cosmological Perturbation Theory. Numerous forms of such perturbations were thus explored, more or less as they have appeared historically, and a number of practical formalisms related thereto were developed in the process. This is indeed one drawback to Cosmological Perturbation Theory - the proliferation of widely differing formalisms, each of which exhibiting non-trivial mathematical machinery. However, in discussing both the pros and cons of these formalisms, it became increasingly evident that depending upon what specifically was desired or being studied, a formalism peculiar to that application only was optimal. For example, if one desired a set of perturbation variables which were complete and consistent in the mathematical sense, gauge-invariant and applicable to a general space-time foliation, then the Bardeen formalism was optimal; However, the drawback was that this more than often resulted in a cumbersome formulation of *physically* meaningful variables. If on the other hand one desired naturally formulated perturbation variables in the comoving frame, then the covariant formulation was optimal, but at the price of not having a complete, well-defined set of standard perturbation quantities. Ultimately one could conclude that there does not seem to be any all-encompassing standard formalism - one has to improvise according to the application in question.

In viewing the various formalisms a number of significant results and predictions were encountered. The most noticeable of these is probably the Sachs-Wolfe effect apropos of the cosmic microwave background radiation. Although this was developed here in its original context, the years subsequent to its inception have seen a veritable revolution in Observational Cosmology, resulting in a complete revision and refinement of these early crude estimates especially in the light of new observations, starting with COBE in the early 1990's, and culminating to date with the latest BOOMERanG results. This has given modern Cosmologists strengthened belief in Big Bang theory, and has provided clearer insight into credible, physically relevant cosmological models, especially the inflationary scenario.

With an eye on the future, the exact numerical values of three highly significant cosmological quantities are still wanting: the density parameter, Hubble's constant and the cosmological constant; their significance being self-explanatory through the fundamental Raychaudhuri and Friedmann equations, as evidenced in Chapter 4. Once these have been determined the cosmological picture will be almost complete.

As motivated in the introduction, it was also the intention here to reconcile the Variational Principle with Gravitation and Cosmological Perturbation Theory. To this end a general variational formulation to gravitation was developed and extended to perturbation theory, and the pros and cons explored. However, the extensive, and now indeed quite routine, application of the Variational Calculus to almost every branch of Physics would evidently render the aims and results of the preceding chapter mere formalities. Nonetheless, in this regard the following significant point needs to be made.

As with much of Science, many principles appear far simpler in theory than in practice - this is indeed the case with the Variational Principle which, as evidenced in Chapter 3, is formulated in relatively simple mathematics, and in a fairly self-explicit manner. However, in application to a physical situation one has firstly the initial problem of motivating and formulating a meaningful Lagrangian as discussed in chapter 3, and secondly the sometimes quite horrendous ensuing mathematics in the variational process, as encountered in the previous two chapters. The latter reason forms the motivation for the extensive mathematical treatment covered in the preliminary chapters: firstly the mathematics of the Variational Calculus, intending to provide a means to understanding the principle intuitively and practically, while at the same time investigating numerous applicable auxiliary techniques and results associated thereto; and secondly the non-trivial mathematics of General Relativity, namely Differential Geometry, the technical framework of the physical theory that forms the subject of application here.

As outlined above, the Variational Principle indeed takes on an esoteric and highly abstract nature; however, as with most concepts in Science, theoretical or otherwise, an intuitive understanding thereof is optimally inculcated through application. It is intended that the preceding undertaking has indeed achieved this objective not only for the Variational Principle, but also for Cosmology in general, and that the synthesis of the two disciplines has strengthened this understanding.

Appendix A

Topology

As encountered in the opening chapter on Differential Geometry, the concept of a manifold was defined in terms of a *topological space*. A topological space $(\mathcal{X}, \mathcal{T})$ is defined as a set of elements $\mathcal{X}$ along with a collection of subsets $\mathcal{T}$ of $\mathcal{X}$ satisfying the following:

- Any arbitrary union of these sets itself forms a subset of the topology;
- Any intersection of a finite number of the subsets is itself a subset of the topology;
- The entire set $\mathcal{X}$ along with the empty set are both subsets of the topology.

In addition to the above definition, one refers the above defined sets as being *open sets of the topology*. In application to Differential Geometry, it is crucial to note that the manifold definition of Chapter 2 imposes no restrictions on the nature of the topology, only that one has 'a' topological space to start with. Secondly one notices that the Einstein field equation s determine the Geometry of space-time, *not* the topology; that is, there are no equations at all governing the topology of space-time. In the physical context of General Relativity though, the only additional assumptions made concerning the underlying topological space is that it is *Hausdorff* and *paracompact*:

- A topological space is said to be Hausdorff if for any two distinct t points in the space, one can find two open sets of the topology, each one containing one of the points such that the two sets have no intersection.

The notion of paracompactness can, for the sake of practical applications, be defined directly in terms of the manifold via the nature of the constituent atlas in terms of it being *locally finite*:

- An atlas is said to be locally finite if every point in the manifold has an open neighbourhood which intersects with only a finite number of sets in the topology.
- A manifold is then said to be paracompact if for every atlas $\mathcal{A}$ there exists a locally finite atlas $\mathcal{A}_{(fin)}$ with each open set of $\mathcal{A}_{(fin)}$ contained in an open set of $\mathcal{A}$.

In a sense, one can view this latter constraint as preventing the manifold from becoming 'too large' in a global sense. It has the important consequences of admitting a Riemannian metric to the manifold, as well as imposing second countability: i.e. one can always cover the manifold with a locally finite, countable family of charts. It also guarantees the existence of a *partition of unity*, a concept crucial for global generalisations of local results and definitions, such as for the definition of integration over a manifold, as encountered in Chapter 2. With regard to the latter, this is defined formally as follows:

Given a locally finite atlas $\mathcal{A}_{(fin)}$, a partition of unity on a manifold is defined to be a set of smooth functions $\{f_\alpha\}$ satisfying:

- The support of f_α is contained within an open set of $\mathcal{A}_{(fin)}$;
- $0 \leq f_\alpha \leq 1$;
- $\sum_\alpha f_\alpha = 1$

such that the *support* of f_α is defined to be the closure of the set upon which f_α is non-vanishing.

Appendix B

Four-Dimensional Spherical Harmonics

In full generality, one can define the harmonics as quantities which are eigenfunctions of a specific differential operator on a given well-behaved space, and such that the full such set of eigenfunctions forms a basis of that space. In Cosmology one normally utilises harmonics $\mathcal{Q}$ which are eigenfunctions of the Laplace-Beltrami operator on three-space hypersurfaces of constant curvature with spatial metric $o_{\alpha\beta}$:

$$\mathcal{Q}^{;\alpha}{}_{;\alpha} + k^2 \mathcal{Q} = 0 \quad , \tag{B.1}$$

where, by analogy with plane waves, k is interpreted as the wave number; and as usual, the covariant derivative is taken with respect to the spatial metric $o_{\alpha\beta}$; the above quantities are also assumed to be functions of the spatial variables only. Assuming then that any physical quantity $\mathcal{B}$ has separable spatial and temporal dependence, one can expand it in terms of the relevant harmonics, whether they be scalar, vectorial or tensorial as will be discussed shortly; one therefore has:

$$\mathcal{B} = \sum_k B(\tau)^{(k)} \mathcal{Q}^{(k)} \quad , \tag{B.2}$$

where $\mathcal{Q}^{(k)}$ is the k^{th} harmonic, and $B(\tau)^{(k)}$ is the 'amplitude' which has only time dependence, as assumed.

In the above formulations one has the identification $k^2 = n^2 \mp 1$, such that n corresponds to the order of the harmonic while the plus and minus signs correspond to pseudospherical and spherical cases respectively. Note that the above has been defined with respect to some, as yet unspecified, foliation of space-time; the spatial metric thereof being $o_{\alpha\beta}$ as in, for example the ADM formalism. This makes the following analysis fully general.

The construction of vectorial and tensorial quantities from the above scalar harmonics is unique up to a scalar normalisation factor; letting $\mathcal{Q}$ denote the scalar harmonics, one can construct, for example, the following two equivalent vectorial quantities:

$$\mathcal{Q}_\alpha \equiv \frac{1}{k} \mathcal{Q}_{;\alpha} \tag{B.3}$$

$$\mathcal{Q}_\alpha \equiv -\frac{1}{k^2}\mathcal{Q}_{;\alpha} \quad . \tag{B.4}$$

Note that in the Cosmological context, $\frac{k}{S}$ corresponds to the *physical* wave vector. Similarly, one may define associated tensorial quantities:

$$\mathcal{Q}^{(t)\beta}_{\alpha} \equiv \frac{1}{3}\delta^\beta_\alpha \mathcal{Q} \tag{B.5}$$

$$\mathcal{Q}^{;\beta}_{;\alpha} \equiv \frac{1}{k^2}\mathcal{Q}^{;\beta}_{\alpha} + \mathcal{Q}^{(t)\beta}_{\alpha} \quad , \tag{B.6}$$

where the first tensorial quantity is formulated so as to have a trace equal to $\mathcal{Q}$, while the second one is traceless. The *vector* harmonics $\mathcal{S}_\alpha$ can be interpreted as vectors which satisfy the vectorial Helmholtz equation:

$$\mathcal{S}^{\alpha;\beta}_{;\beta} + k^2\mathcal{S}^\alpha = 0 \quad , \tag{B.7}$$

and have, for example, associated tensorial quantities:

$$\mathcal{S}^{\alpha\beta} \equiv -\frac{1}{k}\mathcal{S}^{(\alpha;\beta)} \tag{B.8}$$

$$\mathcal{S}^{\alpha\beta} \equiv 2\mathcal{S}^{(\alpha;\beta)} \quad , \tag{B.9}$$

where, as per convention, round brackets indicate symmetrisation. Lastly, the Tensorial harmonics $\mathcal{G}_{\alpha\beta}$ obey then, by definition, the tensorial Helmholtz equation:

$$\mathcal{G}^{\alpha\beta;\gamma}_{;\gamma} + k^2\mathcal{G}^{\alpha\beta} = 0 \quad . \tag{B.10}$$

For the specific space-time foliation which is defined with respect to the matter frame with normalised four-velocity u_a and spatial metric h_{ab}, one may calculate, as with the above, the associated harmonics $\mathcal{Y}$, bearing in mind the following relations to the above formulation:

$$h_{ab} = S^2\delta^\alpha_a\delta^\beta_b o_{\alpha\beta} \tag{B.11}$$

$$h^b_a = \delta^\alpha_a\delta^b_\beta o^\beta_\alpha \tag{B.12}$$

$$h^{ab} = \frac{1}{S^2}\delta^a_\alpha\delta^b_\beta o^{\alpha\beta} \quad . \tag{B.13}$$

Hence, for any scalar quantity U, one has that:

$$^{(3)}\nabla_a U = \delta^\alpha_a U_{;\alpha} \tag{B.14}$$

$$^{(3)}\nabla^a U = \frac{1}{S^2}\delta^a_\alpha U^{;\alpha} \quad . \tag{B.15}$$

Using this, and the requirement that the harmonics $\mathcal{Y}$ be independent of co-ordinate time, one can easily verify the following relationships for the scalar harmonics:

$$\mathcal{Q} = \mathcal{Y} \tag{B.16}$$

$$\mathcal{Q}_a = S\delta_a^\alpha \mathcal{Y}_\alpha \tag{B.17}$$

$$\mathcal{Q}^a = \frac{1}{S}\delta_\alpha^a \mathcal{Y}^\alpha \tag{B.18}$$

$$\mathcal{Q}_{ab} = S^2\delta_a^\alpha \delta_b^\beta \mathcal{Y}_{\alpha\beta} \tag{B.19}$$

$$\mathcal{Q}^{ab} = \frac{1}{S^2}\delta_\alpha^a \delta_\beta^b \mathcal{Y}^{\alpha\beta} \quad , \tag{B.20}$$

as well as the following:

$$\begin{array}{rcl} \dot{\mathcal{Q}} = 0 & \Leftrightarrow & u^0\partial_0\mathcal{Y} = 0 \\ \dot{\mathcal{Q}}_a = 0 & \Leftrightarrow & u^0\partial_0\mathcal{Y}_\alpha = 0 \\ \dot{\mathcal{Q}}_{ab} = 0 & \Leftrightarrow & u^0\partial_0\mathcal{Y}_{\alpha\beta} = 0 \end{array} \quad . \tag{B.21}$$

One can verify the above also for the vector and tensor harmonics.

Bibliography

[1] EM Lifshitz. *1946 J. Phys., 10, 116*

[2] WB Bonnor. *1957 MNRAS, 117, 104*

[3] V Fock *1959 'Theory of Space, Time and Gravitation'*

[4] EM Lifshitz and IM Khalatnikov. *1963 Adv. Phys., 12, 185*

[5] F Hoyle and JV Narlikar. *1963 Proc. Roy. Soc., A273, 1*

[6] SW Hawking. *1966 Ap. J., 145, 544*

[7] RK Sachs and A.M. Wolfe *1967 Ap. J., 147, 73*

[8] JM Stewart and M Walker. *1974 Proc. Roy. Soc., A341, 49*

[9] DW Olson. *1976 Phys. Rev. D, 14, 2*

[10] JM Bardeen. *1980 Phys. Rev. D, 22, 1882*

[11] R Brandenberger, R Kahn and WH Press. *Phys. Rev. D, 28, 8*

[12] JM Bardeen. *1988 ed. Liu-Zhi Fang, 1*

[13] GFR Ellis and M Bruni. *1989 Phys. Rev. D, 40, 6*

[14] GFR Ellis, J Hwang and M Bruni. *1989 Phys. Rev. D, 40, 6*

[15] M Bruni, PKS Dunsby and GFR Ellis. *1992 Ap. J., 395, 35*

[16] PKS Dunsby, M Bruni and GFR Ellis. *1992 Ap. J., 395, 54*

[17] VF Mukhanov, HA Feldman and RH Brandenberger. *1992 Phys. Rep. 215, Nos. 5,6 203-333*

[18] GF Smoot *et al. 1992 APJ, 396, L1*

[19] JC Jackson. *1993 MNRAS, 264, 729*

[20] P Peldan. *1994 CQG, 11, 1087*

[21] GFR Ellis. *Cargese lectures in Physics* Ed. E Schatzmann (Gordon and Beach)(1973)

[22] PKS Dunsby, BACC Bassett and GFR Ellis. *1997 CQG, 14, 1215-1222*

[23] A Melchiorri *et al. 1999 New Astron. Rev., 43, 289-296*

[24] P de Bernardis *et al. 2000 Nature, 404, 955-959*

[25] T Gebbie, PKS Dunsby and GFR Ellis. *2000 An. Phys., 282, 2*

[26] J Bernstein. *Kinetic theory in expanding universes*(Cambridge)

[27] SR de Groot, WA van Leeuwen and CG van Weert. *Relativistic kinetic theory*(North-Holland)(1980)