Acceleration of the Cosmological Expansion as an Effect of Inhomogeneities

Teppo Mattsson

Helsinki Institute of Physics
P.O. Box 64, FIN-00014 University of Helsinki, Finland

ACADEMIC DISSERTATION
To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium CK112 of Exactum, Gustaf Hällströmin katu 2b, on March 27th, 2009, at 2 p.m.

Helsinki 2009
Ihminen joka elää luonnossa, tuottaa luonnostaan luonnollisia ajatuksia.

Kati Sinenmaa, kosmoksen asiantuntija.
# Contents

Abstract .................................................. iii  
Acknowledgements ........................................ iv  
List of included papers ................................... v  

1 Introduction ............................................ 1  

2 Classical gravitation ..................................... 3  
  2.1 Preliminaries .......................................... 3  
  2.1.1 The equivalence principle .............................. 3  
  2.1.2 Space, time and spacetime ........................... 4  
  2.1.3 Tensors .............................................. 4  
  2.1.4 Connection ........................................... 5  
  2.1.5 Metric ............................................... 6  
  2.1.6 Special relativity .................................... 7  
  2.1.7 Riemannian curvature as the gravitational field .... 8  
  2.2 The gravitational field equations ...................... 9  
  2.2.1 Geodesics ............................................ 10  
  2.3 Averaging in general relativity ....................... 10  
  2.3.1 The scalar parts of the Einstein equation .......... 11  
  2.3.2 Non-commutativity of time-evolution and averaging .. 13  

3 A solution of general relativity: Lemaître-Tolman-Bondi metric 17  
  3.1 Lemaître-Tolman-Bondi (LTB) spacetime ................. 17  
  3.2 Light propagation in the LTB spacetime ................ 22  
  3.3 Classifying inhomogeneities in the LTB solution ....... 23  
  3.4 Useful analytic formulae of the LTB solution .......... 24  
  3.5 Bubbleessence ......................................... 26  

4 Cosmological observations and the standard FRW model .......... 28  
  4.1 Cosmological observations ............................. 28  
  4.1.1 Cosmic microwave background ........................ 28  
  4.1.2 Large-scale galaxy distribution ...................... 31  
  4.1.3 Type Ia supernovae .................................. 32  
  4.1.4 The unforeseen message: accelerated expansion along our line of sight .. 35  
  4.2 Spatially homogeneous and isotropic FRW models ........ 35  
  4.3 Dark energy and dark matter ........................... 36  
  4.3.1 Problems with dark energy driven acceleration .... 37
5 Inhomogeneous generalizations of the FRW cosmology

5.1 Accelerated expansion from cosmic inhomogeneities ........................................ 39
5.1.1 Backreaction due to inhomogeneous expansion ........................................ 40
5.1.2 Special location in space: a local void ..................................................... 41
5.1.3 Selection effects due to opaque structures ............................................... 42
5.1.4 Total effect of the inhomogeneities ......................................................... 43
5.1.5 The generalized Dyer-Roeder method ...................................................... 44

5.2 A concordant inhomogeneous model of the universe ....................................... 47
5.2.1 The cosmic microwave background ....................................................... 48
5.2.2 Galaxy distribution ................................................................................. 48
5.2.3 Type Ia supernovae ............................................................................... 50
5.2.4 The Hubble Key Project ......................................................................... 50
5.2.5 The Big Bang Nucleosynthesis .............................................................. 51
5.2.6 A concordant nonlinear CHDM model with no dark energy ............... 51
5.2.7 Future improvements of the model ......................................................... 53

6 Summary ............................................................................................................. 55

A Power expansions of the LTB scale function .................................................. 58
Abstract

The cosmological observations of light from type Ia supernovae, the cosmic microwave background and the galaxy distribution seem to indicate that the expansion of the universe has accelerated during the latter half of its age. Within standard cosmology, this is ascribed to dark energy, a uniform fluid with large negative pressure that gives rise to repulsive gravity but also entails serious theoretical problems. Understanding the physical origin of the perceived accelerated expansion has been described as one of the greatest challenges in theoretical physics today.

In this thesis, we discuss the possibility that, instead of dark energy, the acceleration would be caused by an effect of the nonlinear structure formation on light, ignored in the standard cosmology. A physical interpretation of the effect goes as follows: due to the clustering of the initially smooth matter with time as filaments of opaque galaxies, the regions where the detectable light travels get emptier and emptier relative to the average. As the developing voids begin to expand the faster the lower their matter density becomes, the expansion can then accelerate along our line of sight without local acceleration, potentially obviating the need for the mysterious dark energy.

In addition to offering a natural physical interpretation to the acceleration, we have further shown that an inhomogeneous model is able to match the main cosmological observations without dark energy, resulting in a concordant picture of the universe with 90% dark matter, 10% baryonic matter and 15 Gyr as the age of the universe. The model also provides a smart solution to the coincidence problem: if induced by the voids, the onset of the perceived acceleration naturally coincides with the formation of the voids. Additional future tests include quantitative predictions for angular deviations and a theoretical derivation of the model to reduce the required phenomenology.

A spin-off of the research is a physical classification of the cosmic inhomogeneities according to how they could induce accelerated expansion along our line of sight. We have identified three physically distinct mechanisms: global acceleration due to spatial variations in the expansion rate, faster local expansion rate due to a large local void and biased light propagation through voids that expand faster than the average. A general conclusion is that the physical properties crucial to account for the perceived acceleration are the growth of the inhomogeneities and the inhomogeneities in the expansion rate. The existence of these properties in the real universe is supported by both observational data and theoretical calculations. However, better data and more sophisticated theoretical models are required to vindicate or disprove the conjecture that the inhomogeneities are responsible for the acceleration.
Acknowledgements

This thesis is based on research carried out in the Helsinki Institute of Physics, at the University of Helsinki. The financial support from the Helsinki University Science Foundation, the Magnus Ehrnrooth Foundation and the EU 6th Framework Marie Curie Research and Training network “UniverseNet” (MRTN-CT-2006-035863), is gratefully acknowledged.

I thank my supervisor Kari Enqvist for his guidance and support. Several lecturers in theoretical physics deserve special thanks for their inspiring courses; in particular, I acknowledge Juha Honkonen, Claus Montonen, Hannu Koskinen and Hannu Kurki-Suonio. I have been lucky to be influenced by their enthusiasm in physics.

It is a pleasure to thank Tony Liimatainen, Sami Nurmi, Vappu Reijonen and Tomi Koivisto for several enjoyable and enlightening discussions in and outside of physics, as well as Kimmo Kainulainen, Syksy Räsänen, Tuomas Multamäki, Subir Sarkar, Morad Amarzguioui and Håvard Alnes for valuable discussions mainly about cosmology. In addition to the ones mentioned above, an incomplete list of friends and colleagues that I also thank includes Aleksi, Anna-Stiina, Ari, Janne, Juhani, Matti, Olli, Reijo, Timo, Vesa and Ville. Naturally, my beloved wife deserves the warmest thanks for always being there for me.

Finally, I express my gratitude to my mother for her support and patience.

Teppo Mattsson
Helsinki, November 2008
List of included papers

The three articles included in this thesis are:


The contribution of the present author to the joint publications

In the first publication, I had the idea of studying light propagation in the inhomogeneous LTB model. I also performed the analytic calculations and the numerical computations. K. Enqvist helped in writing and suggested to apply the averaging formalism to the LTB model, as well as provided valuable supervision.

In the second paper, both the calculations and the writing were done jointly with M. Mattsson (nee Ronkainen).
Chapter 1

Introduction

Centuries of observations and theoretical research have revealed to us that natural phenomena can be understood as an interplay between four fundamental interactions: gravitation, electromagnetism, strong and weak forces. Whereas gravitation and electromagnetism have infinite range, the strong and weak forces become important only at a subatomic length scale or at very high energies. Thus, presuming the cosmos is electrically neutral, the cool and sparse late universe forms perhaps the purest gravitational system possible. Furthermore, being a large and cool system, it seems plausible to assume that, at cosmic scales, the late universe is governed by the laws of classical physics.

Einstein’s general theory of relativity, or general relativity for short, is the state-of-the-art description of classical gravitation. Since its completion in 1915, the theory has provided several novel predictions that precision measurements have confirmed and, on the other hand, it has not been shown to be in disagreement with any observation [1]. General relativity is a deterministic theory: given the initial conditions in the early universe, the later evolution is uniquely determined. Nevertheless, there are two practical issues that make the solving of this evolution challenging: 1) we do not know the initial conditions of the universe, and 2) being an extremely complex system, it is unfeasible to calculate the time evolution of the universe using the exact equations.

A popular hypothesis in modern cosmology is that the initial conditions in the hot early universe were (nearly) Gaussian random fluctuations around a homogeneous background, as predicted by the simplest models for the origin of structure [2]. Another widely adopted simplifying assumption has been to treat the universe as spatially homogeneous and isotropic with the growth of structure described as linear perturbations evolving on the smooth background. These assumptions have made it possible to calculate the evolution of the universe and make predictions for cosmological observables. Despite their simplicity, such models have provided good fits for various cosmological observations.

However, since the late 1990’s, the cosmological observations of light from type Ia supernovae [3], the cosmic microwave background [4] and the galaxy distribution [5] have brought us an unforeseen message: during the latter half of its age, the expansion of the universe seems to have increased roughly by a factor of 3/2 along our line of sight, relative to the expectation from the homogeneous and isotropic Friedman-Robertson-Walker (FRW) models of the universe based on general relativity and ordinary energy forms.
Understanding the physical origin of the perceived accelerated expansion has been described as one of the greatest challenges in theoretical physics today. The standard conception has been that it is because 75% of the energy in the universe consists of mysterious dark energy with large negative pressure that gives rise to repulsive gravity [6–8]. However, the enormous fine-tuning needed to explain both the size and the timing of such an energy component has raised serious doubts about its correctness and thus motivates the search for alternatives.

In this thesis, we discuss the possibility that, instead of dark energy, the perceived increase in the expansion rate would be an effect of nonlinear inhomogeneities, ignored in the standard FRW description. The idea is based on the fact that the inhomogeneities can have a similar effect on the cosmological observations of light as dark energy. Although it has gained more attention only after the acceleration was detected, the actual idea of using more sophisticated models than the simple FRW model to describe the real universe is not a new one; see [9,10]. After all, there appears to be little or no justification beyond mere simplicity to calculate the observable properties of light from the homogeneous and isotropic FRW metric. In other words, the evidence for dark energy is based on the unproved assumption that the effects of the observed nonlinear inhomogeneities on the detectable light average out over cosmological distances.

We have identified three suitable, physically distinct mechanisms for how the cosmic inhomogeneities could lead to accelerated expansion: 1) spatial variations in the expansion rate inducing global acceleration, 2) faster local expansion rate due to a large local void and 3) biased light propagation through voids that expand faster than the average. After introducing the necessary background in Chapters 2–4, we elaborate the mechanisms in Chapter 5. In Papers 1–3 [11–13], the inhomogeneities have been discussed from various aspects: Paper 1 studies supernova observations in spherically symmetric inhomogeneous models, Paper 2 concentrates on scale-dependent volume averages and Paper 3 considers a phenomenological fit to various cosmological observations. In Chapter 6, we summarize the main results of the research.
Chapter 2

Classical gravitation

In this chapter, we introduce Einstein’s general relativity – the classical theory of gravity. This material forms the theoretical foundation for the thesis, while in the rest of the chapters we then apply Einstein’s theory to the study of the inhomogeneous late universe.

2.1 Preliminaries

Before discussing gravitational physics, we first set up some necessary formalism. The aim in this section is to introduce the part of mathematics of relativity essential to this thesis and also to establish our conventions.

2.1.1 The equivalence principle

Mathematically, the description of the classical gravitational field is based on Riemannian geometry, that is the geometry of curved spaces. A heuristic way to motivate the use of such formalism in describing gravitation is to consider the observed equivalence of inertial and gravitational masses.

The mass of an object has two properties: firstly, it defines the inertia of the object, i.e. how difficult it is to change the state of motion of the body and secondly, it acts as the gravitational charge, i.e. it is responsible for the gravitational field of the object. In principle, the gravitational mass and the inertial mass would not need to have equal values; however, they have been observed to be equal to an accuracy of $\sim 10^{-13}$ [14, 15]. The hypothesis that these two concepts of mass have exactly equal values carries the name of the equivalence principle.

The equivalence principle implies that, in a given gravitational field, any object falls with the same acceleration under the influence of gravity$^1$. This also implies that a point-like observer can feel no gravitational force. A real observer of finite size, instead, feels inhomogeneities of the gravitational field as tidal forces deforming her body.

Apparently, it was the equivalence principle that led Einstein to the idea of describing gravitation with non-Euclidean geometry: as objects with different mass fall identically under gravity, they

---

$^1$However, as all objects gravitate, it is non-trivial to arrange objects with different masses in exactly identical fields, because their (different) masses change the gravitational field (differently); this is a consequence of the nonlinearity of gravity, discussed more in Sect. 2.2.
can be thought to move in spacetime like particles sliding on a curved surface, tending towards the deeps. This obviously intertwines gravitation tightly together with space and time.

2.1.2 Space, time and spacetime

Space and time are intuitive concepts: space provides the stage for happenings while time measures their duration and defines the chronology. Time is usually considered to flow inevitably forwards, whereas in space one is free to move back and forth. However, when discussing the universe at a theoretical level, time is just another coordinate that can be wound back and forth as well. Moreover, due to the finite speed of light, by looking far away in the universe, we look into the past – not to our past, but the past of the other parts of the universe which, assuming statistical spatial homogeneity, can provide us information about the average properties of the universe from different cosmological eras.

A fundamental aspect of general relativity is that together time and space form an entity, a four-dimensional manifold known as the spacetime, whose points are called events. What is more, the presence of objects curves the spacetime so that its geometry becomes non-Euclidean. This curvature is what we perceive as gravity.

Before discussing the curvature of spacetime, we must understand the concept of a manifold or, more precisely, a Riemannian manifold. A Riemannian manifold \( M \) is a set of points endowed with mathematical structures that make operations such as derivatives etc. well defined. Formally, \( M = (S, \tau, A, g, \nabla) \), where \( S \) is the set of all events, \( \tau \) is the topology of the spacetime, \( A \) is an atlas or a set of coordinate systems whose domains together cover the whole spacetime, \( g \) is the metric of the spacetime and \( \nabla \) is a connection. The connection and the metric are introduced in Sects. 2.1.4 and 2.1.5; for an account of different topologies and atlases accessible to physicists, we refer to Szekeres’ textbook [16].

In cosmology, there is a desire to divide the spacetime into temporal and spatial parts. The division leads to the issue of dependence on the time-slicing: in order to take spatial averages, we must artificially break the symmetry between time and space inherent in general relativity. This becomes more explicit when discussing spatial averages in Sect. 2.3.

2.1.3 Tensors

Some physical quantities are specified by a single number, such as speed or mass, but many require a set of numbers, such as velocity or gravitational field. These sets of numbers form objects that are called tensors. In line with this terminology, single-number quantities or scalars are called tensors of zeroth rank. Similarly, vectors are called first rank tensors. In practice, however, by tensors one usually refers to tensors of rank two or higher.

A general physical requirement, commonly attributed to general relativity, is that physics should not depend on the coordinates used. Thus, it is natural to demand tensors to be coordinate-independent or invariant.

Let \( G \) be an invariant tensor of rank two, representing some physical entity. Although independent of coordinates, for explicit computations we must however employ a coordinate system, in which the tensor can be given in terms of a basis:

\[ G = G^\mu_{\nu} \partial_\mu \otimes dx^\nu, \]  

(21)
where \( \otimes \) is a tensor product, \( \partial_\mu \equiv \partial / \partial x^\mu \) and \( dx^\nu \) represent basis vectors and basis dual vectors respectively, such that \( \partial_\mu(dx^\nu) = dx^\nu(\partial_\mu) = \delta^\nu_\mu \).

In general, vectors are linear mappings from dual vectors to scalars and vice versa. Moreover, as the notation of the basis vectors indicates, a vector \( X \) has also the character of a differential operator: it maps any smooth scalar function \( f \) to another smooth scalar function \( Xf = f' \).

Tensors of higher rank can be formed by taking tensor products of \( n \) vectors and \( m \) dual vectors, defined such that the result is a multilinear map, called a \((n, m)\) tensor, from \( m \) vectors and \( n \) dual vectors to scalars. The components of a \((n, m)\) tensor have \( n \) upper and \( m \) lower indices and it is said to be of rank \( n + m \).

A common practice is to identify the components of the tensor \( G^\mu_\nu \) with the tensor itself. This is pragmatic, as long as one appreciates the fundamental difference between them: whereas \( G \) remains invariant under a coordinate transformation from \( \{ x^\alpha \} \) to \( \{ \tilde{x}^\alpha \} \), the relation between the components of the different coordinate systems is given by the Jacobians:

\[
G^\mu_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\sigma} \frac{\partial \tilde{x}^\sigma}{\partial x^\nu} \tilde{G}^\alpha_\sigma ,
\]

as can be seen by requiring the invariance of \( G \) in Eq. (2.1) and applying the obvious relations between the basis vectors:

\[
\partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial \tilde{x}^\sigma}{\partial x^\mu} \frac{\partial}{\partial \tilde{x}^\sigma} = \frac{\partial \tilde{x}^\sigma}{\partial x^\mu} \tilde{\partial}_\sigma \\
dx^\nu = \frac{\partial x^\nu}{\partial \tilde{x}^\sigma} d\tilde{x}^\sigma .
\]

More generally, the transformation relation has always a Jacobian for each upper index and an inverse Jacobian for each lower index.

### 2.1.4 Connection

On a manifold, each point \( p \) has its own vector space \( T_p \) in which the vectors\(^2\) located at \( p \) live in. Thus, to define a coordinate-independent way to compare tensors at different points – for example in order to take covariant derivatives – we need an operator that connects tensors between different points \( p \neq q \). This operator is called a connection and denoted by \( \nabla \). It takes two vectors \( X \) and \( Y \) as arguments and returns a vector, \( \nabla_X Y = Z \). Given a vector \( X \) as the first argument, the connection provides a derivative operator, called a covariant derivative \( \nabla_X \), directed along the vector \( X \). In addition to being linear in both of its arguments, the connection is defined to satisfy the following properties:

\[
\nabla_X f = Xf ,
\]

where \( f \) is a scalar function and

\[
\nabla_X(Y \otimes Z) = (\nabla_X Y) \otimes Z + Y \otimes (\nabla_X Z) .
\]

For a given basis \( \{ \partial_\mu \} \), we define the connection coefficients \( \Gamma^\alpha_\mu_\nu \) via the relation

\[
\nabla_\mu \partial_\nu = \Gamma^\alpha_\mu_\nu \partial_\alpha ,
\]

\(^2\)By a vector \( X \), we really mean a smooth vector field: an assignment of a vector \( X_p \) to each point \( p \) such that for every smooth scalar function \( f \), \( Xf \) is smooth and similarly for tensors of higher rank.
where $\nabla_{\nu} \partial_{\nu} \equiv \nabla_{\partial_{\nu}} \partial_{\nu}$. By using Eqs. (2.5)-(2.7), we obtain a coordinate-expression for the covariant derivative of a $(1,0)$ tensor $V$:

$$
\nabla_{\mu} V^{\alpha} = \partial_{\mu} V^{\alpha} + \Gamma_{\mu\nu}^{\alpha} V^{\nu} .
$$  \hspace{1cm} (2.8)

Note the common abuse of notation: by $\nabla_{\mu} V^{\nu}$ we really mean the $\nu$th component of the derivative $(\nabla_{\mu} V)^{\nu}$, *not* the derivative of the $\nu$th component (which would simply reduce to the partial derivative $\partial_{\mu} V^{\nu}$).

The covariant derivative can be generalized to operate also on dual vectors through the relation:

$$
\nabla_{\mu}(V(w)) = (\nabla_{\mu} V)w + V(\nabla_{\mu} w) ,
$$  \hspace{1cm} (2.9)

where $V$ is a vector and $w$ a dual vector. Since $V(w)$ is just a scalar, we can combine Eqs. (2.5) and (2.9) to obtain the coordinate-expression for the covariant derivative of a $(0,1)$ tensor $w$:

$$
\nabla_{\mu} w_{\alpha} = \partial_{\mu} w_{\alpha} - \Gamma_{\mu\alpha}^{\sigma} w_{\sigma} .
$$  \hspace{1cm} (2.10)

More generally, Eqs. (2.6), (2.8) and (2.10) imply that the covariant derivative of a $(p,q)$ tensor has $p$ $\Gamma$-terms with positive sign and $q$ $\Gamma$-terms with negative sign. For example, the covariant derivative of a $(1,1)$ tensor has the expression:

$$
\nabla_{\mu} T_{\nu}^{\alpha} = \partial_{\mu} T_{\nu}^{\alpha} + \Gamma_{\mu\alpha}^{\sigma} T_{\nu}^{\sigma} - \Gamma_{\alpha\nu}^{\sigma} T_{\mu}^{\sigma} .
$$  \hspace{1cm} (2.11)

### 2.1.5 Metric

The metric $g$ of a spacetime is a symmetric $(0,2)$ tensor, i.e. $g_{\mu\nu} = g_{\nu\mu}$, that determines the spacetime interval as follows:

$$
d s^2 = g_{\mu\nu} dx^\mu dx^\nu .
$$  \hspace{1cm} (2.12)

The metric also defines the inner products of vectors $X \cdot Y = g_{\mu\nu} X^\mu Y^\nu$, implying a connection between quantities with lower and upper indices: $X_{\mu} = g_{\mu\nu} X^\nu$ and $X^\mu = g^{\mu\nu} X_{\nu}$, where $g^{\mu\nu}$ is the inverse metric, defined such that $g_{\mu\nu} g^{\nu\sigma} = \delta_{\mu}^{\sigma}$. Physically, the metric is a gravitational tensor potential that replaces Newton’s scalar potential $\Phi$.

The requirements (2.5) and (2.6) do not yet single out a unique connection. In general relativity, the further requirements for the connection to be compatible with the metric,

$$
\nabla_X g = 0 ,
$$  \hspace{1cm} (2.13)

and symmetric with respect to the lower indices,

$$
\Gamma_{\mu\nu}^{\alpha} = \Gamma_{\nu\mu}^{\alpha} ,
$$  \hspace{1cm} (2.14)

determine a unique connection. More specifically, Eqs. (2.13) and (2.14) imply that the connection can be given in terms of the metric as follows:

$$
\Gamma_{\mu\nu}^{\alpha} = \frac{1}{2} g^{\alpha\sigma} (\partial_{\mu} g_{\nu\sigma} + \partial_{\nu} g_{\mu\sigma} - \partial_{\sigma} g_{\mu\nu}) ,
$$  \hspace{1cm} (2.15)

in which case the quantities $\Gamma_{\mu\nu}^{\alpha}$ are known as the Christoffel symbols. Thus, in standard general relativity, the connection and the metric are not independent fields but are related by Eq. (2.15).
Being a \((0, 2)\) tensor, the components of the metric obey the transformation relation:
\[
g_{\mu\nu} = \frac{\partial \tilde{x}^\sigma}{\partial x^\mu} \frac{\partial \tilde{x}^\alpha}{\partial x^\nu} \tilde{g}_{\sigma\alpha} .
\]  
(2.16)

The determinant of Eq. (2.16) yields:
\[
det[g_{\mu\nu}] = \left( \det \left[ \frac{\partial \tilde{x}^\sigma}{\partial x^\mu} \right] \right)^2 \det[\tilde{g}_{\sigma\alpha}] .
\]  
(2.17)

Since the volume-element \(d^n x\) of an \(n\)-dimensional spacetime transforms with the Jacobian determinant
\[
d^n x = \det \left[ \frac{\partial x^\mu}{\partial \tilde{x}^\nu} \right] d^n \tilde{x} ,
\]  
(2.18)

Eq. (2.17) implies an important relation between the volume-elements of different coordinate systems:
\[
\sqrt{\left| \det[g_{\mu\nu}] \right|} d^n x = \sqrt{\left| \det[\tilde{g}_{\mu\nu}] \right|} d^n \tilde{x} .
\]  
(2.19)

2.1.6 Special relativity

The framework for physical theories in the spacetime with no gravity is known as special relativity. An experimental basis of special relativity is that the speed of light\(^3\) \(c\) is the same in all inertial frames. For a homogeneous and isotropic spacetime, this implies that the spacetime interval
\[
ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu
\]
(2.20)

has the same value in all inertial coordinate systems and the same form in all inertial frames with cartesian coordinates. Eq. (2.20) can be regarded as a generalization of Pythagoras’ theorem in the spacetime with zero gravity. The symbol \(\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)\) is known as the Minkowski metric.

In an accelerating or non-inertial frame, the spacetime interval takes the more general form
\[
ds^2 = g_{\mu\nu}(x^\sigma) dx^\mu dx^\nu ,
\]
(2.21)

where \(g_{\mu\nu}\) represent the components of the metric tensor which can now be functions of the spacetime point. As long as gravity is absent, there is always a coordinate transformation that brings the line-element (2.21) globally to the Minkowskian form (2.20). However, with nonzero gravity it is no longer possible to transform the line-element \textit{globally} to the form (2.20), though it can be done \textit{locally} around any event, implying that special relativity is always applicable to local physics. In other words, albeit locally indistinguishable from coordinate artifacts, globally gravity is a real force. This is perhaps most evident in cosmology, where the expansion of space is actually properly understood as a gravitational phenomenon.

\(^3\)We employ the units in which \(c = 1\) or \(299792458\) m = 1 s.
2.1.7 Riemannian curvature as the gravitational field

In general relativity, the gravitational field is described by a rank four tensor, the Riemann curvature tensor, defined such that the equation

\[ [\nabla_\mu, \nabla_\nu] V^\alpha = R^\alpha_{\sigma \mu \nu} V^\sigma \]  
(2.22)

holds for any vector field \( V^\alpha \). By applying Eq. (2.11) to Eq. (2.22), we obtain the expression for the quantities \( R^\alpha_{\sigma \mu \nu} \) in terms of the Christoffel symbols:

\[ R^\alpha_{\sigma \mu \nu} = \partial_\mu \Gamma^\alpha_{\nu \sigma} - \partial_\nu \Gamma^\alpha_{\mu \sigma} + \Gamma^\alpha_{\mu \lambda} \Gamma^\lambda_{\nu \sigma} - \Gamma^\alpha_{\nu \lambda} \Gamma^\lambda_{\mu \sigma}. \]  
(2.23)

The Riemann tensor contains the full information about the curvature of spacetime. Eqs. (2.15) and (2.23) tell us that the Riemann tensor is given by the metric and its first and second derivatives, in accordance with the interpretation of the metric as a gravitational tensor potential. As the metric thus uniquely determines the spacetime curvature, it can be regarded as a fundamental variable.

In \( n \) dimensions, the Riemann tensor has \( n^4 \) components but, due to symmetries, not all of them are independent. A convenient way to see the symmetries is to consider the Riemann tensor with lower indices \( R_{\alpha \sigma \mu \nu} = g_{\alpha \tau} R^\tau_{\sigma \mu \nu} \). The following symmetry relations hold:

\[ R_{\alpha \sigma \mu \nu} = -R_{\alpha \mu \sigma \nu} \]  
(2.24)

\[ R_{\alpha \sigma \mu \nu} = -R_{\sigma \alpha \mu \nu} \]  
(2.25)

\[ R_{\alpha \sigma \mu \nu} + R_{\alpha \nu \sigma \mu} + R_{\mu \alpha \nu \sigma} = 0 \]  
(2.26)

\[ R_{\alpha \sigma \mu \nu} = R_{\mu \nu \sigma \alpha}. \]  
(2.27)

Note, however, that the last symmetry (2.27) follows from the first three.

The symmetries (2.24)–(2.26) imply that in \( n \) dimensions, there are \( N(n) = n^2(n^2 - 1)/12 \) independent quantities \( R^\alpha_{\sigma \mu \nu} \), so \( N(4) = 20 \) for the four-dimensional spacetime. In addition to the algebraic symmetries of the Riemann tensor, which constrain the number of independent components at any point, it obeys a differential identity, which constrains its relative values at different points. This so-called Bianchi identity reads as:

\[ \nabla_{[\lambda} R_{\alpha \sigma] \mu \nu} = 0, \]  
(2.28)

where the square-bracketed indices denote antisymmetrization:

\[ A_{[\mu \nu]} \equiv \frac{1}{2!} (A_{\mu \nu} - A_{\nu \mu}) \]  
(2.29)

\[ B_{[\mu \nu \sigma]} \equiv \frac{1}{3!} (B_{\mu \nu \sigma} - B_{\nu \mu \sigma} + B_{\sigma \mu \nu} - B_{\sigma \nu \mu} + B_{\nu \sigma \mu} - B_{\mu \sigma \nu}) \]  
(2.30)

and similarly for objects with more indices.

One can decompose \( R^\alpha_{\sigma \mu \nu} \) into the trace

\[ R_{\sigma \nu} \equiv R^\alpha_{\alpha \sigma \nu}, \]  
(2.31)

\[ R_{\sigma \nu} \equiv R^\alpha_{\alpha \sigma \nu}. \]  

called the Ricci tensor, and the traceless part

\[ C_{\alpha\sigma\mu
u} = R_{\alpha\sigma\mu
u} - \frac{2}{(n-2)} (g_{\alpha\mu} R_{\nu\sigma} - g_{\alpha\nu} R_{\mu\sigma}) + \frac{2}{(n-1)(n-2)} g_{\alpha\mu} g_{\nu\sigma} R \, , \]  

(2.32)
called the Weyl tensor, where \( n = 4 \) hereafter and we have further defined the Ricci scalar

\[ R \equiv g^{\mu\nu} R_{\mu\nu} \, . \]  

(2.33)

It is customary to define a trace-reversed Ricci tensor or Einstein tensor

\[ G^{\mu\nu} \equiv R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu} \, , \]  

(2.34)

which, as can be seen by contracting the Bianchi identity (2.28) twice, has the nice property of being covariantly constant:

\[ \nabla_\mu G^{\mu\nu} = 0 \, . \]  

(2.35)

### 2.2 The gravitational field equations

We have now set up the required machinery to discuss the quantitative relation between the curvature of spacetime and the objects living in it. The local part of this relation is provided by the Einstein equation

\[ G^{\mu\nu} = 8\pi G T^{\mu\nu} \, , \]  

(2.36)

where \( T^{\mu\nu} \) is the energy-momentum tensor, and the non-local part by a corollary of the Bianchi identity (2.28):

\[ \nabla^{\alpha} C_{\alpha\sigma\mu\nu} = 8\pi G \left( \nabla_\mu T_\nu + \frac{1}{3} g_{\sigma\nu} \nabla_\mu T^\sigma \right) \, , \]  

(2.37)

where Eq. (2.36) has been used to eliminate the derivatives of the Einstein tensor. Eqs. (2.36) and (2.37) reveal that the trace of the Riemann tensor or the Einstein tensor is determined by the local value of the non-gravitational energy-momentum, whereas the traceless part or the Weyl tensor corresponds to the part of the gravitational field caused by the non-local sources.

We note that \( T^{\mu\nu} \) contains only the non-gravitational energy-momentum. In this thesis, we do not consider the notoriously difficult issue of the relativistic gravitational energy-momentum; for a review, see [17, 18]. It has, however, been suggested by Wiltshire that gravitational energy may be of central relevance also to the dark energy problem [19–24].

In the presence of matter, a unique specification of the gravitational field requires eight boundary condition functions, given for example on a spacelike hypersurface or along the light cone. The functions can be e.g. the mass distribution, the three components of its velocity distribution plus the two polarization modes of the gravitational waves and their time derivatives. There are other choices as well, but, in the absence of symmetries, the required number of independent functions is nevertheless always the same. Given the eight independent boundary condition functions, the spacetime dynamics is then governed by the field equations (2.36) and (2.37).

When comparing the gravitational field e.g. with the electromagnetic field, a fundamental difference occurs: whereas the Maxwell equations that govern the dynamics of electromagnetism form a set of linear differential equations, the gravitational field equations (2.36) and (2.37) are nonlinear. Physically, this can be understood as a consequence of the fact that the gravitational
field itself carries energy, which is the source of gravity, whereas the electromagnetic field is chargeless and hence not a source of electromagnetism. At a quantum level, this means that the carriers of the electromagnetic interaction or photons are electrically neutral, whereas the hypothetical gravitons carry energy and thus lead to gravitational self-interactions that show up as nonlinearity in the field equations. The non-linearity has far-reaching consequences, e.g., for the time evolution of volume averages, discussed in Sect. 2.3.2.

2.2.1 Geodesics

The gravitational field equations imply that test objects, or objects with negligible energy, follow geodesics of the spacetime [25]. Geodesics are curves that parallel transport their own tangent vectors:

\[
\frac{D}{d\lambda} \frac{dx^\alpha}{d\lambda} \equiv \frac{dx^\mu}{d\lambda} \nabla_\mu \frac{dx^\alpha}{d\lambda} = 0 ,
\]

(2.38)

where \( D/d\lambda \) is a covariant directional derivative along a curve parameterized by an affine parameter \( \lambda \). The requirement (2.38) yields the geodesic equation in a coordinate basis:

\[
\frac{d^2x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0 ,
\]

(2.39)

which holds for both null and timelike geodesics with, however, the additional constraint \( ds^2 = 0 \) for null geodesics.

Strictly speaking, objects with finite energy do not follow geodesics of a fixed background due to the influence of the object itself on the spacetime. Treating small bodies as test objects is, however, often an excellent approximation.

2.3 Averaging in general relativity

For example in cosmology, one has to describe a complex inhomogeneous and anisotropic spacetime in terms of a coarse-grained model that is simple enough to solve. A central question is then how to average relativistic gravitational systems within Einstein’s theory.

As pointed out by Shirokov and Fisher [26] and later made more popular in the observational cosmology programme by Ellis and collaborators [27–30], it seems physically more correct to first calculate the Einstein field \( G(g) \) for the exact metric \( g \) and only then average \( \langle G(g) \rangle \), than to calculate the Einstein field for an averaged metric \( G(\langle g \rangle) \) as in the standard FRW approach. The reason is that the Einstein field is more closely related to physical quantities whereas the metric corresponds to a gravitational tensor potential, whose derivatives determine the physics.

In general relativity, the field \( G \) depends non-linearly on the metric \( g \), so its evaluation does not commute with averaging and the non-vanishing difference, \( \langle G(g) \rangle - G(\langle g \rangle) \neq 0 \), gives rise to what is known as the non-linear backreaction; see [31–33]. Hence the issue is not only a conceptual one, but in general, the two approaches yield identical predictions only in the absence of nonlinear inhomogeneities.

Even if performed in the above explained order, the averaging is not free of problems. Firstly, the only practicable way to average tensors seems to be the manifestly coordinate-dependent averaging of components: \( \langle G \rangle \equiv \langle G_{\mu\nu} \rangle \). As we would like the coordinate-invariance of general relativity to hold also for the coarse-grained or averaged quantities, only the inherently invariant
rank 0 tensors or scalars appear to have well-defined averages. Secondly, the division of the spacetime into temporal and spatial parts leads to the issue of dependence on the time-slicing; in order to take spatial averages, we must artificially break the symmetry between time and space inherent in general relativity.

As demonstrated by Buchert [34], in an irrotational dust universe, these problems can be alleviated by the physically well justified choice\(^4\) of using proper time of dust to determine the time-slicing and by averaging only scalars. Therefore, we only consider averages of the scalar parts of the Einstein tensor equation (2.36).

2.3.1 The scalar parts of the Einstein equation

By assuming that a fluid with four-velocity \(u^\mu\) exists at each spacetime point, one can separate out three independent scalar equations from the Einstein tensor equation, using the following three operators: the fluid four-velocity \(u^\mu\), the metric \(g_{\mu\nu}\) and the covariant derivative \(\nabla_\mu\). We consider only a single-component pressureless fluid or dust, as that appears to be the simplest and most relevant case for the late universe. The energy-momentum tensor then takes the form:

\[
T^{\mu\nu} = \rho u^\mu u^\nu ,
\]

where \(\rho\) is the energy density of the dust in its rest frame and is thus a scalar. This reduces the Einstein equation (2.36) to the form:

\[
G_{\mu\nu} = 8\pi G \rho u_{\mu} u_{\nu} .
\] (2.41)

In accordance with the twice contracted Bianchi identity (2.35), we have also the energy-momentum continuity equation for dust:

\[
\nabla_\mu (\rho u^\mu u^\nu) = 0 .
\] (2.42)

Since rotation is expected to be important only at small scales, we consider only irrotational dust, in which case the metric can be written in synchronous coordinates [10]:

\[
ds^2 = -dt^2 + g_{ij}dx^i dx^j ,
\] (2.43)

where the sum over the Latin indices \((i, j)\) runs from 1 to 3.

To find the scalar parts of the Einstein equation, we first consider a division of the spacetime into spatial slices of constant time coordinate \(t\). The four-velocity of the dust \(u^\mu\) is then orthogonal to the spatial hypersurfaces and normalized by definition:

\[
u = \frac{dx^\mu}{d\tau} \frac{dx_{\mu}}{d\tau} = \frac{ds^2}{d\tau^2} = -1 .
\] (2.44)

Due to the absence of pressure (gradients), the fluid particles follow geodesics so we also have

\[
u^{\mu} \nabla_\mu u^\nu = 0 .
\] (2.45)

With the help of the projection tensor

\[
P_{\mu\nu} = g_{\mu\nu} + u_{\mu} u_{\nu} ,
\] (2.46)

\(^4\)Also supported by a comparison with exact observables performed in Paper 2.
we define the extrinsic curvature tensor as
\[ K_{\mu\nu} = P_{\mu}^{\alpha} P_{\nu}^{\beta} \nabla_{\alpha} u_{\beta} . \]  
(2.47)

Due to the relations (2.44) and (2.45), Eq. (2.47) reduces to
\[ K_{\mu\nu} = \nabla_{\mu} u_{\nu} . \]  
(2.48)

A useful expression for the extrinsic curvature is
\[ K_{\mu\nu} = \sigma_{\mu\nu} + \frac{1}{3} \theta P_{\mu\nu} , \]  
(2.49)

where the shear tensor \( \sigma_{\mu\nu} \) and the expansion scalar \( \theta \) are defined as:
\[ \sigma_{\mu\nu} = \frac{1}{2} (\nabla_{\mu} u_{\nu} + \nabla_{\nu} u_{\mu}) - \frac{1}{3} (g_{\mu\nu} + u_{\mu} u_{\nu}) \nabla_{\alpha} u^{\alpha} \tag{2.50} \]
\[ \theta \equiv \nabla_{\mu} u^{\mu} . \tag{2.51} \]

In the synchronous coordinates of Eq. (2.43), the expansion scalar has the expression:
\[ \theta = \partial_{\mu} u^{\mu} + \Gamma_{\mu\nu}^{\mu} u^{\nu} = \Gamma_{\mu}^{\mu0} = \frac{\partial_{\mu} \sqrt{\det[g_{ij}]} }{\sqrt{\det[g_{ij}]}}, \tag{2.52} \]

where in the first step we have used Eq. (2.8).

Gauss’ equation relates the Ricci scalar of the spatial 3-surfaces, \( (3)R \), to the spacetime Ricci scalar \( R \) as follows [10]:
\[ (3)R = R + 2R_{\mu\nu} u^{\mu} u^{\nu} - K^{2} + K^{ij} K_{ij} , \tag{2.53} \]

where \( K \equiv K^{i}_{i} \). From Eq. (2.47) we see that \( K^{2} = \theta^{2} \) and \( K^{ij} K_{ij} = 2\sigma^{2} + \theta^{2}/3 \), where \( \sigma^{2} \equiv \sigma^{ij} \sigma_{ij}/2 \). Plugging these into Eq. (2.53) yields
\[ R_{\mu\nu} u^{\mu} u^{\nu} + \frac{1}{2} R = \frac{1}{2} (3)R + \frac{1}{3} \theta^{2} - \sigma^{2} . \tag{2.54} \]

By contracting then the Einstein equation for dust (2.41) on both sides by \( u^{\mu} u^{\nu} \), we obtain
\[ R_{\mu\nu} u^{\mu} u^{\nu} + \frac{1}{2} R = 8\pi G \rho . \tag{2.55} \]

Substituting Eq. (2.54) into Eq. (2.55) finally yields the first scalar equation,
\[ \frac{1}{3} \theta^{2} = 8\pi G \rho - \frac{1}{2} R + \sigma^{2} , \tag{2.56} \]

known as the Hamiltonian constraint.

Let us then derive the second scalar equation. Multiplying the definition of the Riemann curvature tensor (2.22) by \( u^{\nu} \) and setting \( \alpha = \mu \), yields
\[ u^{\nu} \nabla_{\alpha} \nabla_{\nu} u^{\alpha} - u^{\nu} \nabla_{\nu} \theta = R_{\sigma\nu} u^{\sigma} u^{\nu} . \tag{2.57} \]

The first term on the left can be written as
\[ u^{\nu} \nabla_{\alpha} \nabla_{\nu} u^{\alpha} = \nabla_{\alpha} (u^{\nu} \nabla_{\nu} u^{\alpha}) - \nabla_{\nu} u^{\alpha} \nabla_{\alpha} u^{\nu} . \tag{2.58} \]
With the help of the geodesic equation (2.45), Eq. (2.58) simplifies to

\[ u^\nu \nabla_\rho u^\rho = -\nabla_\rho u_\nu \nabla^\mu u^\nu \, . \]  
(2.59)

Moreover,

\[ \nabla_\rho u_\nu \nabla^\mu u^\nu = K^\mu_\nu K_\rho^\nu = 2\sigma^2 + \frac{1}{3}\theta^2 \, , \]  
(2.60)

where we have used the tracelessness of the shear tensor \( \sigma^\mu_\mu = 0 \), manifest in Eq. (2.50), and the result \( \sigma^\mu_\nu u^\nu = 0 \), which follows from Eqs. (2.44), (2.45) and (2.48).

By inserting Eqs. (2.59) and (2.60) into Eq. (2.57), we obtain

\[ \dot{\theta} + 2\sigma^2 + \frac{1}{3}\theta^2 = -R_{\sigma\nu}u^\sigma u^\nu \, , \]  
(2.61)

where \( \dot{\theta} \equiv u^\nu \nabla_\mu \theta = \partial_t \theta \). Furthermore, substituting Eq. (2.61) into Eq. (2.55) yields

\[ \dot{\theta} = -2\sigma^2 - \frac{1}{3}\theta^2 - 4\pi G \rho \, , \]  
(2.62)

where we have used \( R = 8\pi G \rho \), which follows from the trace of the Einstein equation (2.41). Eq. (2.62) is called the Raychaudhuri equation [35, 36].

The third scalar equation, obtained simply by contracting the energy-momentum continuity equation (2.42) by \( u_\nu \) and using the normalization equation (2.44) and the geodesic equation (2.45), reads as:

\[ \dot{\rho} + \theta \rho = 0 \, , \]  
(2.63)

where \( \dot{\rho} \equiv u^\mu \nabla_\mu \rho = \partial_t \rho \).

We finally collect together the three exact, local, covariant equations (2.56), (2.62) and (2.63):

\[ \frac{1}{3}\theta^2 = 8\pi G \rho - \frac{1}{2}(3)R + \sigma^2 \]  
(2.64)

\[ \dot{\theta} = -2\sigma^2 - \frac{1}{3}\theta^2 - 4\pi G \rho \]  
(2.65)

\[ \dot{\rho} = -\theta \rho \, . \]  
(2.66)

Note that all the quantities in Eqs. (2.64)–(2.66) have both spatial and temporal dependence and can contain arbitrary large inhomogeneities. The price to pay for reducing the Einstein tensor equation to a set of scalar equations is that the system is not closed: there are three equations for four independent variables.

### 2.3.2 Non-commutativity of time-evolution and averaging

We define the spatial average of a quantity \( S(x, t) \) as follows:

\[ \langle S(x, t) \rangle_D \equiv \frac{\int_D S(x, t) M(x, t) d^3x}{\int_D M(x, t) d^3x} \, , \]  
(2.67)

where \( M(x, t) d^3x \) is an integration measure and \( D \) an averaging domain, i.e. a portion of the spatial sections. To assure coordinate-invariance of the averages, it follows from Eq. (2.19) that we must have \( M(x, t) = \sqrt{\det[g_{ij}]} \), where \( g_{ij} \) is the spatial metric in Eq. (2.43).
The definition (2.67) implies the non-commutativity of time-evolution and averaging as a fundamental property for systems with a time-dependent integration measure, that is

\[ \frac{\partial}{\partial t} \langle f(x, t) \rangle_D \neq \langle \frac{\partial}{\partial t} f(x, t) \rangle_D , \tag{2.68} \]

if \( \partial_t M(x, t) \neq 0 \). The exact commutation relation between the time-derivative and spatial averaging can be obtained by taking the time-derivative of Eq. (2.67):

\[
\partial_t \langle S \rangle_D = \frac{\int_D \sqrt{\det|g_{ij}|} \partial_t S + S \partial_t (\sqrt{\det|g_{ij}|})}{\int_D \sqrt{\det|g_{ij}|} d^3x} \biggl[ \frac{\int_D \sqrt{\det|g_{ij}|} d^3x}{\int_D \sqrt{\det|g_{ij}|} d^3x} \biggr]^3 - \frac{\int_D S \sqrt{\det|g_{ij}|} d^3x}{\int_D \sqrt{\det|g_{ij}|} d^3x} \biggl[ \frac{\int_D \sqrt{\det|g_{ij}|} d^3x}{\int_D \sqrt{\det|g_{ij}|} d^3x} \biggr]^3 , \tag{2.69} \]

which can be simplified by writing it in terms of the expansion scalar (2.52):

\[
\partial_t \langle S \rangle_D = \langle \partial_t S \rangle_D + \langle S \theta \rangle_D - \langle S \rangle_D \langle \theta \rangle_D . \tag{2.70} \]

In order to compare the average evolution with homogeneous and isotropic FRW models, we have to define a scale factor. The simplest way to generalize the notion of an overall scale factor to an inhomogeneous and anisotropic spacetime is to use the volume of the spatial hypersurface:

\[
a_D(t) \equiv \left( \frac{\int_D \sqrt{\det|g_{ij}|} d^3x}{\int_D \sqrt{\det|g_{ij}(t = t_0)|} d^3x} \right)^\frac{4}{3} . \tag{2.71} \]

By using the definition of the average (2.67), we can write the average of the expansion rate (2.51) in terms of the scale factor (2.71) as

\[
\langle \theta \rangle_D = \frac{\int_D 3 d^3x \partial_t (\sqrt{\det|g_{ij}|})}{\int_D d^3x \sqrt{\det|g_{ij}|}} = 3 \frac{\dot{a}_D}{a_D} \equiv 3 \mathcal{H}_D , \tag{2.72} \]

where we have introduced the domain dependent Hubble expansion function \( \mathcal{H}_D \).

Let us then take the average of the scalar equations (2.64)–(2.66), using Eq. (2.67) and commuting the time derivatives with the help of Eq. (2.70), to obtain the equations satisfied by the scale factor (2.71). These equations were derived by Buchert in 1999 [34]; we follow the convention introduced by Räisänen to call them the Buchert equations [37].

We start with the Hamiltonian constraint, Eq. (2.64):

\[
\frac{1}{3} \langle \theta^2 \rangle_D = 8\pi G \langle \rho \rangle_D - \frac{1}{2} \langle (3) R \rangle_D + \langle \sigma^2 \rangle_D . \tag{2.73} \]

By defining the backreaction,

\[
Q_D(t) \equiv \frac{2}{3} \left( \langle \theta^2 \rangle_D - \langle \theta \rangle_D^2 \right) - 2 \langle \sigma^2 \rangle_D , \tag{2.74} \]

and using the relation between the average expansion rate and the scale factor (2.72), Eq. (2.73) simplifies to:

\[
3 \left( \frac{\dot{a}_D}{a_D} \right)^2 = 8\pi G \langle \rho \rangle_D - \frac{1}{2} \langle (3) R \rangle_D - \frac{1}{2} Q_D . \tag{2.75} \]
Eq. (2.75) is a generalization of the Friedman equation (4.11).

Next consider the average of the Raychaudhuri equation (2.65):

$$ \langle \partial_t \theta \rangle_D = -2 \langle \sigma^2 \rangle_D - \frac{1}{3} \langle \theta^2 \rangle_D - 4 \pi G \langle \rho \rangle_D . \tag{2.76} $$

Commuting the time derivative in Eq. (2.76) according to the relation (2.70) yields

$$ \partial_t \langle \theta \rangle_D = -2 \langle \sigma^2 \rangle_D + \frac{2}{3} \langle \theta^2 \rangle_D - \langle \theta \rangle_D^2 - 4 \pi G \langle \rho \rangle_D . \tag{2.77} $$

Furthermore, we use Eq. (2.74) in Eq. (2.77) to obtain

$$ \partial_t \langle \theta \rangle_D = Q_D - \frac{1}{3} \langle \theta \rangle_D^2 - 4 \pi G \langle \rho \rangle_D . \tag{2.78} $$

Then, by inserting Eq. (2.72) in Eq. (2.78), we have the Buchert acceleration equation:

$$ 3 \frac{\ddot{a}_D}{a_D} = -4 \pi G \langle \rho \rangle_D + Q_D . \tag{2.79} $$

Finally, the average of the third scalar equation (2.66) reads as

$$ \langle \partial_t \rho \rangle_D + \langle \theta \rho \rangle = 0 . \tag{2.80} $$

With the help of Eqs. (2.70) and (2.72), we obtain a continuity equation for the averages:

$$ \partial_t \langle \rho \rangle_D + 3 \frac{\dot{a}_D}{a_D} \langle \rho \rangle_D = 0 . \tag{2.81} $$

Let us then collect together the Buchert equations (2.75), (2.79) and (2.81):

$$ 3 \frac{\ddot{a}_D(t)}{a_D(t)} = -4 \pi G \langle \rho \rangle_D(t) + Q_D(t) \tag{2.82} $$

$$ 3 \left( \frac{\dot{a}_D(t)}{a_D(t)} \right)^2 = 8 \pi G \langle \rho \rangle_D(t) - \frac{1}{2} (3) R_D(t) - \frac{1}{2} Q_D(t) \tag{2.83} $$

$$ \partial_t \langle \rho \rangle_D(t) = -3 \frac{\dot{a}_D(t)}{a_D(t)} \langle \rho \rangle_D(t) . \tag{2.84} $$

Note that the difference between the Buchert acceleration equation (2.82) and its dust FRW counterpart (Eq. (4.9) with $p = 0$) is given by the backreaction $Q_D(t)$, defined in Eq. (2.74). We discuss the implications of these equations in Chapter 5.

In the context of cosmological averaging, with the overall dynamics of the universe given by Eqs. (2.82), (2.83) and (2.84), one can make the approximation that the average metric takes the form:

$$ ds^2 = -dt^2 + a_D^2(t) \left[ \frac{dr^2}{1 - k_D(t)r^2} + r^2(d\theta^2 + \sin^2 \theta d\varphi^2) \right] , \tag{2.85} $$

which makes it possible to calculate estimates for the observable properties of light; see e.g. [38, 39]. Although, on spatial slices, the form of the metric (2.85) is the same as in the homogeneous and isotropic FRW universe, the time evolution of the scale factor $a_D(t)$ and the spatial curvature $k_D(t)$ are in general different from the FRW case, where $a(t)$ is determined by the Friedman
equation (4.8) and $k = \text{constant}$. To obtain more accurate estimates for the observable properties of light in the inhomogeneous and anisotropic universe, the calculations should be done from the first principles [40], instead of using the average metric (2.85).

In addition to the backreaction, compared to the corresponding FRW equations, there is also another difference in the averaged equations (2.82)–(2.84): the explicit dependence on the smoothing scale $D$. We have studied the implications of the scale-dependence on the observations in Paper 2.
Chapter 3

A solution of general relativity: Lemaître-Tolman-Bondi metric

Although the relativistic gravitational field equations, discussed in Sect. 2.2, form an extremely complex set of non-linear differential equations, it is possible to solve them exactly under some simplifying symmetries. A particular case that has a wide range of applicability in inhomogeneous cosmology is spatial spherical symmetry. The exact spherically symmetric dust solution of the Einstein equations was discovered by Lemaître in 1933 [41]. We have exploited this solution extensively in Papers 1 and 2.

3.1 Lemaître-Tolman-Bondi (LTB) spacetime

Consider a spacetime with a worldline around which the spatial sections are spherically symmetric. The general form of the line-element in coordinates respecting the spherical symmetry can be written as

\[ ds^2 = -b^2(\tilde{r}, \tilde{t})d\tilde{r}^2 + c(\tilde{r}, \tilde{t})d\tilde{r}d\tilde{t} + \tilde{X}^2(\tilde{r}, \tilde{t})d\tilde{t}^2 + \tilde{A}^2(\tilde{r}, \tilde{t})(d\theta^2 + \sin^2 \theta d\phi^2) \]  

(3.1)

where \( b(\tilde{r}, \tilde{t}), c(\tilde{r}, \tilde{t}), \tilde{X}(\tilde{r}, \tilde{t}) \) and \( \tilde{A}(\tilde{r}, \tilde{t}) \) are functions of the radial and time coordinates and the spatial origin \( \tilde{r} = 0 \) has been chosen as the symmetry center. The coordinates \( r \) and \( t \) can be subjected to a transformation \( \tilde{r} = f_r(r, t), \; \tilde{t} = f_t(r, t) \) that makes the term \( c(\tilde{r}, \tilde{t})d\tilde{r}d\tilde{t} \) vanish. Furthermore, the spherical symmetry implies vanishing vorticity so that in the case of a dust universe, we can choose the time coordinate \( t \) to measure the proper time of the comoving fluid. The line-element takes then the form

\[ ds^2 = -dt^2 + X^2(r, t)dr^2 + A^2(r, t)(d\theta^2 + \sin^2 \theta d\phi^2) \]  

(3.2)

The nonzero Christoffel symbols (2.15) for the metric of Eq. (3.2) are

\[
\begin{align*}
\Gamma^0_{11} &= X\tilde{X} & \Gamma^0_{22} &= A\tilde{A} & \Gamma^0_{33} &= A\tilde{A}\sin^2 \theta \\
\Gamma^1_{10} &= \Gamma^1_{01} = \frac{\tilde{X}}{X} & \Gamma^1_{11} &= \frac{X'}{X} & \Gamma^1_{22} &= -\frac{AA'}{X^2} & \Gamma^1_{33} &= -\frac{AA'}{X^2}\sin^2 \theta \\
\Gamma^2_{02} &= \Gamma^2_{20} = \frac{\tilde{A}}{A} & \Gamma^2_{12} &= \Gamma^2_{21} = \frac{A'}{A} & \Gamma^2_{33} &= -\sin \theta \cos \theta
\end{align*}
\]
\[ \Gamma_{03}^3 = \Gamma_{30}^3 = \frac{\dot{A}}{A}, \quad \Gamma_{13}^3 = \Gamma_{31}^3 = \frac{A'}{A}, \quad \Gamma_{23}^3 = \Gamma_{32}^3 = \cot \theta, \]

where \( f' \equiv \partial_r f \) and \( \dot{f} \equiv \partial_t f \). Calculating the Einstein tensor (2.34) from the above connection coefficients gives the following nonvanishing components:

\[
\begin{align*}
G_0^0 &= \frac{2A''}{AX^2} - \frac{2AX}{AX} - \frac{2A'X'}{AX^3} - \frac{(1 + \dot{A}^2)}{A^2} + \frac{A'^2}{A^2} X^2, \\
G_0^1 &= \frac{2A'}{A} - \frac{2A\dot{X}}{AX} - \frac{(1 + \dot{A}^2)}{A^2}, \\
G_1^1 &= \frac{A'^2}{A^2 X^2} - \frac{2\ddot{A}}{A} - \frac{(1 + \dot{A}^2)}{A^2}, \\
G_2^2 &= G_3^3 = \frac{A''}{AX^2} - \frac{\dot{A}}{A} - \frac{\ddot{X}}{X} - \frac{A\dot{X}' - A'X'}{AX^3}.
\end{align*}
\]  

(3.3)

The energy momentum tensor in the above defined coordinates is given by

\[ T_{\mu}^\nu = -\rho_M(r, t)\delta_{0}^{\mu}\delta_{0}^{\nu} - \rho_A\delta_{\mu}^{\nu}, \]

(3.7)

where \( \rho_M(r, t) \) is the energy density of dust, \( u^\mu = \delta_0^\mu \) represents the components of the 4-velocity-field of the fluid and we have kept the vacuum energy \( \rho_A \) for generality.

When Eqs. (3.3)-(3.7) are applied to the Einstein equation (2.36), the following four algebraically different equations arise:

\[
-2 \frac{A''}{AX^2} + 2 \frac{A'X'}{AX^3} + 2 \frac{\ddot{A}X}{AX} + \frac{1}{A^2} + \left( \frac{\dot{A}}{A} \right)^2 - \left( \frac{A'}{AX} \right)^2 = 8\pi G(\rho_M + \rho_A)
\]

(3.8)

\[
\dot{A}' = \frac{A'\ddot{X}}{X}
\]  

(3.9)

\[
2 \frac{\ddot{A}}{A} + \frac{1}{A^2} + \left( \frac{\dot{A}}{A} \right)^2 - \left( \frac{A'}{AX} \right)^2 = 8\pi G\rho_A
\]

(3.10)

\[
- \frac{A''}{AX^2} + \frac{\dot{A}}{A} - \frac{\ddot{X}}{X} + \frac{A'X'}{AX^3} + \frac{\dot{X}}{X} = 8\pi G\rho_A.
\]

(3.11)

By solving the functions \( \ddot{X} \) and \( \ddot{X} \) from Eq. (3.9), and the functions \( A'^2 \) and \( A'' \) from Eq. (3.10), we find that Eq. (3.11) is trivially satisfied. Thus, only three of the differential equations (3.8)-(3.11) are independent. Furthermore, we can solve Eq. (3.9) to obtain

\[
X(r, t) = \frac{A'(r, t)}{\sqrt{1 - \frac{k(r)}}},
\]

(3.12)

where \( k(r) \) is a function determined by the boundary conditions. Although staying at fixed spatial coordinates, the fluid can move physically in the radial direction; this movement is encoded in \( \sqrt{g_{11}} = A'(r, t)/\sqrt{1 - k(r)} \).

With the help of Eq. (3.12), the components of the full Riemann tensor (2.23) for the LTB metric
(3.2) become:

\[
R^0_{\ 101} = \frac{A' \dot{A}'}{1 - k} \\
R^0_{\ 202} = \dot{A} \dot{A} \\
R^0_{\ 303} = R^0_{\ 202} \sin^2 \theta \\
R^1_{\ 212} = \frac{1}{2 \lambda} (2 \dot{A}' + k') \\
R^1_{\ 313} = R^1_{\ 212} \sin^2 \theta \\
R^2_{\ 323} = (\dot{A}' + k) \sin^2 \theta ,
\]

where the unlisted components are either zero or related to (3.13)–(3.18) by the symmetries (2.24)–(2.26). Similarly, Eq. (3.12) simplifies the Einstein tensor of the LTB solution to:

\[
G^0_0 = -\dot{A}^2 + k \cdot \frac{2 \dot{A}' + k'}{AA'} \\
G^1_1 = -2 \dot{A} \cdot \frac{\dot{A}^2 + k}{A^2} \\
G^2_2 = G^3_3 = -\frac{\dot{A}}{A} - \frac{\dot{A}'}{A'} - \frac{2 \dot{A}' + k'}{2AA'} .
\]

Clearly, the Riemann tensor (3.13)–(3.18) has four, and the Einstein tensor (3.19)–(3.21) three algebraically independent components, so the Weyl tensor (2.32) of the LTB spacetime contains the one remaining independent degree of freedom. This represents the non-local part of the gravitational field caused by the spherically symmetric spatial inhomogeneities.

The relation (3.12) also reduce the number of independent differential equations in (3.8)–(3.11) from three to two:

\[
\frac{\ddot{A}^2 + k(r)}{A^2} + \frac{2 \ddot{A}' + k'(r)}{AA'} = 8\pi G (\rho_M + \rho_\Lambda) \\
\ddot{A}^2 + 2 \ddot{A} + k = 8\pi G \rho_\Lambda A^2 ,
\]

The first integral of Eq. (3.23) is

\[
\dot{A}^2 = \frac{F(r)}{A^3} \cdot \frac{8\pi G}{3} \rho_\Lambda - \frac{k(r)}{A^2} ,
\]

where \(F(r)\) is a non-negative function. Substituting Eq. (3.24) into Eq. (3.22) gives

\[
\frac{F'}{AA^2} = 8\pi G \rho_M .
\]

By combining Eqs. (3.22) and (3.23) we can also construct a generalized acceleration equation:

\[
\frac{2 \ddot{A}}{3 A} + \frac{1}{3 A} \frac{\dot{A}'}{A'} = -\frac{4\pi G}{3} (\rho_M - 2\rho_\Lambda) .
\]

This equation tells us that the total acceleration, represented by the left hand side, is everywhere negative unless the vacuum energy is large enough: \(\rho_\Lambda > \rho_M / 2\). However, it does not exclude the possibility of having radial acceleration or \(\dot{A}'(r, t) > 0\), even in the pure dust universe, if
the angular scale factor $A(r, t)$ is decelerating enough and vice versa. Already a simple example like this demonstrates how the very notion of accelerated expansion becomes ambiguous in the presence of the inhomogeneities; see also Ref. [42] for a discussion.

The boundary condition functions $F(r)$ and $k(r)$ are specified by the exact inhomogeneity profile, most conveniently given on a spatial hypersurface at a fixed time $t = t_0$. Their resemblance to the more familiar quantities – the Hubble constant $H_0$ and the density parameter $\Omega_M$ – can be recognized by comparing Eq. (3.24) with the Friedman equation (4.11). Indeed, the similarity between Eqs. (3.24) and (4.11) motivates us to define a local angular Hubble rate

$$H(r, t) \equiv \frac{\dot{A}(r, t)}{A(r, t)}, \quad \text{(3.27)}$$

and matter density through

$$F(r) \equiv H_0^2(r)\Omega_M(r)A_0^2(r), \quad \text{(3.28)}$$

with

$$k(r) \equiv H_0^2(r)(\Omega_M(r) + \Omega_A(r) - 1)A_0^2(r), \quad \text{(3.29)}$$

where $A_0(r) \equiv A(r, t_0)$, $H_0(r) \equiv H(r, t_0)$, $\Omega_A(r) \equiv 8\pi G \rho_A / 3H_0^2(r)$ and $\Omega_M(r)$ is a measure of the matter density at $t = t_0$, given by:

$$\Omega_M(r) \equiv \frac{\rho_M(r, t_0)}{\rho_{\text{crit}}(r, t_0)} \equiv \frac{8\pi G}{3H_0^2(r)} \int_{B(r)} \frac{\rho_M(r, t_0) d^3r}{\int_{B(r)} d^3r}, \quad \text{(3.30)}$$

where $\int_{B(r)}$ denotes spatial integration over an origin-centered ball of radius $r$. With these definitions, Eq. (3.24) takes the physically more transparent form

$$H^2(r, t) = H_0^2(r) \left[ \Omega_M(r) \left( \frac{A_0}{A} \right)^3 + \Omega_A(r) + \Omega_c(r) \left( \frac{A_0}{A} \right)^2 \right], \quad \text{(3.31)}$$

where $\Omega_c(r) \equiv 1 - \Omega_A(r) - \Omega_M(r)$. The difference between the homogeneous Friedman equation (4.11) and its LTB generalization (3.31) is that all the quantities in the LTB case depend on the $r$-coordinate. This is true even for the scaling freedom of the scale function: whereas in the FRW case the present value of the scale factor $a(t_0)$ can be chosen to be any positive number, the corresponding present-day scale function of the LTB model $A_0(r)$ can be chosen to be any smooth and invertible positive function. In this thesis, we employ the conventional choice:

$$A_0(r) = r. \quad \text{(3.32)}$$

Although the vacuum energy density $\rho_A$ is constant, its value in the units of critical density $\Omega_A(r) \equiv \rho_A / \rho_{\text{crit}}(r)$ is not. This is because the critical density itself has spatial dependence: $\rho_{\text{crit}}(r) \equiv 3H_0^2(r) / 8\pi G$. The converse is also true: if e.g. $\Omega_M(r) = \text{constant}$, the matter distribution $\rho_M(r, t)$ itself has spatial dependence as long as $H_0(r) \neq \text{constant}$.

The integration of Eq. (3.31) w.r.t. time gives the second integral of Eq. (3.23):

$$t_0 - t = \frac{1}{H_0(r)} \int_{A_0(r)}^{A(r(t_0))} \frac{dx}{x\sqrt{\Omega_M(r)x^{-3} + \Omega_c(r)x^{-2} + \Omega_A(r)}}, \quad \text{(3.33)}$$

20
For any point in the LTB spacetime with coordinates \((t, r, \theta, \varphi)\), Eq. (3.33) gives the function \(A(r, t)\) and all its derivatives. This uniquely determines the metric

\[
ds^2 = -dt^2 + \frac{A^2(r, t)}{1 - k(r)} \, dr^2 + A^2(r, t) (d\theta^2 + \sin^2 \theta d\varphi^2), \tag{3.34}
\]

so given the inhomogeneities \(H_0(r)\) and \(\Omega_M(r)\) (and \(\rho_{\Lambda}\)), all the observable quantities can be computed. The metric (3.34) was first studied by Lemaître [41], Tolman [43] and Bondi [44]; later, it has been used in various astrophysical contexts [9, 10]. The homogeneous FRW metric is only a special case of Eq. (3.34), obtained in the limit: \(A(r, t) \to a(t)r\) and \(k(r) \to kr^2\), where \(a(t)\) is the FRW scale factor and \(k\) is the curvature constant.

In general, Eq. (3.33) has to be integrated numerically but for some special cases, it can also be done analytically. We list here the results that can be given in terms of elementary functions:

1. \(\Omega_M(r) = 1\) and \(\Omega_{\Lambda}(r) = 0\):

\[
A(r, t) = A_0(r) \left[ 1 + \frac{3}{2} (t - t_0) H_0(r) \right]^{2/3}.
\tag{3.35}
\]

2. \(\Omega_M(r) \equiv \Omega(r) > 1\) and \(\Omega_{\Lambda}(r) = 0\):

\[
(t - t_0) H_0(r) = \frac{\Omega(r)}{(\Omega(r) - 1)^{3/2}} \left[ \arcsin \sqrt{\frac{\Omega(r) - 1}{\Omega(r)}} \left( \frac{A(r, t)}{A_0(r)} \right) - \arcsin \sqrt{\frac{\Omega(r) - 1}{\Omega(r)}} \right] + \\
+ \frac{1}{(\Omega(r) - 1)} \left[ 1 - \sqrt{\Omega(r)} \left( \frac{A(r, t)}{A_0(r)} \right) - (\Omega(r) - 1) \left( \frac{A(r, t)}{A_0(r)} \right)^2 \right].
\tag{3.36}
\]

3. \(\Omega_M(r) \equiv \Omega(r) < 1\) and \(\Omega_{\Lambda}(r) = 0\):

\[
(t - t_0) H_0(r) = \frac{\Omega(r)}{(1 - \Omega(r))^{3/2}} \left[ \arcsinh \sqrt{\frac{1 - \Omega(r)}{\Omega(r)}} - \arcsinh \sqrt{\frac{1 - \Omega(r)}{\Omega(r)}} \left( \frac{A(r, t)}{A_0(r)} \right) \right] + \\
- \frac{1}{(1 - \Omega(r))} \left[ 1 - \sqrt{\Omega(r)} \left( \frac{A(r, t)}{A_0(r)} \right) - (\Omega(r) - 1) \left( \frac{A(r, t)}{A_0(r)} \right)^2 \right],
\tag{3.37}
\]

which can alternatively be written as

\[
(t - t_0) H_0(r) = \frac{\Omega(r)}{(1 - \Omega(r))^{3/2}} \ln \left( \frac{\sqrt{1 - \Omega(r)} + 1}{\sqrt{(1 - \Omega(r)} A(r, t) A_0(r) + \sqrt{(1 - \Omega(r)} A(r, t) A_0(r) + \Omega(r)} \right) + \\
- \frac{1}{(1 - \Omega(r))} \left[ 1 - \sqrt{\Omega(r)} \left( \frac{A(r, t)}{A_0(r)} \right) - (\Omega(r) - 1) \left( \frac{A(r, t)}{A_0(r)} \right)^2 \right].
\tag{3.38}
\]

Eq. (3.38) is more useful in numerical applications, as it behaves well also in the limit \(\Omega(r) \to 0\).

4. \(\Omega_M(r) + \Omega_{\Lambda}(r) = 1\):

\[
A(r, t) = A_0(r) \left[ \cosh \left( \frac{3}{2} H_0(r) \Omega_{\Lambda}^{1/2}(r)(t - t_0) \right) + \frac{\sinh \left( \frac{3}{2} H_0(r) \Omega_{\Lambda}^{1/2}(r)(t - t_0) \right)}{\Omega_{\Lambda}^{1/2}(r)} \right]^{2/3}
\tag{3.39}
\]

21
3.2 Light propagation in the LTB spacetime

To compare the inhomogeneous LTB model with observations of cosmic light, we need to relate the redshift and the energy flux of light with the exact nature of the inhomogeneities. For this, we must study light propagation in the LTB spacetime. We derive the appropriate equations in this section; a more general derivation for an off-center observer can be found in Ref. [45].

Due to the spherical symmetry of the solution, it is clear that light can travel radially, that is, there exist geodesics with \( d\theta = d\phi = 0 \). Moreover, since light always travels along null geodesics, we have \( ds^2 = 0 \). Inserting these conditions into the equation for the line element (3.34), we obtain the constraint equation for radial light rays

\[
\frac{dt}{du} = -\frac{dr}{du} \frac{A'(r,t)}{\sqrt{1 - k(r)}},
\]

where \( u \) is a curve parameter and the minus sign indicates that we are studying radially \textit{incoming} light rays.

Consider two light rays with solutions to Eq. (3.40) given by \( t_1 = t(u) \) and \( t_2 = t(u) + \lambda(u) \). Inserting these into Eq. (3.40), we obtain

\[
\frac{d}{du} t_1 = \frac{dt(u)}{du} = -\frac{dr}{du} \frac{A'(r,t)}{\sqrt{1 - k(r)}}
\]

\[
\frac{d}{du} t_2 = \frac{dt(u)}{du} + \frac{d\lambda(u)}{du} = -\frac{dr}{du} \frac{A'(r,t)}{\sqrt{1 - k(r)}} + \frac{d\lambda(u)}{du}.
\]

On the other hand, we also have

\[
\frac{d}{du} t_2 = -\frac{dr}{du} \frac{A'(r,t) + \lambda(u)}{\sqrt{1 - k(r)}} = -\frac{dr}{du} \frac{A'(r,t) + \dot{A'}(r,t)\lambda(u)}{\sqrt{1 - k(r)}},
\]

where Taylor expansion has been used in the last step and only terms linear in \( \lambda(u) \) have been kept. Equating the right hand sides of Eqs. (3.42) and (3.43) gives the relation

\[
\frac{d\lambda(u)}{du} = -\frac{dr}{du} \frac{\dot{A'}(r,t)\lambda(u)}{\sqrt{1 - k(r)}}.
\]

Differentiating the definition of the redshift, \( z \equiv (\lambda(0) - \lambda(u))/\lambda(u) \), we obtain

\[
\frac{dz}{du} = -\frac{d\lambda(u)}{du} \frac{\lambda(0)}{\lambda^2(u)} = \frac{dr}{du} \frac{(1 + z)\dot{A'}(r,t)}{\sqrt{1 - k(r)}},
\]

where in the last step we have used Eq. (3.44) and the definition of the redshift. Finally, we can combine Eqs. (3.29), (3.40) and (3.45) to obtain the pair of differential equations

\[
\frac{dt}{dz} = \frac{-A'(r,t)}{(1 + z)A'(r,t)}
\]

\[
\frac{dr}{dz} = \frac{\sqrt{1 + H_0^2(r)(1 - \Omega_M(r) - \Omega_L(r))A_0^2(r)}}{(1 + z)A'(r,t)},
\]
determining the relations between the coordinates and the observable redshift: $t(z)$ and $r(z)$. Now that we have related the redshift to the inhomogeneities, we still need the relation of the redshift to other, independent observables such as the energy flux $F$ or the angular size $\Theta$ of the object seen on the sky. A widespread convention in astronomy is that, instead of considering $F$ and $\Theta$ directly, one defines two observable distance measures: the luminosity distance

$$d_L \equiv \sqrt{L/4\pi F},$$

where $L$ is the total power radiated by the source, and the angular diameter distance

$$d_A \equiv \frac{s}{\Theta},$$

where $s$ is the proper size of the object and $\Theta$ its observed angular diameter on the sky.

There is a general relation between the angular diameter and luminosity distances,

$$d_L(z) = (1 + z)^2 d_A(z),$$

which holds for geodesic light in any spacetime as proved by Etherington in 1933 [46,47]. We may thus sometimes speak simply about an observable distance-redshift relation $d(z)$ which can refer to both $d_A(z)$ and $d_L(z)$.

In the LTB model, the angular diameter distance is related to the metric simply by [10]:

$$d_A(z) = A(r(z), t(z)),$$

so due to Etherington’s theorem (3.50), we also have

$$d_L(z) = (1 + z)^2 A(r(z), t(z)).$$

As the relations $t(z)$ and $r(z)$ are determined by Eqs. (3.46) and (3.47) and the scale function $A(r, t)$ by Eq. (3.33), using Eqs. (3.51) and (3.52), we can calculate the observables $d_A$ and $d_L$ for a given $z$. All of these relations have a manifest dependence on the inhomogeneities, i.e. on the functions $H_0(r)$ and $\Omega_M(r)$. What remains is a comparison of Eqs. (3.51) and (3.52) with the observed $d_A(z)$ and $d_L(z)$. A comparison of the observed luminosity distances from type Ia supernova with the LTB function $d_L(z)$ for various inhomogeneity profiles has been done in Paper 1.

In the FRW model, the parameters that best describe our universe are found by maximizing the likelihood function $e^{-\chi^2(H_0, \theta_0, \Omega_0)}$ constructed from the observations. However, to find the boundary conditions of the LTB model that best describe our universe, we should in principle maximize the likelihood functional $e^{-\chi^2(H_0(r), \Omega_M(r))}$. In practice, this is impossible. Therefore we have considered some physically motivated types for the functions $H_0(r)$ and $\Omega_M(r)$ that contain free parameters; these have then been fitted to the supernova observations by maximizing the leftover likelihood function.

### 3.3 Classifying inhomogeneities in the LTB solution

There are different ways to classify the inhomogeneities of the LTB spacetime. Firstly, they can be classified according to their physical interpretation on a spatial hypersurface at $t = t_0$: inhomogeneities in the matter distribution $\rho_M(r, t_0)$ and inhomogeneities in the expansion rate $\theta(r, t_0)$.
or, equivalently, in the velocity distribution of the matter. Secondly, we can consider differences in the time evolution, i.e. divide the inhomogeneities to growing and decaying modes. Although our quantitative discussion concentrates only on the spherically symmetric LTB solution, the same division applies to arbitrary inhomogeneities.

The growing modes can be found by demanding homogeneity in the early universe or simultaneity of the Big Bang (BB). In a simultaneous Big Bang, the divergent curvature of the initial singularity or the condition \( A(r, t_{BB}) = 0 \) coexists everywhere at the same time coordinate \( t = t_{BB} \) whereas in an inhomogeneous Big Bang, \( t_{BB}(r) \) is a function of the spatial location. Thus, for a simultaneous Big Bang, the age of the universe must be independent of the spatial coordinates: \( t_{age} = t_0 - t_{BB} = \text{constant} \). To obtain this constraint we need the expression for the age of the LTB universe which can be calculated by performing the integral in Eq. (3.33) with \( A(r, t) = 0 \). In the cases where the result is an elementary function of the boundary conditions, the results are:

1. \( \Omega_M(r) < 1 \) and \( \Omega_A(r) = 0 \):

\[
t_{age}(r) = \frac{\sqrt{1 - \Omega_M(r)} - \Omega_M(r) \arcsinh \sqrt{\frac{1 - \Omega_M(r)}{\Omega_M(r)}}}{H_0(r)(1 - \Omega_M(r))^{3/2}}. \tag{3.53}
\]

2. \( \Omega_M(r) > 1 \) and \( \Omega_A(r) = 0 \):

\[
t_{age}(r) = \frac{\Omega_M(r) \arcsin \sqrt{\frac{\Omega_M(r) - 1}{\Omega_M(r)}} - \sqrt{\Omega_M(r) - 1}}{H_0(r)(\Omega_M(r) - 1)^{3/2}}. \tag{3.54}
\]

3. \( \Omega_M(r) = 1 \) and \( \Omega_A(r) = 0 \):

\[
t_{age}(r) = \frac{2}{3H_0(r)}. \tag{3.55}
\]

4. \( \Omega_M(r) + \Omega_A(r) = 1 \):

\[
t_{age}(r) = \frac{2}{3H_0(r)} \arcsinh \sqrt{\frac{1 - \Omega_M(r)}{\Omega_M(r)}}. \tag{3.56}
\]

In each case, the requirement \( t_{age}(r) = \text{constant} \) sets the constraint between the boundary condition functions \( \Omega_M(r) \) and \( H_0(r) \) that preserves only the growing modes. If the condition \( t_{age}(r) = \text{constant} \) is instead violated, decaying modes are present as well. In the case that \( \Omega_M(r) = 1 \) or \( k(r) = 0 \), we have only decaying modes. For a more detailed discussion on growing and decaying modes in the LTB model, see [48].

### 3.4 Useful analytic formulae of the LTB solution

Let us next list some previously unpublished expressions for the LTB solution that we have derived and found useful in our research, particularly in numerical computations. Since the CMB observations, discussed in Sect. 4.1.1, suggest the early universe was very close to homogeneous, we consider here only LTB models with simultaneous Big Bang. Furthermore, as our main interest is in voids, we restrict the considerations only to the case where \( \Omega(r) \leq 1 \). We also
assume \( \rho_A = 0 \). These requirements imply that the age of the universe is given by Eq. (3.53) and must be independent of \( r \), yielding the following constraint between the boundary condition functions \( \Omega(r) \) and \( H_0(r) \):

\[
H_0(r) = \frac{1}{t_0} \left[ \sqrt{1 - \Omega(r)} - \frac{\Omega(r) \text{arsinh} \sqrt{1 - \Omega(r)}}{(1 - \Omega(r))^{3/2}} \right],
\]

where \( t_0 = \text{constant} \) is the age of the universe. Thus, the model under consideration can be determined by a free dimensionless function, \( \Omega(r) \), and a free parameter, \( t_0 \). We give the expressions for the scale function and its most relevant derivatives in Eqs. (3.58)-(3.62).

Firstly, the integral (3.33) yields an algebraic implicit equation that determines the scale function:

\[
\left( \frac{t}{t_0} \right) \left[ \Omega \text{arsinh} \sqrt{\frac{1 - \Omega}{\Omega} - \sqrt{1 - \Omega}} \right] = \\
\left[ \Omega \text{arsinh} \sqrt{a \left( \frac{1 - \Omega}{\Omega} \right) - \sqrt{1 - \Omega} \sqrt{\Omega a + (1 - \Omega)a^2}} \right],
\]

where \( a \equiv a(r, t) \equiv A(r, t)/A_0(r) \) and \( \Omega \equiv \Omega(r) \). Taking the derivative \( \partial / \partial r \) of Eq. (3.58) and solving for \( a'(r, t) \) gives

\[
a'(r, t) = \frac{\Omega'}{\Omega(1 - \Omega)} \left[ a - \left( \frac{t}{t_0} \right) \sqrt{\frac{\Omega}{a} + (1 - \Omega)} \right],
\]

where \( a(r, t) \) must be solved from Eq. (3.58). By differentiating Eq. (3.59) w.r.t. time and using Eq. (3.31) to eliminate \( \dot{a} \), we have

\[
\dot{a}'(r, t) = \frac{\Omega'}{(1 - \Omega)^{3/2}} \frac{1}{t_0} \left[ \sqrt{\frac{\Omega}{a} + (1 - \Omega)} \left\{ \sqrt{1 - \Omega} - \text{arsinh} \sqrt{\frac{1 - \Omega}{\Omega}} \right\} \right] + \\
+ \frac{1}{2a^2} \left( \frac{t}{t_0} \right) \left\{ \sqrt{1 - \Omega} - \Omega \text{arsinh} \sqrt{\frac{1 - \Omega}{\Omega}} \right\}. \]

Similarly, we obtain the higher order derivatives:

\[
a''(r, t) = \frac{a}{\Omega(1 - \Omega)} \left[ \Omega'' + \frac{2\Omega'^2}{1 - \Omega} \right] - \left( \frac{t}{t_0} \right) \left[ \Omega'' + \frac{3}{2} \frac{\Omega'^2}{1 - \Omega} \right] \times \\
\sqrt{\frac{\Omega}{a} + (1 - \Omega)} \left( \frac{\Omega^2}{2\Omega(1 - \Omega)^2} \right) \frac{1}{a^2} \left( \frac{t}{t_0} \right)^2,
\]

\[
\ddot{a}'(r, t) = \frac{\sqrt{\Omega/a + (1 - \Omega)}}{\Omega(1 - \Omega)} \left\{ \frac{\sqrt{1 - \Omega} - \Omega \text{arsinh} \sqrt{1 - \Omega}}{(1 - \Omega)^{3/2}} \right\} \times \\
\left( \Omega'' + \frac{\Omega'^2}{1 - \Omega} \left[ 2 + \frac{1}{a^3} \left( \frac{t}{t_0} \right)^2 \right] \right) - \Omega'' - \frac{3}{2} \frac{\Omega'^2}{1 - \Omega} \times \\
+ \frac{1}{t_0} \left( \frac{t/t_0}{a^2(1 - \Omega)} \right) \times 
\]
\[
\left\{ \frac{1}{2} \left( \sqrt{1 - \Omega - \Omega \sinh \sqrt{\frac{1 - \Omega}{1 + \Omega}}} \right) \left( \frac{\Omega'' + 3 \Omega' - \Omega' - \frac{\Omega'^2}{\Omega(1 - \Omega)}}{2} \right) \right\}.
\]

When \( \Omega \to 1 \), the expressions (3.58)–(3.62) take the unstable form "0/0" and are hence useless for numerical computations near the points where \( \Omega = 1 \). Since the scale function \( a(r, t) \) and its derivatives nevertheless behave well if the limit \( \Omega \to 1 \) is taken properly, the numerical difficulties can be avoided by considering the series expansions of Eqs. (3.58)–(3.62) in powers of \( \omega \equiv 1 - \Omega \). We give these expansions in Appendix A.

### 3.5 Bubblessence

When studying inhomogeneous cosmological models, it is instructive to find an effective scalar field that reproduces the observable distance-redshift relations within the FRW description. For this, we consider a FRW model with a classical free field coupled only to gravity. The collective homogeneous fluid formed by matter, spatial curvature and the auxiliary field is assumed to satisfy the instantaneous equation of state

\[
p = w(z)\rho ,
\]

valid for all redshifts, where the function \( w(z) \) defines the effective equation of state for the inhomogeneities. We name the effective field Bubblessence according to the cosmic voids or bubbles that are perhaps the most important type of inhomogeneity having an effect on the observations; see the discussion in Chapter 5.

In order for the effective FRW model to produce the same observable distance-redshift relations as the inhomogeneous LTB spacetime under consideration, we demand the equality:

\[
d_{A}^{\text{FRW}}(z) = d_{A}^{\text{LTB}}(z).
\]

In the homogeneous FRW model, with the collective fluid satisfying Eq. (3.63), the angular diameter distance can be calculated from Eqs. (3.50) and (4.12). Written in terms of \( w(z) \), the result is:

\[
d_{A}^{\text{FRW}}(z) = \frac{H_{0}^{-1}(1 + z)}{1 - \Omega_{0}} S_{k}\left\{ 1 - \Omega_{0}\int_{0}^{z} \exp \left( -\frac{3}{2} \int_{0}^{x} \frac{1 + w(u)}{1 + u} du \right) dx \right\},
\]

where \( S_{k}(x) \) is defined as

\[
S_{k}(x) = \begin{cases} 
\sin(x) & \text{if } \Omega_{0} < 1 \\
x & \text{if } \Omega_{0} = 1 \\
\sin(x) & \text{if } \Omega_{0} > 1
\end{cases}
\]

\( H_0 \) is the Hubble constant and \( \Omega_0 \equiv \rho_0/(3H_0^2/8\pi G) \) the total present-day energy density of the FRW model. The angular diameter distance of the LTB model is given by Eq. (3.51) which, together with Eqs. (3.64) and (3.65), implies the following expression for the equation of state function of the Bubblessence field:

\[
w(z) = -\frac{2}{3}(1 + z) \frac{\partial^2 A}{\partial z^2} [1 + z)A] + \frac{2}{3} (1 - \Omega_0) A(1 + z)^2 A(1 - \Omega_0)^2(1 - \Omega_0) - 1 .
\]
For simplicity, we take a flat FRW model as the effective solution, so \( \Omega_0 = 1 \) and Eq. (3.67) reduces to
\[
 w(z) = -\frac{2}{3}(1 + z) \frac{\partial^2}{\partial z^2}[(1 + z)A(r(z), t(z))] - 1 .
\] (3.68)

For the computations we rewrite Eq. (3.68) as
\[
 w(z) = -\frac{2}{3} \left[ \frac{f(z) + (1 + z)\partial_z f(z)}{f(z) + A(r(z), t(z))} \right] - 1 ,
\] (3.69)

where \( f(z) \equiv (1 + z)\partial_z A(r(z), t(z)) \). Writing \( \partial_z \) in terms of \( \partial_r \) and \( \partial_t \) by the chain rule, \( f(z) \) reads as:
\[
 f(z) = \frac{A'(r, t)}{A'(r, t)} \left[ \sqrt{1 - k(r)} - A(r, t) \right] \left. \right|_{(r, t) = (r(z), t(z))} .
\] (3.70)

Similarly, the first derivative of \( f(z) \) takes the form
\[
 \frac{\partial f(z)}{\partial z} = \frac{1}{1 + z} \left[ \left( \frac{A''}{A'^2} - \frac{A' A''}{A^3} \right) (1 - k) + \left( \frac{\dot{A}' A'^2}{A'^3} + \frac{\ddot{A} A' A'}{A'^2} - \frac{A'' A'}{A'^2} - \frac{2 A'}{A'^3} \right) \sqrt{1 - k} + \right.
\]
\[
\frac{A' \dot{A}}{A'} + \frac{A'^2 \ddot{A}}{2 A'^2} - \frac{A' A'' \dot{A}}{A'^3} \left. \right|_{(r, t) = (r(z), t(z))} .
\] (3.71)

Note that Etherington’s relation (3.50) holds true for both the FRW and the LTB solutions, the condition (3.64) leads to the equivalence of the different distance measures (3.48) and (3.49) between the physical LTB spacetime and the effective FRW model, though only for an observer sitting at the LTB symmetry center. Furthermore, as the effect of a local inhomogeneity on the CMB radiation comes mainly via the change it induces to the angular diameter distance, the use of Eq. (3.69) offers a possibility to fit an inhomogeneous LTB model with a local void to the observed CMB power spectrum, discussed in Sect. 4.1.1.
Chapter 4

Cosmological observations and the standard FRW model

Physical cosmology studies the basic laws of the universe from the largest scales all the way down to the Planck length or, if possible, even below. The cosmological theories are tested mainly against astrophysical observations, but also against data that is collected from high energy accelerators on Earth, such as the brand-new 7 TeV Large Hadron Collider at Cern [49].

In this chapter, we outline the so-called standard cosmology as well as briefly review the observations which have led to the unexpected conclusion that the expansion of the late universe has accelerated. The three main assumptions of the standard cosmology are: 1) standard general relativity, 2) homogeneous and isotropic FRW metric and 3) small Gaussian density fluctuations in the energy distribution of the early universe. Rather than a scientific consensus, the standard model of cosmology is still more like a working hypothesis – a useful set of assumptions which yields models that are simple enough to be tested against the cosmological observations.

4.1 Cosmological observations

The three important observations that can be simultaneously accounted for only by assuming accelerated cosmic expansion along our line of sight are: the cosmic microwave background radiation, the large scale galaxy distribution and the type Ia supernovae. There are also various other cosmological data, such as gravitational lensing statistics [50], the Hubble Key Project [51, 52] and the Big Bang nucleosynthesis [53], that set valuable constraints on the cosmological models but do not alone provide evidence for the acceleration of the expansion. For a thorough and open-minded review of the current observational status, see [54].

4.1.1 Cosmic microwave background

The detection of the 3 K microwave radiation that bathes the Earth from all directions was reported in 1965 by Penzias and Wilson [55]. Evidently, the cosmic microwave background is a remnant of the Big Bang: it is believed to be the photon gas from the early hot universe, cooled down by the enormous factor the universe has expanded since the early stages. The frequency distribution of the CMB is closest to the black body spectrum ever observed. By measuring the
temperature and the small anisotropy of the CMB radiation over the sky, we can gain knowledge about the expansion history of the universe.

The basic measurement from the CMB is its temperature distribution over the sky,

$$ T(\theta, \phi) = \bar{T} + \delta T(\theta, \phi) , $$

where $\bar{T} = 2.725 \pm 0.001$ is the observed average temperature and the function $\delta T(\theta, \phi)$ represents the small anisotropy. Conventionally, the observed anisotropy is expanded in terms of spherical harmonics $Y_{lm}(\theta, \phi)$, a complete set of functions on a sphere [56],

$$ \delta T(\theta, \phi) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \phi) . $$

![Image of temperature anisotropy map](image)

Figure 4.1: The temperature anisotropy map of the CMB radiation over the sky from WMAP 5-year data. Figure by WMAP Science Team [57].

A standard assumption is that the statistical properties of the temperature distribution do not depend on the orientation of the $(\theta, \phi)$ coordinates. One can thus average over the index $m$ to obtain the observed temperature power spectrum:

$$ C_l = \sum_{m=-l}^{l} |a_{lm}|^2 . $$

By assuming the small anisotropy to be a (Gaussian) random variable, one then compares the observed spectrum (4.3) with the theoretical one,

$$ \tilde{C}_l = \langle |\tilde{a}_{lm}(P_i)|^2 \rangle , $$

29
where $\langle \cdot \rangle$ stands for an ensemble average and $P_i$ represents the cosmological parameters of the model; a comparison between the observed and a theoretical power spectra is plotted in Fig. 4.2. By iterating the comparison with different values of the parameters $P_i$, one can find a best fit where the theoretical spectrum $C_l$ and the observed one $C_l$ deviate from each other as little as possible.

![CMB power spectrum comparison](image)

**Figure 4.2:** The theoretical CMB temperature power spectrum of the best fit accelerating $\Lambda$CDM model with binned data points from different experiments. Figure by WMAP Science Team [58].

The most prominent feature in the CMB is the mode $l = 1$, that is the dipole anisotropy with the observed maximum value $\Delta T/T = 1.23 \cdot 10^{-3}$. It corresponds to a pattern which is hot in one direction and cold in the opposite direction, with a smooth transition between them. CMB maps are usually shown with the dipole removed, such as the map from the WMAP 5-year data in Fig. 4.1. The origin of the dipole can be both kinematical, i.e. due to the motion of the Earth relative to the CMB rest frame, and cosmological, i.e. due to inhomogeneities of the universe. Assuming our motion were alone responsible for the dipole, it would correspond to Sun having a velocity of 370 km/s [59]. Alternatively, a potential cosmological origin is in voids, as elaborated in Chapter 5.

We have only a single temperature power spectrum but a set of parameters $P_i$, typically of the order of ten, to constrain from the data. A rule of thumb is that one can constrain as many non-degenerate parameters as there are features in the spectrum. However, due to parameter degeneracies, it is not feasible to extract specific values for most of the parameters from the CMB data alone. Moreover, the best-fit estimates for the parameters often depend heavily on the underlying assumptions or priors of the model. When the CMB is combined with other data sets, however, more stringent constraints can be found.
As an important example of the parameter degeneracies, the current CMB data cannot by itself determine whether the expansion of the late universe has accelerated or not. This is made explicit by a flat matter dominated FRW model, the so-called CHDM model by Hunt and Sarkar [60, 61], which decelerates at all times but yields statistically even a better fit to the CMB data than the accelerating standard ΛCDM model. However, due to e.g. a low Hubble constant $H_0 \simeq 44 \text{ km/s/Mpc}$ in the CHDM model, it appears hard to reconcile the model with other data sets, unless the observable distance-redshift relations are not determined by the FRW metric, e.g. due to inhomogeneities in the expansion rate, discussed in Chapter 5.

In addition to the temperature of the CMB radiation, it is also possible to measure its polarization. The observed polarization sets additional constraints on the cosmological models but, due to the quality of the current data, they are still weak relative to those obtained from the temperature measurements. However, more accurate constraints are expected from the polarization measurements of the upcoming Planck satellite experiment [62].

### 4.1.2 Large-scale galaxy distribution

The largest observable features or inhomogeneities in the universe consist of clusters and filaments of galaxies with nearly empty voids between them [63–66]; a glimpse of this structure can be seen in Fig. 4.3. The galaxy surveys seem to indicate that a homogeneity scale of the universe is at least $>100 \text{ Mpc}$, if such a scale exists at all [67–69]. By studying this structure, we can learn about the total density of matter and how matter clusters. As the clustering of matter depends on the expansion history of the universe, the galaxy surveys can give us information about the potential acceleration of the expansion.

The clustering properties of the galaxies are analyzed statistically by looking at how the density of galaxies is correlated. A quantitative measure for this is the two-point correlation function $\xi(s)$, which is defined as the excess probability, relative to a homogeneous distribution, of finding a galaxy at a distance $s$ from another galaxy. By writing the deviation from a homogeneous matter distribution as $\rho(x) = \bar{\rho}(1 + \delta(x))$, where $\bar{\rho}$ is the background FRW density and $\delta(x)$ is the perturbed density contrast, the two-point correlation function reads as

$$
\xi(s) = \langle \delta(x)\delta(x) \rangle ,
$$

where the averaging is taken over all realizations of the density field under the condition $s = |x - \bar{x}|$. The two-point correlation function from a recent galaxy survey is plotted in Fig. 4.4, showing the recently detected baryon acoustic oscillation peak in the SDSS data [5]. In practice, one often deals with the matter power spectrum $P(k)$, which is the Fourier transform of $\xi(s)$:

$$
P(k) = \int \xi(x)e^{-ik\cdot x} \, d^3x .
$$

An important piece of information about the expansion history of the universe is contained in the observable distance $d(0.35)$ to the redshift $z = 0.35$, corresponding to the baryon acoustic oscillation scale. As the physical scale associated with these oscillations is set by the sound horizon at recombination, the baryon oscillation peak in the galaxy spectrum and the first peak of the CMB angular power spectrum should correspond to the same physical feature\(^1\) just seen

---

1 Assuming the statistical properties of the primordial perturbations are the same at different spatial locations in the universe.

31
at different redshifts $z = 0.35$ and $z = 1100$, i.e. at different epochs and spatial locations. By measuring how large this length scale appears in the galaxy distribution and combining it with the information of the associated scale of the peak in the CMB spectrum, we obtain a constraint on the ratio $d(0.35)/d(1100)$. The observational value for this ratio provides independent evidence for the acceleration of the universe [5]. Recently, the galaxy surveys have been able to determine also the best fit value for the distance $d(0.2)$. Interestingly, when combined with the supernova observations, there seems to be tension between the deduced value for the ratio $d(0.35)/d(0.2)$ and the accelerating $\Lambda$CDM models [71].

4.1.3 Type Ia supernovae

Supernovae are extremely violent nuclear explosions of stars whose brightness during the explosion can exceed even the host galaxy formed by some $10^{10}$ stars. The so-called type Ia supernovae
Figure 4.4: The two-point correlation function from the SDSS data. The curve with no peak represents a matter dominated model with no acceleration; other curves correspond to accelerating models with different total matter densities. Figure from Ref. [5].

have turned out to be good standard candles, i.e. objects with nearly constant intrinsic luminosity. More specifically, they should be called standardizable, since it is crucial to use their light curves to reduce the scatter between the different type Ia supernovae [72, 73].

The nearly constant intrinsic luminosity makes type Ia supernovae useful objects to measure the relation between redshifts and distances in the universe. Light emitted at the explosion will be redshifted along its path towards the observer due to the expansion of the universe. Furthermore, the light will appear dimmer due to the distance to the supernova, with the exact amount of dimming depending on the expansion of the universe along our line of sight to the object. Thus, by measuring the redshifts and energy fluxes from supernovae at different distances from us, we can constrain the expansion of the universe from today back until the time corresponding to the emission of light from the farthest supernova.

According to the observations reported since 1998, the type Ia supernovae at large redshifts appear fainter than expected in a perfectly homogeneous and isotropic matter dominated FRW universe [74]. The natural interpretation for this so-called excess dimming of the distant supernovae seems to be that the expansion of the universe has accelerated during the latter half of its age. Of the different observations, the supernovae provide the most direct evidence for the acceleration, as they probe the integrated cosmic expansion to several redshifts in the range $z = 0...2$. The fact that the present supernova data disfavors models with no acceleration is illustrated in Fig. 4.5.
Figure 4.5: The type Ia supernova data seems to support accelerated expansion at a high confidence level. Figure from Ref. [75].
4.1.4 The unforeseen message: accelerated expansion along our line of sight

Due to the finiteness of the speed of light, our line of sight meets the universe along the past light cone. Since the cosmological data supporting the acceleration is based on observations of light, the acceleration has been detected only along our past light cone, which does not imply the stronger condition of local acceleration of the expansion.

When the CMB data is combined with the BAO scale inferred from the galaxy distribution statistics, we obtain a constraint that the expansion of the universe should have increased along our line of sight such that \( Ht \sim 2/3 \) at large redshifts whereas today \( H_0 t_0 \sim 1 \). Independently from these two, a similar conclusion comes from the type Ia supernova data. Thus, when these three data sets are combined, a consistent picture seems to arise: during the latter half of the universe age, the expansion appears to have increased roughly by a factor of \( 3/2 \) along our line of sight. As we explain in Sect. 4.3, under the assumption of the homogeneous and isotropic FRW metric, this implies that an exotic fluid with large negative pressure, called dark energy, contains about 75\% of the energy of the universe.

4.2 Spatially homogeneous and isotropic FRW models

The simplest non-trivial models for the large scale structure of the universe are obtained by assuming perfect spatial homogeneity and isotropy but allowing the universe to evolve in time. The metric in these models is a special case of the LTB metric (3.34), with \( A(r,t) = a(t)r \) and \( k(r) = kr^2 \):

\[
ds^2 = -dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right]. \tag{4.7}
\]

The Einstein equations for the metric (4.7) reduce to the Friedman equation:

\[
H^2(t) \equiv \frac{\dot{a}^2(t)}{a^2(t)} = \frac{8\pi G}{3} \rho - \frac{k}{a^2}, \tag{4.8}
\]

where \( \rho \) is the total energy density, and the acceleration equation:

\[
\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \tag{4.9}
\]

where \( p \) stands for pressure. We have also the continuity equation,

\[
\dot{\rho} + 3H(\rho + p) = 0, \tag{4.10}
\]

which, however, is not independent of the first two equations (4.8) and (4.9). By using Eq. (4.10) to solve the time evolution of the different energy components, the Friedman equation (4.8) can be written as:

\[
H^2(t) = H_0^2 \left[ \Omega_M \left( \frac{a_0}{a} \right)^3 + \Omega_R \left( \frac{a_0}{a} \right)^4 + \Omega_\Lambda + (1 - \Omega_M - \Omega_R - \Omega_\Lambda) \left( \frac{a_0}{a} \right)^2 \right], \tag{4.11}
\]

where \( a_0 \equiv a(t_0) \), while \( \Omega_M, \Omega_R \) and \( \Omega_\Lambda \) represent the present-day energy densities of the different components - matter, radiation and dark energy - in the units of critical density.

As the FRW spacetime is only a special case of the more general LTB spacetime of Chapter 3, the equations for the light propagation can be obtained from the LTB equations (3.46) and
(3.47) in Sect. 3.2. The resulting luminosity-distance redshift relation can in this case be written as the integral:

\[
d_L^{FRW}(z) = (1 + z) \frac{H_0^{-1}}{\sqrt{|1 - \Omega_0|}} S_k \left[ \sqrt{|1 - \Omega_0|} H_0 \int_0^z \frac{d\tilde{z}}{H(\tilde{z})} \right],
\]

(4.12)

where \( S_k(x) \) is given by Eq. (3.66).

Since the expansion integrated over the past light cone, \( \int_0^z H^{-1}(\tilde{z})d\tilde{z} \), yields the main contribution to the relation (4.12), one sometimes loosely speaks about the expansion \( H(z) \) and the distance measure \( d(z) \) interchangeably.

A convenient measure of the expansion rate is given by the dimensionless factor obtained by multiplying the Hubble parameter \( H \) with the age of the universe \( t \). By integrating the Friedman equation (4.11), we find the following expressions:

\[
H t = \begin{cases} 
1 & \text{if } \Omega_m = \Omega_v = 0 \\
2/3 & \text{if } \Omega_m = \Omega_v = 1 \\
2/sqrt{\Omega_\Lambda} \arcsinh \sqrt{\Omega_\Lambda} & \text{if } \Omega_\Lambda + \Omega_m = 1
\end{cases}
\]

(4.13)

To avoid a common confusion, we note that although space with \( \Omega_0 \simeq 0 \) expands faster \( (Ht = 1) \) than space with \( \Omega_0 = 1 \) \( (Ht = 2/3) \), in the \textit{mathematical limit} of perfectly empty space \( \Omega_0 \equiv 0 \), the expansion vanishes. This may seem unphysical, but the contrary is the case: some matter or radiation is needed to make the expansion observable via its effect on that very substance.

That is, the concept of expansion is meaningful only as long as there is something it can have observable consequences on. In the real universe, at least the ubiquitous CMB radiation ensures that the expansion remains detectable and thus physical.

### 4.3 Dark energy and dark matter

Conventionally, the cosmic acceleration or the increase in the expansion rate from \( H t \simeq 2/3 \) at \( z \sim 10^3 \) to \( H t_0 \simeq 1 \) at \( z \sim 0 \) has been accounted for by introducing a cosmological constant \( \Lambda \) or vacuum energy \( \Omega_\Lambda \). Indeed, the analysis of the cosmological data within the FRW model with both the vacuum energy and matter components included yields a consistent picture of the universe, known as the concordance \( \Lambda \)CDM model, with the following best fit values for the density parameters:

- CMB alone: \( \Omega_m + \Omega_\Lambda \sim 1 \)
- Galaxy surveys alone: \( \Omega_m \sim 0.3 \)
- CMB + BAO scale from galaxy surveys: \( \Omega_\Lambda \sim 0.7 \)
- Type Ia supernovae\(^2\): \( \Omega_m \sim 0.3 \) and \( \Omega_\Lambda \sim 0.7 \)

\(^2\)In the supernova data analysis spatial flatness \( (\equiv \Omega_m + \Omega_\Lambda = 1) \) has been assumed. However, as shown in Fig. 4.5, these values are within the 1\( \sigma \) error of the best fit values without the flatness constraint: \( \Omega_m \sim 0.5 \) and \( \Omega_\Lambda \sim 1 \).

36
Moreover, at least the supernova data seems to require $\Omega_\Lambda > 0$ at a high confidence level [3, 75]. Non-trivial dynamics of the vacuum energy is permitted, but not required by the data. Generically, in cosmology the vacuum energy, whether constant or not, is known as dark energy. Besides dark energy, astrophysical observations suggest the existence of another new ingredient: dark matter [76]. Dark matter is assumed to be simply ordinary kind of matter that is just not seen and is probably made of (yet unknown) weakly interacting particles. Unlike dark energy, dark matter is not theoretically problematic; in fact, a generic expectation from the completion of the particle physics theories is new particles. Observationally, there are several independent sources of evidence for dark matter, such as the motions of stars in galaxies, the motions of clusters, the peak structure of the CMB power spectrum, the early formation of structures, gravitational lensing, as well as direct measurements of the matter density combined with the baryon density from the Big Bang Nucleosynthesis. In contrast, the explanation of the observations without dark energy requires only the relatively small ($\sim 3/2$) change in the expansion rate along our past light cone.

### 4.3.1 Problems with dark energy driven acceleration

Although homogeneous and isotropic models with dark energy can provide a phenomenological fit to the cosmological data, there are various suspicious issues that motivate the considerations of alternative solutions. We formulate these issues into five questions that we feel one should be able to give a natural answer in order to consider the explanation to the perceived acceleration plausible.

1. If nonzero, why would the cosmological constant be so tiny, $\Lambda \sim 10^{-123} \text{ G}^{-1}$, whereas naive dimensional analysis suggests $\Lambda \sim G^{-1}$?

2. Why would dark energy appear at a very low temperature $T \simeq 4 \text{ K}$ while it has had the huge temperature range from $T_{\text{Planck}} \gtrsim 10^{32} \text{ K}$ to $T_0 \simeq 2.7 \text{ K}$ available? A naive expectation for the probability of a quantum field to arise at a temperature $T$ is $P \sim e^{-T_{\text{Planck}}/T}$.

3. Why would $\Lambda$ be fine-tuned such that today the universe undergoes nearly free expansion, $H_0 t_0 = \frac{2}{3\sqrt{\Omega_\Lambda}} \text{arsinh} \sqrt{\frac{\Omega_\Lambda}{1 - \Omega_\Lambda}} \simeq 1$, whereas a slightly bigger $\Lambda$ would yield $H_0 t_0 > 1$?

4. Why would the effects of dark energy appear just when large voids start to form?

5. What justifies the use of the FRW metric to model light propagation in the real universe?

Despite extensive efforts, various alternatives to dark energy that introduce new theories, such as modifications of general relativity [77], have fallen short of satisfactory answers [6–8]. Instead, as supported by our studies, inhomogeneous cosmological models seem to offer a natural solution to these problems within standard theories. Thus, in the rest of the thesis, the discussion is devoted to the inhomogeneous generalizations of the FRW cosmology.
Chapter 5

Inhomogeneous generalizations of the FRW cosmology

The spatially homogeneous and isotropic FRW spacetime, with the growth of structure described as linear perturbations evolving on the smooth background, has become the widely adopted framework of modern cosmology. Accordingly, the standard model of cosmology is built on the assumption that the effects of the evident nonlinear inhomogeneities on the detectable light average out\(^1\) over cosmological distances. This assumption — although perhaps in concordance with Newtonian intuition — lacks a convincing demonstration within general relativity, not to mention a mathematically rigorous proof. In fact, the assumption was criticized already in the 60’s by Zel’dovich \cite{78}, Feynman\(^2\), Bertotti \cite{79}, Gunn \cite{80}, Kantowski \cite{81} and in the 70’s by Dyer and Roeder \cite{82,83}, who suggested that the clumping of matter in the real universe has consequences on the observable distance-redshift relations that the simplified FRW description fails to capture.

At the time of this pioneering research, the cosmological observations were too inaccurate to distinguish the predictions of the proposed inhomogeneous models from the simple and elegant FRW solutions. Consequently, the efforts towards a more thorough description of the effects of the nonlinear structures on the distance-redshift relations never became part of the mainstream research. Although the structure formation itself has since then developed into a major part of modern cosmology, the emphasis is usually given either to the general relativistic but perturbative description of structures on large scales, or to the Newtonian description of the small scale structure e.g. in the context of galaxy formation.

However, the unparalleled precision of the recent cosmological observations has made the debate about the effect of the nonlinear structures on the distance-redshift relations topical. Indeed, the observations indicate the existence of inhomogeneities that seem to be on large enough scales to have cosmological significance but too strong to be described within the linear theory — in particular the observed voids and the clustering of galaxies as filaments between the voids \cite{63-66}. The increase in the precision of the data naturally demands a corresponding improvement in the theoretical model; otherwise there is an increased risk of drawing incorrect conclusions from the observations.

\(^1\)Apart from gravitational lenses that are occasionally taken into account as additional corrections.

\(^2\)Apparently, R. P. Feynman was also one of the pioneers of this subject, as \cite{80,81} cite a colloquium by Feynman in 1964.
In fact, the discovery that the standard FRW cosmology needs a mysterious, severely fine-tuned dark energy component in order to account for observations (see Sect. 4.3), suggests that the precision of the data may already have surpassed the precision of the employed theoretical model. This is also backed up by the strange coincidence in the standard cosmology, pointed out by Schwarz [84] and Wetterich [85] and elaborated by Räsänen [31–33]: the effects of the elusive dark energy happen to appear just during the same era when nonlinear structures start to form at cosmologically significant scales. As in general, a dynamical explanation would be preferred.

Whereas the fine-tuning of dark energy only suggests that something fundamental is lacking from the standard picture, the coincidence problem provides evidence for a causal connection between dark energy and structure formation. Indeed, perhaps the most natural explanation would be, that the standard FRW description breaks down when the nonlineairities become dominant in the late-time universe and the need for a tiny cosmological constant is only a manifestation of this breakdown, not evidence for dark energy [31]. In recent years, more and more authors have addressed this issue and substantially increased the general understanding of the subject, but crucial open questions still remain [39]. In this thesis, we address the issue by considering phenomenological models that aim to: 1) clarify the physical interpretation for how the structure formation can actually mimic dark energy and 2) provide a quantitative fit for the observations.

5.1 Accelerated expansion from cosmic inhomogeneities

Formally, the cosmological inhomogeneities pose the following problem: we have a physical system, the universe, with $> 10^{80}$ degrees of freedom\(^3\) but we want a model with $\lesssim 10$ degrees of freedom, represented by the cosmological parameters. For a mathematical description of the universe, we thus need a coarse graining map, a physiomorphism: $10^{80}$ dof $\mapsto$ 10 dof, that preserves the relevant physical structure of the spacetime but removes its mathematical complexity. As our cosmological knowledge is firmly based on observing light, the relevant structure means the correct optical properties of the universe. That is, we want the observable distance-redshift relations to be unaltered by the coarse graining.

A common critique against acceleration from inhomogeneities is that linearly perturbed FRW solutions seem to describe well even highly nonlinear density contrasts, $\delta \rho / \bar{\rho} \sim 10^{30}$, while still having insignificant backreaction [86]. This is because the relevant perturbed gravitational potential can satisfy $\Phi \ll 1$ even for the high density contrasts $\delta \rho / \bar{\rho} \gg 1$. However, this argument overlooks the crucial part of the nonlinear structures, namely inhomogeneities in the expansion rate [32]. Even a relatively small spatial variation in the expansion rate, $\delta \theta / \bar{\theta} \sim \mathcal{O}(0.1)$, can invalidate the linear perturbation theory by rendering the uniformly expanding FRW background inapplicable. Sometimes the assumption $\delta \theta / \bar{\theta} \ll 1$ is hidden in the requirement of small peculiar velocities [87], which restricts the spatial variations in the expansion rate to be small as well [88]. As consolidated by all of our studies, spatial variations in the expansion rate of $\mathcal{O}(0.1)$ can indeed induce accelerated expansion via various mechanisms.

In brief, the inhomogeneities can cause deviations from FRW cosmology due to the facts that: 1) the average evolution of the real inhomogeneous universe differs from the evolution of the corresponding homogeneous FRW universe, and 2) the well-observed light does not reflect the average properties of the universe, i.e. we do not see "average" light. The latter point can be further divided into two subcases: A) our location in space can be special so that global averages

\(^3\)($H_0 G m_p)^{-1} c^3 \sim 10^{30}$ gives a rough estimate for the number of protons in the universe.
fail to describe the local universe and B) there can be selection effects in the cosmological observations. Note that these are all independent mechanisms, as it is in principle possible to have a universe: 1) with perfectly transparent structures but highly nonlinear variation of the expansion rate to produce large backreaction, 2) where both the backreaction and the selection effects are negligible, but we are located in a large local void or 3) with large, completely opaque structures that induce significant selection effects but give negligible backreaction. Of course, in the real universe, all of these could in principle be important simultaneously. We next expound the physics behind all the three mechanisms; in Sect. 5.1.1, we discuss the case 1, whereas Sects. 5.1.2 and 5.1.3 focus on the cases 2A and 2B respectively. There can also be other mechanisms that might not be reducible to any of these; see Refs. [19–24].

5.1.1 Backreaction due to inhomogeneous expansion

In Sect. 2.3, we discussed averaging in general relativity. By considering a general irrotational dust universe, we derived a set of averaged scalar equations, the Buchert equations (2.82)–(2.84), describing the average evolution of an inhomogeneous universe. These equations differ from the conventional FRW equations by the backreaction term and by the dependence on the averaging scale, the focus of Paper 2. An important implication is that averages of inhomogeneous quantities do not evolve in time like the corresponding homogeneous quantities.

From Eq. (2.82) it is obvious that, regardless of decelerating locally everywhere, with large enough backreaction \( Q_D(t) > 4\pi G(\rho_D(t)) \), the average expansion can accelerate without an exotic fluid with negative pressure or a modification of gravity. Intuitively, the global acceleration is possible because the volume can become dominated by fast-expanding regions [33]. This is realized with large enough variance of the expansion rate, if the counterbalancing average shear is not too large. The variance becomes large when contracting \((\theta < 0)\) and expanding \((\theta > 0)\) regions coexist. This is exactly what one would expect in the late universe with structures forming via gravitational collapse [42, 89], so it has been conjectured that the average acceleration could explain the cosmological observations [32].

To clarify the essential point, consider two disconnected decelerating regions \( U \) and \( O \) that have initially the same volumes \( V_i \). Let the region \( U \) represent an underdensity that initially expands at the rate \( \theta_U > 0 \) and let \( O \) represent an overdensity that initially contracts at the rate \( -\theta_i < 0 \). Clearly, the volume-averaged expansion rate is initially zero: \( \langle \theta \rangle_i = \langle V_U \theta_U + V_O (-\theta_i) \rangle / 2V_i = 0 \).

Later, the region \( U \) has expanded to take up a volume \( V_U > V_i \) which is still expanding, \( \theta_U > 0 \), though slower than initially, \( \theta_U < \theta_i \), whereas the region \( O \) has shrunken to a small virialized structure with volume \( V_O \ll V_U \) which is essentially static, \( \theta_O \approx 0 \). In the process, the volume-averaged expansion rate has become positive: \( \langle \theta \rangle = (V_U \theta_U + V_O \theta_O) / (V_U + V_O) \approx \theta_U > 0 \), confirming that the global expansion has accelerated regardless of decelerating locally at each point. A more realistic treatment takes into account the junction conditions between the two regions and, indeed, the emergence of the acceleration has been verified within exact solutions of general relativity [89, 90].

The current studies suggest that while backreaction seem to be significant, perhaps increasing the average expansion by a factor \( \sim 1.3 \), it may not be enough to account for the total increase factor 1.5 alone [19]; however, it has been argued that the approximations in the current calculations may underestimate the backreaction [33], though one could also argue that neglecting the shear would do the opposite. A more realistic estimate for the backreaction may require a large
5.1.2 Special location in space: a local void

Clearly, for an observer sitting in a special location, the spatial averages over the whole system may fail to describe local physics around the observer. For example, the average surface temperature of the earth – although useful in studying global changes in the climate – does not help to predict the local weather in wintry Rovaniemi. Similarly, if our location in the universe were special, the global spatial averages could fail to describe our cosmological observations.

When looking at the large-scale structure of the universe, perhaps the most eye-catching feature is that the major part of the universe appears to be taken up by voids of size $O(10) \ldots O(100)$ Mpc [63–66], nearly empty regions expanding faster than the whole universe on average. Even larger voids have been detected [91] and it has recently been suggested that the correlation between the cold spot in the CMB and a deficit of radiogalaxies in the same direction might be evidence for a void of size $\sim 300$ Mpc [92]; though see [93]. It is in principle also possible that we would happen to live inside such a void [94, 95, 97], often known as Hubble Bubble in this context.

A particularly useful metric for describing voids is the spherically symmetric LTB metric$^4$ (3.34). In 1997, Mustapha, Hellaby and Ellis argued that any isotropic set of observations can be explained by appropriate inhomogeneities in the LTB model [98]. Two years later – after the first indications of accelerated expansion in the supernova observations – Célérier made use of this to explicitly demonstrate that the supernova data can as well be accounted for by suitable inhomogeneities in the LTB solution without dark energy [99]. In 2005, Alnes, Amarzguioui and Gron showed that the suitable inhomogeneity profile was in fact physically describing a local void [100], whose existence had a few years earlier been speculated by Zehavi et. al. [94] and Tomita [95] and, even before the acceleration was detected, by Moffat and Tatarksi [96]. During the recent years, several authors have studied various observations in the LTB models [11, 101–111].

A key notion in understanding the physical basis of the inhomogeneity induced perception of acceleration is that the cosmological observations are made along our past light cone. A way to illustrate this is to consider the directional derivative along the light cone

$$\frac{d}{dz} = \frac{\partial x^i}{\partial z} \frac{\partial}{\partial x^i} + \frac{\partial t}{\partial z} \frac{\partial}{\partial t} \approx H_0^{-1} \left( \frac{\partial}{\partial r} - \frac{\partial}{\partial t} \right), \quad (5.1)$$

where the approximation in the last step is more accurate for small distances, $r \ll H_0^{-1}$, but the sign is correct even for larger $r$. Observationally, Eq. (5.1) tells us that negative radial variation roughly corresponds to positive time variation and that their relative contributions to the total variation w.r.t. the observable redshift cannot be distinguished. This is natural since by looking at a source, we simultaneously look into the past (i.e. along the negative $t$-axis) and spatially further (i.e. along the positive $r$-axis).

Within the LTB model, the argument can be made more precise by comparing the expression of

$^4$ Another option is to apply an extended averaging method, considered in Paper 2, where global averages are replaced with local ones for each object separately.
the luminosity distance in a matter dominated local void [12]

\[ d_{\text{LTB}}^L(z) = H_0^{-1}(0) \left[ z + \left( \frac{1}{4} - \frac{H_0'(0)}{H_0^2(0)} \right) z^2 + \left( -\frac{1}{8} + \frac{1}{3} \frac{H_0'(0)}{H_0^2(0)} + 2 \frac{H_0'(0)}{H_0^3(0)} - \frac{1}{2} \frac{H_0'(0)}{H_0^3(0)} \right) z^3 + \mathcal{O}(z^4) \right] , \]  

where \( \dot{r} \equiv \partial/\partial r \), with its expression in the homogeneous and flat \( \Lambda \)CDM model:

\[ d_{\text{LCDM}}^L(z) = H_0^{-1} \left[ z + \left( \frac{1}{4} + \frac{3}{4} \Omega_\Lambda \right) z^2 + \left( -\frac{1}{8} - \Omega_\Lambda + \frac{9}{8} \Omega_\Lambda^2 \right) z^3 + \mathcal{O}(z^4) \right] . \]  

The comparison between Eqs. (5.2) and (5.3) confirms that negative radial variation of the expansion rate, \( H_0'(0) < 0 \), is needed to mimic positive cosmological constant, \( \Omega_\Lambda > 0 \), or to induce accelerated expansion along our line of sight. In other words, for the expansion rate to increase towards us along the past light cone, the LTB expansion function \( H_0(r) \) must decrease as \( r \) grows. This is exactly what an observer inside a void would see, and has been shown to account for various cosmological observations without local acceleration of the expansion [11,100,103,108].

However, the explanation of the perceived cosmic acceleration as an effect induced by a local void does not come for free. Most notably, in order to respect the cosmological principle, we should not sit in the dead center of a spherically symmetric void, as in the simplest LTB models. By considering off-center observers, it was shown in [112] that the supernova data does not impose severe restrictions for our location in the void, but perhaps the most relevant constraint comes from the dipole anisotropy of the CMB. Indeed, as shown in [45], for the LTB model to remain consistent with the observed CMB dipole, we should reside within a few percent of the size of the void from its center.

Greater off-center distances are of course allowed if the induced large cosmological dipole would be counterbalanced by the kinematic dipole caused by a suitable peculiar velocity. However, perhaps the amount of fine-tuning needed for this would be nothing else but going out of the frying pan into the fire. In addition, even if the void were not perfectly spherical and even if we were not near the center of the void, the cosmological principle would still be, at least weakly, violated: there are less galaxies in a void and if thrown to a random galaxy in the universe, it is much more probable to end up in an environment with lots of galaxies than in a relatively empty void. Altogether, without a natural dynamical explanation for this issue, the scenario with a large local void – if taken literally – is perhaps not a probable alternative. Nevertheless, an interpretation for these models could be a coarse grained description of more complex inhomogeneities, such as a group of several smaller voids.

### 5.1.3 Selection effects due to opaque structures

Even if we were not located inside a large local void, the observed cosmic network of smaller voids can still significantly affect the observable distance-redshift relations. Firstly, as explained in Sect. 5.1.1, the spatial variations in the expansion rate can increase the volume-averaged global expansion rate. In addition, as we have pointed out in Paper 3, the opacity of the cosmic structures can bias the observations, such that the light we see has traveled mostly in freely expanding voids at late times, and thereby affect the observed distance-redshift relations,
independently from the dynamical backreaction. Although it may be difficult to distinguish the observational consequences of the selection effects and the backreaction, they are in principle distinct mechanisms: one could have large opaque structures that induce selection effects but give negligible backreaction or, on the other hand, perfectly transparent but highly nonlinear structures that produce backreaction but no selection effects. Naturally, in the real universe, both mechanisms may have significant effects on the observations. Mathematically, the difference between the purely dynamical backreaction and the biased light propagation in voids can be seen as the difference between the volume-averaged expansion rate and the average expansion rate along our line of sight.

We justify the selection effects in the observations of cosmic light simply by the fact that the emptier a region in space, the easier it is to see an object through it. The bias can be further categorized roughly as two different cases: 1) an opaque structure on the foreground simply screens the further object behind it, or 2) the further object is seen but, due to foreground contamination, not clearly enough that the object would contribute to the high quality data. The first case is perhaps relatively rare, since only the nearest galaxies, including the Milky Way, cover a notable fraction of the celestial sphere. Instead, the second case may be more relevant as, due to their long distance, the cosmological objects tend to be dim, which means that they can be harder to detect if there are bright foreground objects nearby on the sky. In addition to the brightness of the foreground, the contamination can arise from gravitational lensing. An important implication of the contamination is that the effective area screened by the foreground objects may be significantly larger than their physical cross-section. As an example, large type Ia supernova surveys typically cover only $\sim 10^{-4}$ part of the sky [113]. It thus appears a natural assumption that the well-observed supernovae consist of objects with minimum amount of intervening material on their foreground.

In addition to the kind of bias discussed above, there may be other kind of selection effects, arising from the fact that the dimmer the object the harder it is to detect. It appears non-trivial how this would affect the observed distance-redshift relations, since on one hand it could favor objects with higher luminosity but on the other hand disfavor objects with high redshift. In any case, for the statistically most relevant brighter supernovae at lower redshifts, we would expect this effect to be less significant. Furthermore, if the type Ia supernovae would be perfect standard candles, the bias towards higher intrinsic luminosity should not exist at all. In addition to selection effects, there could also be other systematic effects related to observational techniques [114]. In this thesis, we simply assume that such effects do not play a significant role.

Unfortunately, due to the complexity of the cosmological structures, we have not been able to derive quantitative estimates for the significance of the selection effects. A possible way to obtain such estimates could be a computer simulation where one studies the difference of the average expansion rate along our line of sight relative to the volume-averaged global expansion rate. Altogether, we can currently neither vindicate nor rule out the conjecture that selection effects in the cosmological observations are significant.

5.1.4 Total effect of the inhomogeneities

As we have discussed in Sects. 5.1.1–5.1.3, it appears very difficult to derive from theory a realistic inhomogeneous model that would take into account all the relevant effects of the inhomogeneities. Instead of dealing with these difficulties, in Paper 3, we have employed a phenomenological approach based on physically justified assumptions. The model contains a transition redshift to
the void-dominated era as a new free parameter and the applied mathematical framework is provided by a generalized Dyer-Roeder method, briefly reviewed in Sect. 5.1.5. In this approach, the observable or optical distance-redshift relations are determined by the average properties of the space along our line of sights to the objects, instead of the global volume-averages. The aim of the model is to provide only a crude description of what the overall effect of the cosmic inhomogeneities on the observed distance-redshift relations may be. Although based on assumptions, in the end, one can use the naturalness of the required value for the new parameter to judge the viability of the model, e.g. compared to the standard FRW cosmology with \( \Lambda \) as the additional phenomenological parameter.

Besides the phenomenological Dyer-Roeder method, the mathematically exact Swiss cheese solutions of the Einstein equations have been used to estimate the effect of the nonlinear inhomogeneities on the cosmological observations [81,115–118]. In these models, the cheese has exact FRW geometry whereas the geometry in the holes is described by the LTB metric (3.2). However, it appears difficult to properly capture the relevant effects of inhomogeneities on light propagation within the exact Swiss cheese models, since the holes cannot model true over- or underdensities: in order to fulfill the exact junction conditions, the average mass density of the holes must be the same as the density of the homogeneous background. Due to the smoothness of the exact solution, a light ray going through the holes thus unavoidably travels also through the overdensities, which in reality should make these photons more unlikely to reach us. On the other hand, a light ray propagating only in the cheese is unaffected by the holes since the cheese has the exact FRW geometry. The constraints from the junction conditions thus appear to restrict also the physical degrees of freedom relevant for backreaction. Because of these issues, we prefer the Dyer-Roeder method in modeling light propagation in the inhomogeneous universe.

5.1.5 The generalized Dyer-Roeder method

Perhaps the first to discuss light propagation in the clumpy universe was Zel’dovich in 1963 [78]. Before that, the idealized FRW geometry had been used as the basis in deriving theoretical predictions for the observable distance-redshift relations. Later in the 60’s, Zel’dovich’s idea was elaborated at least by Bertotti [79], Gunn [80] and Kantowski [81]. However, it was the works by Dyer and Roeder in the 70’s that made the idea more famous [82,83]; consequently, the distance-redshift relations allowing for the effects of clumping of matter on light are usually referred to as the Dyer-Roeder relations.

It seems that the interest towards the effects of clumpiness on light propagation suffered a hiccup after 1976, when Weinberg argued against the use of the Dyer-Roeder version of the distance-redshift relations in place of the conventional FRW ones [119]. He proposed that the effect of light propagating in the empty intergalactic space would, on average, be canceled by an equal but opposite effect due to gravitational lensing caused by the matter clumps. However, Weinberg’s argument has later been challenged [120]. In addition, as we pointed out in Paper 3, the original Dyer-Roeder relations do not take two crucial physical properties of the universe into account: the inhomogeneities in the expansion rate and the growth of the inhomogeneities. Indeed, as we further demonstrated, the evolving inhomogeneous expansion can induce much stronger corrections to the observable distance-redshift relations than the mere clumpiness of matter in the original formulation of Dyer and Roeder – and hence unlikely to get counterbalanced by the gravitational lensing.

Before presenting the generalized version of the Dyer-Roeder method, we first discuss the original
formulation to justify the modifications; see Refs. [82, 83, 121-124] and Paper 3 for more details. At low redshifts, the original Dyer-Roeder luminosity distance reads as:

\[ d_L^{\text{DR}}(z) = H_0^{-1} \left[ z + \frac{1}{4} z^2 + \left( -\frac{1}{8} + \frac{1}{4} \alpha \right) z^3 + O(z^4) \right], \tag{5.4} \]

where the constant smoothness parameter \( \alpha \) represents the fraction of the matter density along our line of sight \( \rho_m(z) \) to the volume-averaged matter density \( \bar{\rho}_m(z) \) of the whole universe; that is, \( \rho_m(z) \equiv \alpha \bar{\rho}_m(z) \). The value \( \alpha = 0 \) corresponds to a universe where all the matter is concentrated into opaque clumps and the value \( \alpha = 1 \) to a perfectly homogeneous FRW universe. By comparing the Dyer-Roeder luminosity distance (5.4) with its counterpart in the homogeneous \( \Lambda \)CDM model (5.3), we see that there is no way \( \alpha \) could mimic \( \Lambda \), as it does not even appear in the all-important second-order term. In fact, as the inspection of the higher order terms reveals, the effect of clumpiness via \( \alpha < 1 \) seems to effectively increase the deceleration of the expansion in concordance with the original results of Dyer and Roeder [82, 83].

However, on physical grounds, one could expect just the opposite result: due to the local gravitational attraction of matter, a region of space with low mass density should decelerate less than a denser one. This can also be seen from the Raychaudhuri equation (2.65), which we write as:

\[ -\frac{1}{V^{1/3}} \frac{d^2}{d\tau^2} V^{1/3} = \frac{4\pi G}{3} \rho_m + \frac{1}{3} (\sigma^2 - \omega^2), \tag{5.5} \]

where the shear \( \sigma^2 \) and rotation \( \omega^2 \) manifest tidal effects or the Weyl curvature, \( \tau \) is the proper time of the matter particles and \( V \) is a local volume element, making \( V^{1/3} \) a generalized scale factor. As expected, in a region far enough from the high-density filaments so that the Weyl part can be neglected \( \sigma^2 - \omega^2 \ll 4\pi G \rho_m \), the deceleration or the left hand side of Eq. (5.5) is the smaller the lower the local matter density \( \rho_m \).

On the grounds of the above arguments, we proposed in Paper 3 that in order to make the Dyer-Roeder method more accurate, the global expansion \( \bar{H}(z) \) should also be replaced with the average expansion rate along our line of sight \( H(z) \). For this, we introduced another parameter \( \beta \), the ratio of the expansion rate along our line of sight to the FRW value, i.e. \( H(z) \equiv \beta \bar{H}(z) \). Furthermore, we also included the fact that structures evolve in the universe; that is, we let both of the parameters depend on the redshift: \( \alpha(z) \) and \( \beta(z) \). The redshift-dependence of \( \alpha \) was in fact already proposed by Linder in 1988 [121], but perhaps because the effects of clumping in the original Dyer-Roeder method are so weak (see Eq. (5.4)), the redshift dependence of \( \alpha \) has won very little attention in the literature [126]. Instead, as we demonstrate in Eq. (5.13), it is only after taking into account both the parameter \( \beta \) and its dependence on the redshift that the results start to go more hand in hand with intuition.

To calculate the observable distance-redshift relations in the generalized Dyer-Roeder method, we need to know the functions \( \alpha(z) \) and \( \beta(z) \). An entirely phenomenological approach would be to determine them purely from the cosmological observations. In contrast, an ideal approach would be to derive \( \alpha(z) \) and \( \beta(z) \) from the Einstein equations starting with some initial perturbations given at a fixed moment in the early universe. In Paper 3, we employed an analytic approach that falls somewhere between these two extremes: based on theoretically and observationally motivated assumptions, we deduced a form for the functions \( \alpha(z) \) and \( \beta(z) \) that contain a single free parameter to be determined from the observations.

Firstly, motivated by the observed isotropy of the CMB, we assumed homogeneous early universe, implying a constraint between \( \alpha(z) \) and \( \beta(z) \). Secondly, we assumed a separate LTB universe
approximation in order to obtain an explicit form for the constraint, hence essentially given by Eq. (3.57). For the leftover free function, we used the dimensionless expansion rate (4.13) along our line of sight and denoted it by $h(z)$, not to be confused with $h$ without an argument: $H_0 = 100h$ km s$^{-1}$ Mpc$^{-1}$. Thirdly, we assumed that, in the late universe, light propagates mainly in freely expanding voids, implying $h(0) \simeq 1$, while $h(1100) \simeq 2/3$ for the homogeneous early universe. Finally, by assuming a smooth transition to the void-dominated era, we chose

\[
h(z) = (2 + e^{-z/z_0})/3 ,
\]

(5.6)

where the transition redshift $z_0$ is a free parameter.

We summarize here the relevant equations of the generalized Dyer-Roeder method; the details of the derivation and physical justification of the equations can be found in Paper 3. The differential equations for the observable distance measures are:

\[
(1 + z)^2 H^2(z) d_A''(z) + \left[ 2(1 + z) H^2(z) + (1 + z)^2 H'(z) H(z) \right] d_A'(z) + \left[ 4 \pi G \rho_m(z) + \frac{16}{3} \pi G \rho_r(z) \right] d_A(z) = 0
\]

(5.7)

\[
(1 + z)^2 H^2(z) d_c''(z) + H'(z) H(z) (1 + z)^2 d_c'(z) + \left[ 4 \pi G \rho_m(z) + \frac{16}{3} \pi G \rho_r(z) - H'(z) H(z) (1 + z) \right] d_c(z) = 0
\]

(5.8)

\[
(1 + z)^2 H^2(z) d_L''(z) + \left[ H'(z) H(z) (1 + z)^2 - 2 H^2(z) (1 + z) \right] d_L'(z) + \left[ 4 \pi G \rho_m(z) + \frac{16}{3} \pi G \rho_r(z) + 2 H^2(z) - 2 H'(z) H(z) (1 + z) \right] d_L(z) = 0
\]

(5.9)

where $' \equiv \partial/\partial z$ and $d_L(z) = (1 + z) d_c(z) = (1 + z)^2 d_A(z)$, satisfying the initial conditions: $d_X(0) = 0$, $d_X'(0) = H^{-1}_0$ with $X \in \{ A, c, L \}$ and

\[
H(z) = t_0^{-1} h(z) (1 + z) \sqrt{1 + 9z(1 - h(z))^2 + z(z + 2) \Omega_r}
\]

(5.10)

\[
\frac{16}{3} \pi G \rho_r(z) = \frac{16}{3} \pi G \rho_r(0) (1 + z)^4 = \frac{8}{9} t_0^{-2} \Omega_r (1 + z)^4
\]

(5.11)

\[
4 \pi G \rho_m(z) = \frac{27}{2} t_0^{-2} h^2(z) (1 - h(z))^2 (1 + z)^3
\]

(5.12)

where $\Omega_r \equiv \rho_r(0)/(3H_0^2/8\pi G)$ is the smooth radiation component and $h(z)$ describes average expansion rate along our line of sight. In the data analysis of Sect. 5.2, we use Eq. (5.6) to determine $h(z)$. The age of the universe $t_0$ is related to the FRW Hubble constant via $H_0 = t_0^{-1}/2$, whereas the relation to the optically deduced Hubble constant is simply $H_0 = t_0^{-1}$.

Let us solve Eq. (5.9) as a Taylor series to third power in redshift to illustrate how the nonlinear inhomogeneities could modify the luminosity distance at low redshifts:

\[
d_L(z) = H_0^{-1} \left[ z + \frac{1}{2} (1 - h'(0)) z^2 + \frac{1}{6} \left(2 h'^2(0) - h''(0)\right) z^3 + \mathcal{O}(z^4) \right].
\]

(5.13)

Here the nonlinear structures, encapsulated in the function $h(z)$, have clearly much more notable role than in the original Dyer-Roeder method in Eq. (5.4), where $1 - \alpha$ describes deviation from the homogeneity. In addition, the effect of the inhomogeneities in Eq. (5.13) is importantly in the
opposite direction, mimicking $\Lambda$ with $h'(0) < 0$, as can be seen by comparing to the expression of the luminosity distance for the flat $\Lambda$CDM model (5.3). In fact, the correspondence between Eqs. (5.3) and (5.13) can be made exact to third order in redshift by choosing: $h'(0) = (1 - 3\Omega_\Lambda)/2$ and $h''(0) = 9(1 - (\Omega_\Lambda - 2/3)^2)/4$. The analogy can be taken even further by expanding Eqs. (4.12) and (5.9) to $n$th order and solving the derivatives of $h(z)$ to $(n - 1)$th order in terms of $\Omega_\Lambda$. This works because each new term in the expansion contains a higher derivative of $h(z)$ evaluated at $z = 0$ and thus corresponds to an independent degree of freedom. Obviously, the correspondence is not bijective: one cannot mimic a general structure formation function $h(z)$ with a constant $\Omega_\Lambda$, but this requires a quintessence field with a redshift dependent equation of state parameter $w(z)$. In any case, we are not interested in reproducing the $\Lambda$CDM model, but rather in finding out how the physically justified simple model with $h(z) = (2 + e^{-z/\Omega_\Lambda})/3$ fits the observations.

5.2 A concordant inhomogeneous model of the universe

In this section, we review how the model of Paper 3, where Eqs. (5.7)–(5.9) determine the observable distance-redshift relations, fits various cosmological observations: the cosmic microwave background anisotropy, the position of the baryon acoustic oscillation peak inferred from the galaxy distribution, the magnitude-redshift relations of type Ia supernovae, the local measurements of the Hubble constant and the Big Bang (BB) nucleosynthesis. At large redshifts, the considered model is similar to the linear CHDM model by Hunt and Sarkar [60,61], which is based on a FRW spacetime with the following basic properties: no dark energy or $\Lambda = 0$, cold dark matter proportion $\Omega_{CDM} = 0.8$, hot dark matter proportion $\Omega_{HDM} = 0.1$, baryonic matter proportion $\Omega_B = 0.1$, flat spatial geometry $k = 0$ and a low Hubble constant $H_0 = 44$ km/s/Mpc.

The linear CHDM model fits the CMB angular power spectra better than the standard $\Lambda$CDM model. It is also consistent with the nucleosynthesis constraints and with most of the features in the matter power spectrum of the galaxy distribution surveys. However, the model seems to fail in matching the observed magnitude-redshift relations of type Ia supernovae and the detected position of the baryon acoustic oscillation peak in the matter power spectrum and it is also in disagreement with the local measurements of the Hubble constant. For more details of the linear CHDM model, see [61].

In the late universe or at low redshifts, our model differs from the linear CHDM model as, instead of determining the optical properties from the homogeneous FRW relations, we use the generalized Dyer-Roeder method of Sect. 5.1.5 to allow for the effects of the nonlinear structure formation. We still assume the growth of small deviations from homogeneity to follow the linear perturbation theory, so apart from the changes in the associated distance measures, the perturbation spectra are assumed to be essentially unaltered. We call this the nonlinear CHDM model. Altogether, we pose five observational tests for this model:

1. To obtain the same value for the angular diameter distance to the last scattering surface as in the linear CHDM model.
2. To match the position of the baryon oscillation peak in the matter power spectrum.
3. To fit the observed magnitude-redshift relations of the type Ia supernovae.
4. To explain the locally measured value of the Hubble constant.
5. To remain within the observed limits on the baryon density set by the BB nucleosynthesis.

5.2.1 The cosmic microwave background

For the values $\Omega_B = 0.1$, $h = 0.44$ and $\Omega_0 = 1$, the fitting formula of Hu and Sugiyama [126] yields $z_{\text{dec}} = 1101.35$ as the decoupling redshift, for which we use the approximation $z_{\text{dec}} \simeq 1100$ throughout.

The fit of the linear CHDM model to the WMAP 3-year data has been thoroughly performed by Hunt and Sarkar [61] so we only considered how the nonlinearities could modify this. As the nonlinear structures appear only in late times, it is essentially only the comoving angular diameter distance $d_c(z)$ and the late integrated Sachs-Wolfe effects that can change. However, as the model already fits the CMB data without the nonlinear structures, we require that adding the nonlinearities would not alter the distance to the last scattering surface, $d_c(1100)$. Physically, this is possible because the negative spatial curvature and the faster Hubble expansion of the voids have the opposite effects on the comoving angular diameter distance $d_c(z)$, so with a suitable value for the new parameter $z_0$, their effects can cancel.

By determining $d_c(z)$ numerically from Eq. (5.8), and $\tilde{d}_c(z) = \tilde{d}_L(z)/(1 + z)$ from Eq. (4.12), we find that the value for $z_0$ that sets $d_c(1100)$ and $d_c(1100)$ identical, is $z_0 = 0.347$. Values in the range $z_0 = 0.30...0.39$ give less than 1% discrepancy between $d_c(1100)$ and $d_c(1100)$. Note that although equal around $z \approx 1100$, the functions $d_c(z)$ and $\tilde{d}_c(z)$ are not equal at lower redshifts.

Varying the Hubble constant does not change the limits on $z_0$ as $\dot{H}_0$ has essentially the same effect on both the "background" value $\tilde{d}_c(1100)$ and the value $d_c(1100)$ which takes into account the nonlinearities. To conserve the successful angular power spectrum found in Ref. [61], we fix the FRW Hubble constant to the value $\dot{H}_0 = 44$ kms$^{-1}$Mpc$^{-1}$.

Within the linear perturbation theory, a flat matter dominated FRW model does not have a late integrated Sachs-Wolfe (LISW) effect, because the perturbed gravitational potentials do not evolve in time, $\dot{\Phi} = 0$. However, this does not necessarily hold in a universe with nonlinear structures if the light accepted for the CMB data analysis has propagated mostly in voids. Indeed, with $\dot{\Phi}$ potentially non-zero along our line of sight, one would expect a non-trivial LISW effect from nonlinear inhomogeneities also in a matter dominated universe with initially flat FRW geometry. On the other hand, as the LISW affects the angular power spectrum only at low multipoles $l$, its contribution to the goodness of the overall fit could be insignificant due to the large cosmic variance $\propto (2l + 1)^{-1}$. Moreover, it might in fact be that the effects of the local structures dominate the angular power at the lowest multipoles [127,128]. It is also possible that, even if the selection effects were significant in the supernova observations, they may be less important in the CMB observations. Altogether, along with other more involved issues, we have postponed a quantitative discussion of the LISW to a future work.

Apart from the potential modifications due to the LISW effect at small $l$, the angular power spectrum of the nonlinear CHDM model with $z_0 = 0.347$ has been plotted in Fig. 1 of Paper 3.

5.2.2 Galaxy distribution

As long as the nonlinear inhomogeneities only affect the properties of light, the evolution of the linear perturbations remains the same as in the linear CHDM model. In the case of the CMB power spectrum this seems obvious, since the relevant evolution of the perturbations happened at
$z < 1100$ when the nonlinearities were negligible. However, as the galaxy distribution also reflects the dynamics of the late universe, it is not so obvious that the nonlinear inhomogeneities would not play a dynamical role here. We nevertheless neglect the potential effects of the nonlinearities on the evolution of the perturbations, but remind that, in a more realistic treatment, these effects should be taken into account in the matter power spectrum.

As shown in Refs. [61, 129], the linear CHDM model fits the matter power spectrum apart from the position of the baryon acoustic oscillation peak. The modifications in the distance-redshift relations caused by the nonlinearities open up a possibility that the nonlinear CHDM model would also fit the position of the baryon acoustic oscillation peak. To study this, we used the data from the Sloan Digital Sky Survey given in Ref. [5].

The detected position of the baryon acoustic oscillation peak in the matter power spectrum constrains the distance to the redshift $z = 0.35$, corresponding to the acoustic oscillation scale today. More specifically, due to the different scaling of radial and angular distances, the exact distance measure constrained by the data reads as [5]:

$$ D_V(z) \equiv \left[ \frac{zd_\ell^2(z)}{H(z)} \right]^{1/3}. $$

As the actual observations are made in the redshift space and the distance-redshift relation depends on the cosmological model, the "measured" values quoted in [5] are prior-dependent. We take this into account by using the more conservative $2\sigma$ error limits. The value of the distance (5.14) to the redshift $z = 0.35$ deduced from the SDSS data is

$$ D_V(0.35) = 1370 \pm 150 \text{ Mpc}, $$

(5.15)

where the limits corresponding to $2\sigma$ errors have been projected from Fig. 7 of Ref. [5]. With $H_0 = 66 \text{ kms}^{-1}\text{Mpc}^{-1}$ fixed, we find that the constraint set by Eq. (5.15) on the transition redshift parameter in Eq. (5.6) is $z_0 > 0.25$, whereas the best fit value from the CMB analysis, $z_0 = 0.347$, gives $D_V(0.35) = 1480 \text{ Mpc}$. Fixing $z_0 = 0.347$ requires $H_0 = 71 \text{ kms}^{-1}\text{Mpc}^{-1}$ or $H_0 = 47 \text{ kms}^{-1}\text{Mpc}^{-1}$ to give the best fit value $D_V(0.35) = 1370 \text{ Mpc}$.

It is possible to combine the distance (5.15) inferred from the galaxy surveys with the distance $d_c(1100)$ inferred from the CMB data to obtain an additional constraint on the ratio:

$$ R_{0.35} \equiv \frac{D_V(0.35)}{d_c(1100)}.$$

(5.16)

It should be again noted that due to the prior-dependency of the distance-redshift relations, Eisenstein et. al. quote several "measured" values for $R_{0.35}$. As $d_c(1100) = 13190 \text{ Mpc}$ is the best fit value to the CMB data in the CHDM model, Eq. (5.15) yields:

$$ R_{0.35} = 0.104 \pm 0.011,$$

(5.17)

where the limits should roughly correspond to $2\sigma$ errors in the CHDM model. We find that the values of $z_0$ corresponding to these limits are: $z_0 = 0.29...0.83$, with $z_0 = 0.53$ yielding the best fit value $R_{0.35} = 0.104$. As $R_{0.35}$ is a ratio of distances, it is independent of $H_0$.

Albeit the nonlinear CHDM model with $h(z) = (2 - e^{-z/z_0})/3$ is not as concordant with the SDSS data as with the other data sets, the concordance value $z_0 = 0.35$ still falls within the experimental limits. Moreover, the results represent a huge improvement compared to the linear CHDM model, which yields the values $D_V(0.35) = 1760 \text{ Mpc}$ and $R_{0.35} = 0.134$ that are both way off the rather conservative limits of Eqs. (5.15) and (5.17).
5.2.3 Type Ia supernovae

We use the Riess et. al. gold sample of 182 type Ia supernovae [3] to study how the nonlinear CHDM model fits the observations. For this, we use the conventional $\chi^2$ test:

$$\chi^2 = \sum_{n=1}^{182} \left( \frac{m_{\text{obs}}(z_n) - m(z_n)}{\sigma_n} \right)^2,$$

(5.18)

where $m(z_n) = 5 \log_{10}(d_L(z_n)/\text{Mpc}) + 25$ is the theoretical prediction for the magnitude and $\sigma_n$ is the estimated error for the measured magnitude $m_{\text{obs}}(z_n)$ of a source at the redshift $z_n$.

Just like the flat ΛCDM model with $\Omega_\Lambda$ and $H_0$ as free parameters, the nonlinear CHDM model has two free parameters: the transition redshift $z_0$ and, say, the optical Hubble constant $H_0$. There is, however, a perfect degeneracy between $H_0$ and the maximum intrinsic luminosity of the type Ia supernovae, $L_{\text{Ia}}$, so the uncertainty in the value of $L_{\text{Ia}}$ reflects a similar uncertainty in the value of $H_0$ constrained from this data. For this reason, the combination $(H_0, L_{\text{Ia}})$ is sometimes marginalized in the data analysis but we choose to instead use the fixed value for $L_{\text{Ia}}$ employed by Riess et. al. in Ref. [3].

By letting both parameters vary, we find the best fit values $H_0 = 65.5 \text{ kms}^{-1}\text{Mpc}^{-1}$ and $z_0 = 0.39$ that yield $\chi^2 = 161.1$. A little inspection in the parameter space shows that there is degeneracy between $z_0$ and $H_0$, such that one can obtain good two-parametric fits ($\Delta \chi^2 \lesssim 5$) roughly in the range: $z_0 = 0.2...0.8$. Keeping instead $H_0$ fixed to the concordance value $66 \text{ kms}^{-1}\text{Mpc}^{-1}$, we find that the best fit transition redshift is $z_0 = 0.34$, yielding $\chi^2 = 161.4$. The increase $\Delta \chi^2 = 2.7$ corresponds to $2\sigma$ errors in a one-parametric fit and yields for the limits: $z_0 = 0.34 \pm 0.10$. The absolute fit $161.4/182 = 0.89$ is good and represents a tremendous improvement to the linear CHDM model where $\chi^2 = 1098$ for $\dot{h} = 0.44$ and even the best fit value $\ddot{h} = 0.54$ gives $\chi^2 = 283.4$.

For comparison, the best fit parameters of the ΛCDM model $\Omega_\Lambda = 0.67$ and $\dot{h} = 0.63$ yield $\chi^2 = 305.2$, but this can be fixed by changing the assumed value of $L_{\text{Ia}}$ or equivalently changing $\dot{h}$. Indeed, we find that the value $\ddot{h} = 0.642$ gives the best fit for the concordance $\Omega_\Lambda = 0.73$ ΛCDM model with $\chi^2 = 162.5$. Overall, the conclusion from comparing the results between the nonlinear CHDM model and the ΛCDM model is that the goodness of their fits to the Riess et. al. data are essentially equal. In Fig. 5.1, we plot the magnitude-redshift relations, with the $\ddot{h} = 0.66$ empty Milne universe relation subtracted, for three different models: the linear CHDM model with $\dot{h} = 0.54$, the ΛCDM model with $\Omega_\Lambda = 0.67$ & $\ddot{h} = 0.63$ and the nonlinear CHDM model with $z_0 = 0.34$ and $\dot{h} = 0.66$.

5.2.4 The Hubble Key Project

The best fit values obtained from the different data sets of Sects. 5.2.1, 5.2.2 and 5.2.3 for the optical Hubble constant are all consistent with the low redshift measurements of the expansion rate that should specifically probe the optical Hubble constant; the concordance value $H_0 = 66 \text{ kms}^{-1}\text{Mpc}^{-1}$ sits firmly within the limits of the final results from the Hubble Key Project: $H_0 = 72 \pm 8 \text{ kms}^{-1}\text{Mpc}^{-1}$ [51] or $H_0 = 62.3 \pm 6.3 \text{ kms}^{-1}\text{Mpc}^{-1}$ [52]. In addition, there are persistent indications that one gets systematically lower values for the Hubble constant by observing objects at higher redshifts [54]. For the standard FRW cosmology this feature is an anomaly, while in the nonlinear CHDM model it is a prediction: towards higher redshifts, the
optically perceived expansion should decrease relative to the FRW expansion due to the fact that the earlier universe was more homogeneous.

5.2.5 The Big Bang Nucleosynthesis

Finally, we comment on the nucleosynthesis. The overall result from deducing the abundances of the light elements from the spectra of stars and galaxies is that the baryon density falls within the range $0.017 \leq \Omega_B h^2 \leq 0.024$ at 95% confidence level [53]. As long as the early universe was nearly homogeneous and the constants of nature do not vary, these limits are model-independent. Therefore, the observed value should not be affected by the nonlinear inhomogeneities in the late-time universe so that the value for the baryon density, $\Omega_B h^2 = 0.1 \cdot 0.44^2 = 0.0194$, falls within the above quoted observational limits also in the nonlinear CHDM model.

5.2.6 A concordant nonlinear CHDM model with no dark energy

By combining the results from the analysis of the different data sets in Sects. 5.2.1-5.2.5, we obtain a concordant cosmological model with the following properties: the optical Hubble constant $h = 0.66$, the transition redshift to the era when nonlinear structures become dominant $z_0 = 0.35$, baryon proportion $\Omega_B = 0.1$, cold dark matter proportion $\Omega_{CDM} = 0.8$, hot dark matter proportion $\Omega_{HDM} = 0.1$ and no dark energy $\Lambda = 0$. All the data sets are consistent with these values and, apart from the SDSS data, the concordance value for $z_0$ is actually very close
to the best fit value of each separate data set; see Table 5.1. Moreover, the concordance value for $H_0$ yields $t_0 = H_0^{-1} = 14.8$ Gyr as the age of the universe, consistent with the lower bound determined from the astronomical estimates for the age of the oldest globular clusters: $12.7 \pm 0.7$ Gyr [130,131]. The main properties of the best fit nonlinear CHDM model are summarized in Table 5.2.

Within the considered model, the physical interpretation for the perceived acceleration seems to be clear: due to the clumping of matter and formation of voids, the regions the detectable light traverses become emptier and emptier relative to the average, thus increasing the expansion rate along our line of sight without local acceleration. Optically, the resulting phenomenon bears close resemblance to the local Hubble Bubble, discussed in Sect. 5.1.2; see in particular Eqs. (5.2), (5.3) and the paragraph thereafter. However, unlike with a large void in the local matter distribution, it is our special location in time, not in space, that is here crucial for the effect, so the cosmological principle is not violated. On the contrary, since every observer in the late universe sees the history of the structure formation, the phenomenon should look essentially the same no matter at which spatial point the observer sits in the universe5. The situation is somewhat akin to the Hubble law whereby every observer sees the galaxies rushing away at growing speed with distance.

<table>
<thead>
<tr>
<th>Description</th>
<th>CMB</th>
<th>BAO peak</th>
<th>Type Ia SNe</th>
<th>Concordance value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition redshift</td>
<td>$z_0$</td>
<td>$0.35 \pm 0.04$</td>
<td>$0.53 \pm 0.24$</td>
<td>$0.34 \pm 0.10$</td>
</tr>
</tbody>
</table>

Table 5.1: The constraints on the transition redshift from each data set and the concordance value.

<table>
<thead>
<tr>
<th>Description</th>
<th>Parameter</th>
<th>Concordance value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition redshift</td>
<td>$z_0$</td>
<td>0.35</td>
</tr>
<tr>
<td>Optical Hubble constant</td>
<td>$h$</td>
<td>0.66</td>
</tr>
<tr>
<td>FRW Hubble constant</td>
<td>$\bar{h}$</td>
<td>0.44</td>
</tr>
<tr>
<td>Age of the universe</td>
<td>$t_0$</td>
<td>14.8 Gyr</td>
</tr>
<tr>
<td>Baryon proportion</td>
<td>$\Omega_B$</td>
<td>10%</td>
</tr>
<tr>
<td>Cold dark matter propor</td>
<td>$\Omega_{CDM}$</td>
<td>80%</td>
</tr>
<tr>
<td>Hot dark matter propor</td>
<td>$\Omega_{HDM}$</td>
<td>10%</td>
</tr>
</tbody>
</table>

Table 5.2: Best fit parameters for the concordant nonlinear model with no dark energy. Only one of the three parameters $h$, $\bar{h}$ and $t_0$ is independent. It is noteworthy that a phenomenologically similar model was found by Wiltshire et. al. [21], but with different interpretation for how the voids give rise to the acceleration [19–24].

5This, at least, distinguishes the light propagation in freely expanding voids as the cause for the acceleration from Wiltshire’s proposal whereby it is crucial that the observer sits in a static or gravitationally bound location, such as inside a galaxy, to perceive the inhomogeneity induced acceleration [19].
5.2.7 Future improvements of the model

Due to the approximations employed in the model of Sect. 5.2 to enable an analytical treatment, we expect that the numerical values of the best fit parameters can represent the optical effects of the inhomogeneities only at a relatively crude level. By replacing the analytical approximations with more thorough numerical procedures and performing the fits to all the data sets from scratch — e.g. making the necessary data calibrations by using a nonlinear model as the baseline instead of the linear ΛCDM model — an even better overall fit is a potential outcome. This concerns especially the matter power spectrum, for which the full analysis requires considerably more work. Instead, due to the mutual degeneracy between the nonlinear structure formation and the cosmological constant ΩΛ, the fit does not gain a noteworthy improvement by keeping ΩΛ in the analysis. In fact, due to the severe theoretical problems with a tiny cosmological constant, discussed in Sect. 4.3.1, we do not even see such a fit meaningful, at least if it is possible to account for the perceived acceleration by inhomogeneities alone.

It may appear that the model of Sect. 5.2 would be nothing but a reproduction of a geocentric LTB bubble, but we see the resemblance only as an inevitable consequence of considering observables that represent averages over different directions on the sky. Naturally, with photons from different directions propagating through different void profiles, the isotropic Dyer-Roeder distance-redshift relations can only represent idealized ensemble averages. The actual observations should thus contain angular variations that are relatively smaller when making observations over distances much greater than a homogeneity scale and hence the largest at low redshifts. Indeed, considerable amount of scatter with unidentified cause has been observed e.g. in the magnitude-redshift relations of the type Ia supernovae and importantly, especially at low redshifts [132–134]; for detailed analysis of scatter in HKP data, see [135]. The observed anomalies at the low multipoles of the angular power spectrum represent corresponding features in the CMB data [127, 128]. Moreover, Gurzadyan et. al. have identified features in the randomness of the raw CMB data that could originate from voids [136]. Altogether, it appears a promising option that these apparent peculiarities would in fact be caused, at least partially, by the non-linear inhomogeneities and thus provide more evidence for their cosmological role addressed in this work. However, due to our lack of knowledge from the detailed (dark) matter distribution and expansion rate, it seems difficult to make precise quantitative predictions for the angular deviations. An important task is to search for correlations between the apparent anomalies of the different data sets.

The transition to the nonlinear era, roughly given by z0 = 0.35, might seem to happen rather late, only Δt = 5.4 Gyr back in time, whereas the oldest stars have formed already Δt > 12 Gyr ago. However, there are in fact several natural reasons to expect such a seemingly late transition. Firstly, while traveling through structures that are still collapsing in earlier times, light is blueshifted which can partially counterbalance the excess redshift caused by the light propagation in voids. Later, however, when the structures have virialized also at larger scales, there are more and more structures which neither collapse nor expand. Those photons that are able to penetrate these virialized overdense regions, then do not blueshift anymore, diminishing this partial counterbalance in late times. Secondly, the large amount of weakly interacting dark matter in the universe prefers to form large halos which cluster into smaller clumps later than the baryonic matter which nonetheless forms the luminous galaxies that indicate the clumping to us. It is also possible that the voids can grow large enough for their shear to be negligible only in the late universe. Finally, when contrasting the inhomogeneity induced acceleration along our line of sight with dark energy driven local acceleration, it is essential to note that the rate
at which gravitational phenomena happen is largely independent of the temperature, whereas non-gravitational interactions happen faster at higher temperatures. Therefore, whereas it seems extremely difficult to explain why a quantum field would arise at \( T \simeq 4 \) K especially if it has had the huge temperature range from \( T_P \simeq 10^{32} \) K to \( T_0 \simeq 2.7 \) K available, it is natural that the inhomogeneity induced acceleration would start as "late" as \( \Delta t \simeq 7 \) Gyr ago (corresponding to redshift \( z \simeq 0.52 \) that can be read off e.g. from Fig. 5.1).

Rather than a derivation or proof that the existing inhomogeneities would really be responsible for the perceived accelerated expansion, the model in this section aims to be only a phenomenological description of what the overall effect of the cosmic inhomogeneities on the observed distance-redshift relations could be. Ideally, a thorough derivation of the model from theory would involve an exact numerical simulation employing the full Einstein equations starting with some initial perturbations given at a fixed moment in the early universe. More realistically, the feasible next step is to replace some of the analytic approximations with more careful numerical procedures; e.g. the transition redshift \( z_0 \) should be a derived, not phenomenological, parameter. Such a treatment necessarily involves more realistic theoretical estimates of the backreaction and an evaluation of the selection effects for the different cosmological data sets separately.

In any case, given the physically well-grounded foundations, the observational evidence for the nonlinear structures as well as the large volume taken up by voids, the concordant fit to the cosmological data and the offered smart solutions to some of the most puzzling theoretical problems in the modern cosmology, it seems perfectly plausible that the main physical properties of the model – most notably that dark energy would not exist – could be correct. An important test is to explain quantitatively the observed angular deviations in various data that are an anomaly for the standard \( \Lambda \)CDM model, but a predicted feature in the anisotropic inhomogeneous models.

54
Chapter 6

Summary

The observations of light from various cosmological objects, in particular the type Ia supernovae, the cosmic microwave background and the galaxy distribution, seem to imply that the expansion of the universe has unexpectedly increased by a factor of $3/2$ along our line of sight roughly during the latter half of the age of the universe. Conventionally, the accelerated expansion has been ascribed to dark energy, a uniform fluid with large negative pressure that causes repulsive gravity. However, it appears that accounting for the perceived acceleration by dark energy entails more questions than answers; most notably: how would the effects of the observed cosmic structures average out, why does the onset of acceleration coincide with the formation of voids and galaxy filaments, why would a quantum field arise at such an unnaturally low temperature $T \sim 4$ K and why would $\Lambda \sim 10^{-123}$ $G^{-1}$ and $H_0 t_0 \sim 1$ have such unexpected values.

In this thesis, we have studied the possibility that, instead of dark energy, the perceived acceleration of the expansion would be caused by the formation of nonlinear structures, ignored in the standard FRW cosmology. We have identified three physical mechanisms as to how the inhomogeneities could explain the observations without dark energy: 1) acceleration of the average expansion due to spatial variations in the expansion rate (often called backreaction), 2) faster local expansion rate due to a large local void and 3) propagation of the detectable light mostly through the voids that expand faster than the global average.

In Paper 1, we used the exact spherically symmetric Lemaitre-Tolman-Bondi dust solution to study the effect of inhomogeneities on the observable distance-redshift relations. We found that the inhomogeneities in the expansion rate, rather than in the matter distribution, have the dominant effect on the observations. We also confirmed that as long as the expansion rate along our past light cone increases by the correct factor, it is observationally irrelevant whether the variation is spatial or temporal, implying strong degeneracy between dark energy and the inhomogeneities. In addition, we demonstrated that although a local void is able to produce acceleration along the past light cone, it does not imply acceleration of the volume-averaged expansion, but is instead a physical mechanism independent from the backreaction.

In Paper 2, we studied the role of scale dependence in the Buchert averaging method, using the flat LTB model as a testing ground. The result found in Paper 1, that volume-averaging is unable to take into account large scale inhomogeneities, was found remediable by allowing for the scale-dependence of the averaged quantities in calculating the observable distance-redshift relations. Indeed, we found an $O(1\%)$ precision at redshifts $z < 2$ in the averaged luminosity and angular diameter distances compared to their exact values. Although considered merely under
the assumption of spherical symmetry, we expect the scale-dependent averaging method to show its full advantage only when applied to more irregularly shaped large-scale inhomogeneities. In addition to its computational simplicity, a virtue of the method is that it provides a unified scheme to take into account both the global backreaction and the local deviations from the average expansion within the same formalism.

In Paper 3, we proposed that the opacity of the nonlinear structures can bias the cosmological observations such that they affect the observed distance-redshift relations similarly as dark energy. To model the overall effect of the inhomogeneities quantitatively, we generalized the Dyer-Roeder method to allow for two crucial physical properties of the universe: inhomogeneities in the expansion rate and the growth of the nonlinear structures. We confirmed the previous results that the clumpiness in the matter distribution alone cannot account for the observations, but demonstrated further that once the inhomogeneous expansion is taken into account, it is possible to have accelerated expansion along our line of sight due to the structure formation. As an application of the generalized Dyer-Roeder method, we found a phenomenological fit to the observations from the CMB anisotropy, the position of the baryon oscillation peak from galaxy surveys, the magnitude-redshift relations of type Ia supernovae, the local Hubble flow and the Big Bang nucleosynthesis, resulting in a concordant model of the universe with 90% dark matter, 10% baryons, no dark energy and 15 Gyr as the age of the universe.

Overall, our studies suggest that the observed cosmic inhomogeneities offer an extremely promising candidate to explain the perceived increase in the cosmic expansion within standard theories and with no dark energy. Indeed, in addition to potentially accounting for the observed distance-redshift relations, the structure formation provides a smart solution to the coincidence problem: if induced by the voids, the onset of the perceived acceleration naturally coincides with the formation of the voids. Furthermore, the observed $H_0 t_0 \sim 1$ corresponds to the characteristic expansion rate of freely expanding voids, $t \sim t_0/2$ is a natural time scale for gravity to form structures on the relevant length scales and $\Lambda = 0$ is compatible with the data if void formation is responsible for the acceleration. In addition, there are various observed features in the universe that are mysterious anomalies for the standard FRW cosmology, such as the cold spot in the CMB, the low multipoles of the CMB, the CMB ellipticity and the scatter in the supernova data, but could be explained naturally by the direction-dependence of the cosmic void distribution on the sky; a more exhaustive list of such features can be found in Wiltshire's work [23].

Since our studies of inhomogeneous cosmology are based on relativistic but classical gravitation, the answer to the long-standing debate of why would $\Lambda = 0$ is naturally beyond the scope here. As is well known, the vanishing of $\Lambda$ cannot be explained within the general theory of relativity whereby $\Lambda$ is allowed to be any constant, determined by the initial conditions of the universe. On the other hand, non-gravitational interactions only care about energy differences. So when interpreted physically as the energy density of vacuum, the value of $\Lambda$ is observationally irrelevant as long as gravity is not concerned. Altogether, an explanation to the value of $\Lambda$ can then only be expected from a theory that satisfactorily unifies gravity and the other fundamental interactions.

Of the different kind of inhomogeneities discussed in this thesis, the local void seems to be the only one that entails serious problems. Firstly, to respect the cosmological principle, we should not sit in the center of a spherically symmetric void, like fitting a local void for the observations seem to require; e.g. for the LTB model to remain consistent with the observed CMB dipole without a fine-tuned peculiar velocity, we should reside within a few percent of the size of the void from its center. In addition, even if the void were not perfectly spherical and even if we were not near the center of the void, the cosmological principle would still be, at least weakly,
violated: there are less galaxies in a void and if born in a random galaxy, it is much more probably located in an environment with lots of galaxies than in a relatively empty void. Another issue is that the required size of the local void seems to be a few times bigger than the typical size of the observed large voids. Moreover, the existence of such a void has not been observationally confirmed although, due to most of the matter in the universe being dark, it may not be easy to disprove it either. In any case, without a natural explanation for these issues, the scenario with a large local void – if taken literally – does not appear likely. Instead, if considered to describe observations averaged over the celestial sphere, the LTB solutions may still have relevance as an effective description for light propagation through several smaller voids.

Indeed, considering the various effects of the inhomogeneities, the most plausible alternative appears to be that a cumulative effect from the observed cosmic network of voids is responsible for the perceived acceleration. This could be due to a combined effect from the physical acceleration of the volume-averaged global expansion and an observational bias due to the well-observed light propagating mostly in voids. Both of these effects may be important but, currently, we have only tentative estimates for the backreaction and a qualitative understanding of the selection effects. In order to obtain realistic estimates, it will be crucial to build more sophisticated models for the inhomogeneities, e.g. by running numerical simulations.

However promising, there are still both observational and theoretical uncertainties with inhomogeneities and one cannot rule out the possibility that a completely new explanation for the acceleration will turn out to be more convincing in the future. In any case, the study of inhomogeneous cosmological models is extremely important for the correct interpretation of the data that is getting more and more precise as the technology advances. For example, given that the evidence for dark energy depends crucially on the unproven assumption that light propagation in the real universe would be equivalent to light traveling in a homogeneous space with the FRW metric, it appears that dark energy can at present be regarded only as a hypothesis. Fortunately, there is hope that more accurate theoretical calculations and better observational data will clarify the physical origin of the perceived accelerated expansion in the not so distant future.
Appendix A

Power expansions of the LTB scale function

When crossing or approaching points \( R_i \) of the LTB solution where \( \Omega(R_i) = 1 \), the expressions (3.58)–(3.62) do not behave well in numerical computations due to the factors of 0/0. Therefore, we calculate the expansion of these quantities in powers of \( \omega(r) \equiv (1 - \Omega(r)) \).

By expanding Eq. (3.58) to fourth order, we obtain:

\[
a(r, t) = \tau^{2/3} \left[ 1 - \frac{1}{5} \left( 1 - \tau^{2/3} \right) \omega(r) + \left(-\frac{18}{175} + \frac{3}{25} \tau^{2/3} - \frac{3}{175} \tau^{4/3} \right) \omega^2(r) + 
\right.
\]

\[
\left. + \left(-\frac{74}{1125} + \frac{76}{875} \tau^{2/3} - \frac{3}{125} \tau^{4/3} + \frac{23}{7875} \tau^2 \right) \omega^3(r) + 
\right.
\]

\[
\left. + \left(-\frac{142349}{3031875} + \frac{2708}{39375} \tau^{2/3} - \frac{846}{30625} \tau^{4/3} + \frac{253}{39375} \tau^2 - \frac{1894}{3031875} \tau^{8/3} \right) \omega^4(r) \right] , \quad (A.1)
\]

where \( \tau \equiv t/t_0 \). Similarly, Eq. (3.59) yields:

\[
a'(r, t) = -\Omega'(r) \tau^{2/3} \left[ -\frac{1}{5} \left( 1 - \tau^{2/3} \right) + 2 \left(-\frac{18}{175} + \frac{3}{25} \tau^{2/3} - \frac{3}{175} \tau^{4/3} \right) \omega(r) + 
\right.
\]

\[
\left. + 3 \left(-\frac{74}{1125} + \frac{76}{875} \tau^{2/3} - \frac{3}{125} \tau^{4/3} + \frac{23}{7875} \tau^2 \right) \omega^2(r) + 
\right.
\]

\[
\left. + 4 \left(-\frac{142349}{3031875} + \frac{2708}{39375} \tau^{2/3} - \frac{846}{30625} \tau^{4/3} + \frac{253}{39375} \tau^2 - \frac{1894}{3031875} \tau^{8/3} \right) \omega^3(r) \right] . \quad (A.2)
\]

By expanding Eq. (3.60), we have:

\[
a'(r, t) = \frac{\Omega'(r)}{\tau^{1/3}} \left[ \frac{2}{15} - \frac{4}{15} \tau^{2/3} + \left(\frac{24}{175} - \frac{8}{25} \tau^{2/3} + \frac{12}{175} \tau^{4/3} \right) \omega(r) + 
\right.
\]

\[
\left. + \left(\frac{148}{1125} - \frac{304}{875} \tau^{2/3} + \frac{18}{125} \tau^{4/3} - \frac{184}{7875} \tau^2 \right) \omega^2(r) + 
\right.
\]

\[
\left. + \left(\frac{1138792}{9095625} - \frac{43328}{118125} \tau^{2/3} + \frac{6768}{30625} \tau^{4/3} - \frac{8096}{118125} \tau^2 + \frac{15152}{1819125} \tau^{8/3} \right) \omega^3(r) \right] . \quad (A.3)
\]
The expansion of Eq. (3.61) reads as:

\[ a''(r, t) = -\Omega'(r)\tau^{2/3}\left[ \frac{1}{5}\left(1 - \tau^{2/3}\right) - 2\left(\frac{18}{175} - \frac{3}{25}\tau^{2/3} + \frac{3}{175}\tau^{4/3}\right)\omega(r) + \\
+ 3\left(-\frac{74}{1125} + \frac{76}{875}\tau^{2/3} - \frac{3}{125}\tau^{4/3} + \frac{23}{7875}\tau^2\right)\omega^2(r) + \\
+4\left(-\frac{142349}{3031875} + \frac{2708}{39375}\tau^{2/3} - \frac{846}{30625}\tau^{4/3} + \frac{253}{39375}\tau^2 - \frac{1894}{3031875}\tau^{8/3}\right)\omega^3(r) \right] + \\
+\Omega^2(r)\tau^{2/3}\left[ -\frac{36}{175} + \frac{6}{25}\tau^{2/3} - \frac{6}{175}\tau^{4/3} - \left(\frac{148}{375} - \frac{456}{875}\tau^{2/3} + \frac{18}{125}\tau^{4/3} - \frac{46}{2625}\tau^2\right)\omega(r) + \\
+ \left(-\frac{569396}{1010625} + \frac{10832}{13125}\tau^{2/3} - \frac{10152}{30625}\tau^{4/3} + \frac{1012}{13125}\tau^2 - \frac{7576}{1010625}\tau^{8/3}\right)\omega^2(r) \right]. \tag{A.4} \]

Finally, by expanding Eq. (3.62), we obtain:

\[ \dot{a}''(r, t) = \frac{\Omega''(r)}{\tau^{1/3}}\left[ \frac{2}{15} - \frac{4}{15}\tau^{2/3} + \left(\frac{24}{175} - \frac{8}{25}\tau^{2/3} + \frac{12}{175}\tau^{4/3}\right)\omega(r) + \\
+ \left(\frac{148}{1125} - \frac{304}{875}\tau^{2/3} + \frac{18}{125}\tau^{4/3} - \frac{184}{7875}\tau^2\right)\omega^2(r) + \\
+ \left(\frac{1138792}{9095625} - \frac{43328}{118125}\tau^{2/3} + \frac{6768}{30625}\tau^{4/3} - \frac{8096}{118125}\tau^2 + \frac{15152}{1819125}\tau^{8/3}\right)\omega^3(r) \right] + \\
+ \frac{\Omega^2(r)}{\tau^{1/3}}\left[ -\frac{24}{175} + \frac{8}{25}\tau^{2/3} - \frac{12}{175}\tau^{4/3} + \left(\frac{296}{1125} - \frac{608}{875}\tau^{2/3} - \frac{36}{125}\tau^{4/3} + \frac{368}{7875}\tau^2\right)\omega(r) + \\
+ \left(\frac{8096}{39375}\tau^2 - \frac{1138792}{3031875} - \frac{20304}{30625}\tau^{4/3} + \frac{43328}{39375}\tau^{2/3} - \frac{15152}{606375}\tau^{8/3}\right)\omega^2(r) \right]. \tag{A.5} \]

Note that we have treated both \( \Omega'(r) \) and \( \Omega''(r) \) as first-order quantities, since in all well-behaved models they are small when \( \omega(r) \) is small, that is when \( \Omega(r) \) gets values close to 1. Although given here up to fourth order for completeness, already the first order term is often enough for the required accuracy.
Bibliography


For an English translation, see:


