Abstract

It is shown that the decomposition theorems for symmetric spatial second-rank tensors, such as the perturbed metric and perturbed Ricci tensor, and the spatial fluid velocity vector imply that, for open, flat or closed Friedmann-Lemaître-Robertson-Walker universes, there are exactly two, unique, independent gauge-invariant quantities which describe the perturbations to the energy density and particle number density.

Using these two new quantities, evolution equations for cosmological density perturbations, adapted to non-barotropic equations of state for the pressure, are derived. The new definitions for the perturbations to the energy density and particle number density allow for an exact non-relativistic limit with a time-independent Newtonian potential. It is shown that density perturbations evolve adiabatically if and only if the particle number density does not contribute to the pressure.

PACS: 04.25.Nx; 98.80.-k; 98.80.Jk

KEYWORDS: Post-Newtonian approximation; perturbation theory; related approximations; cosmology; mathematical and relativistic aspects of cosmology
1 Introduction

The theory related to the linearized Einstein equations is important in cosmology because it
describes the growth of all kinds of structures in the expanding universe, such as stars, galaxies
and microwave background fluctuations.

Lifshitz [1] and Lifshitz and Khalatnikov [2] were the first researchers to derive a cosmological
perturbation theory. Due to the general covariance of the Einstein equations and conservation
laws, they encountered the problem that the linear equations have physical solutions as well as
spurious solutions, the so-called gauge modes, which obscure the physics.

In his seminal article Bardeen [3] demonstrated that the use of gauge-invariant quantities,
i.e., quantities that are invariant under the general linear infinitesimal coordinate transformation
(2.3), in the construction of a perturbation theory ensures that it is free of spurious solutions.
As Bardeen has put it: only gauge-invariant quantities have an inherent physical meaning. Stewart
[4] defines gauge-invariant quantities as ‘quantities that are $\xi^{\mu}$ independent’, which is equiva-
 lent to the definition of Bardeen. Stewart [4], Kodama and Sasaki [5] elaborated and clarified
the work of Bardeen. The article of Bardeen has inspired the pioneering works of Ellis et al. [6–8] and Mukhanov et al. [9, 10]. These researchers proposed alternative perturbation theories using gauge-invariant quantities which differ from the ones used by Bardeen. The equivalence of the approach of Bardeen and the formalism of Ellis et al. has been shown by Bruni et al. [11]. The study of density perturbations has therefore been much refined since Lifshitz initiated his research in 1946.

Motivated by the importance of cosmological perturbation theories in the study of structure formation in the early universe, a thorough analysis of the influential work presented by the above mentioned researchers has been conducted. This study has revealed that, in addition to the use of gauge-invariant quantities, the theory of cosmological perturbations can be improved even further. In this article it will be demonstrated that there are exactly two, unique, gauge-invariant quantities (2.40a) which describe the perturbations to the energy density and particle number density. The third gauge-invariant quantity (2.40b) vanishes identically and, therefore, expresses the fact that, in first-order, a local density perturbation does not affect the global expansion of the universe.

The proof of this theorem is based on the fact that the perturbed Einstein equations and conservation laws (2.10) can, for scalar perturbations, be rewritten as (2.29) which is precisely the perturbed counterpart of the background system (2.7). This is a direct consequence of the theorems of York [12], Stewart and Walker [13] and Stewart [4] on the decomposition of perturbed spatial symmetric second-rank tensors, such as the metric and Ricci tensor, and the perturbed spatial part of the fluid four-velocity. This will be discussed in detail in Sections 2.4 and 2.5. From the systems of equations (2.7) and (2.29), it follows that exactly three scalars (2.9) and their first-order perturbations play a role in the evolution of cosmological density perturbations. Consequently, one can construct exactly two (non-zero) gauge-invariant, i.e., measurable, quantities (2.40a) by elimination of the gauge function $\xi_0(t, \mathbf{x})$ occurring in the transformation (2.38). Hence, the uniqueness of these quantities. Using the quantities (2.40a), equations (2.29) become, in the non-relativistic limit, equal to the Poisson equation (4.8) of the Newtonian Theory of Gravity. Therefore, the quantities (2.40a) are the real, measurable, energy density perturbation and particle number density perturbation, respectively, in the General Theory of Relativity. The Poisson equation (4.8) of the Newtonian Theory of Gravity is, again, a direct consequence of the York, Stewart and Walker decomposition theorems, namely (2.19). In the non-relativistic limit there is some gauge freedom left, since the Newtonian Theory of Gravity is invariant under the gauge transformation (4.6).

Applying the algorithm given in the Appendix, the set of equations (2.41) for the four gauge-dependent quantities $\varepsilon_{(1)}, n_{(1)}, \vartheta_{(1)}, R_{(1)}$ is, using the expressions (2.42) and (2.43), recast in the system (2.44) for the gauge-invariant energy density and particle number density contrast functions $\delta_\varepsilon$ and $\delta_n$, respectively. What will be done, essentially, is to rewrite the linearized Einstein equations for the quantities $\varepsilon_{(1)}$ and $n_{(1)}$ in terms of the measurable energy and particle number densities $\varepsilon_{(1)}^{\text{gi}}$ and $n_{(1)}^{\text{gi}}$, just as in electromagnetism the Coulomb and Ampère laws for the potentials $\Phi$ and $\mathbf{A}$ can be rewritten in the form of the Maxwell equations for the measurable fields $\mathbf{E}$ and $\mathbf{B}$. The final results (2.44) are evolution equations for perturbations in the energy density and entropy in open ($K = -1$), flat ($K = 0$) or closed ($K = 1$) Friedmann-Lemaître-Robertson-Walker (FLRW) universes.
The evolution equations (2.44) and their solutions are, by construction, invariant under general infinitesimal coordinate transformations (2.3). Consequently, there is no need to choose a particular coordinate system to solve these equations and interpret their outcome. Moreover, the evolution equations are valid for all scales, so that there is no need to distinguish between sub-horizon and super-horizon perturbations.

In a companion article the evolution equations (2.44) will be solved. Since the solutions are not contaminated with spurious gauge modes, the physical consequences of the new approach will stand out clearly.

2 Derivation of the Evolution Equations for Density Perturbations

In this section the evolution equations (2.44) for density perturbations in FLRW universes are derived. The background equations (2.7) and first-order perturbation equations (2.10), from which the evolution equations will be derived, can be deduced from the set of Einstein equations (97.11)–(97.13) from the textbook [14] of Landau and Lifshitz and the conservation laws $T^\mu{}_{\nu;\nu} = 0$. For a detailed derivation of the basic equations (2.7) and (2.10), see [15], Sections II and III.

2.1 Equations of State

In Section 3.2 it is shown that barotropic equations of state $p = p(\varepsilon)$ do not take into account entropy perturbations. Since heat exchange of a perturbation with its environment could be important for structure formation in the early universe, realistic equations of state are needed. From thermodynamics it is known that both the energy density $\varepsilon$ and the pressure $p$ depend on the independent quantities $n$ and $T$, i.e.,

$$\varepsilon = \varepsilon(n, T), \quad p = p(n, T),$$

(2.1)

where $n$ is the particle number density and $T$ the temperature. Since $T$ can, in principle, be eliminated from these equations of state, a computationally more convenient equation of state for the pressure will be used, namely

$$p = p(n, \varepsilon).$$

(2.2)

This, non-barotropic, equation of state and its partial derivatives has been included in the evolution equations (2.44).

2.2 Choosing a System of Reference

In order to derive the evolution equations a suitable system of reference must be chosen. Due to the general covariance of the Einstein equations and conservation laws, the General Theory of Relativity is invariant under a general coordinate transformation $x^\mu \to \tilde{x}^\mu(x')$, implying that there are no preferred coordinate systems (Weinberg [16], Appendix B). In particular, the linearized Einstein equations and conservation laws are invariant under a general linear coordinate transformation (i.e., a gauge transformation)

$$x^\mu \to x^\mu - \xi^\mu(t, x),$$

(2.3)
where $\xi^\mu(t, x)$ are four arbitrary, first-order (infinitesimal) functions of time, $x^0 = ct$, and space, $x = (x^1, x^2, x^3)$, coordinates, the so-called gauge functions. Since there are no preferred systems of reference and since the solutions of the evolution equations (2.44) are invariant under the infinitesimal coordinate transformation (2.3), i.e., are gauge-invariant, one may use any coordinate system to derive the evolution equations.

In order to put an accurate interpretation on the gauge-invariant quantities (2.40) one needs the non-relativistic limit. In the Newtonian Theory of Gravity space and time are strictly separated, implying that in this theory all coordinate systems are essentially synchronous. In view of the non-relativistic limit, it would, therefore, be convenient to use synchronous coordinates [1, 2, 14] in the background as well as in the perturbed universe. In these coordinates the metric of FLRW universes has the form

$$g_{00} = 1, \quad g_{0i} = 0, \quad g_{ij} = -a^2(t)\tilde{g}_{ij}(x),$$

(2.4)

where $a(t)$ is the scale factor of the universe, $g_{00} = 1$ indicates that coordinate time is equal to proper time, $g_{0i} = 0$ is the synchronicity condition (see Landau and Lifshitz [14], § 84) and $\tilde{g}_{ij}$ is the metric of the three-dimensional maximally symmetric sub-spaces of constant time. The functions $\xi^\mu$ of the general infinitesimal transformation (2.3) become

$$\xi^0 = \psi(x), \quad \xi^i = \tilde{g}^{ik}\partial_k\psi(x)\int^ct\frac{d\tau}{a^2(\tau)} + \chi^i(x),$$

(2.5)

if only transformations between synchronous coordinates are allowed. The four functions $\psi(x)$ and $\chi^i(x)$ cannot be fixed since the four coordinate conditions $g_{00} = 1$ and $g_{0i} = 0$ have already exhausted all four degrees of freedom, see Weinberg [17], Section 7.4 on coordinate conditions.

A property of a synchronous system of reference (2.4), which is important for the derivation of the evolution equations, is that the space-space components of the four-dimensional Ricci curvature tensor $R^{\mu\nu}$ is broken down into two parts such that one part contains exclusively all time-derivatives of the space-space components of the metric tensor and the second part $R^i_j$ is precisely the Ricci curvature tensor of the three-dimensional sub-spaces of constant time, see (97.10) in the textbook of Landau and Lifshitz [14]. This property makes it possible to apply the decomposition theorems of York, Stewart, and Walker to the perturbed metric and perturbed Ricci tensor.

In conclusion, in order to arrive at the non-relativistic limit and to apply the decomposition theorem of symmetric second-rank three-tensors, a synchronous system of reference is the most appropriate coordinate system to derive the evolution equations (2.44).

### 2.3 Basic Equations for FLRW Universes

In this section the background and first-order Einstein equations and conservation laws are written down. Expressions pertaining to density perturbations, i.e., the expansion scalar, the spatial part of the Ricci curvature tensor and its trace and the three-divergence of the spatial part of the fluid four-velocity, are derived.
2.3.1 Background Equations

The complete set of zeroth-order Einstein equations and conservation laws for an open, flat or closed FLRW universe filled with a perfect fluid with energy-momentum tensor

\[ T^{\mu \nu} = (\varepsilon + p)u^\mu u^\nu - pg^{\mu \nu}, \quad p = p(n, \varepsilon), \]  

(2.6)
is, in synchronous coordinates, given by

\[ 3H^2 = \frac{1}{2}R_{(0)} + \kappa\varepsilon_{(0)} + \Lambda, \quad \kappa = 8\pi G_N/c^4, \]  

(2.7a)
\[ \dot{R}_{(0)} = -2HR_{(0)}, \]  

(2.7b)
\[ \dot{\varepsilon}_{(0)} = -3H\varepsilon_{(0)}(1 + w), \quad w \equiv p_{(0)}/\varepsilon_{(0)}, \]  

(2.7c)
\[ \dot{\theta}_{(0)} = 0, \]  

(2.7d)
\[ \dot{n}_{(0)} = -3Hn_{(0)}. \]  

(2.7e)
The \( G_{0i} \) constraint equations and the \( G_{ij}, i \neq j \), dynamical equations are identically satisfied. The \( G_{ii} \) dynamical equations are equivalent to the time-derivative of the \( G_{00} \) constraint equation (2.7a). Therefore, the \( G_{ij} \) dynamical equations need not be taken into account. In equations (2.7) \( \Lambda \) is the cosmological constant, \( G_N \) the gravitational constant of the Newtonian Theory of Gravity and \( c \) the speed of light. Note that \( w \) is only a shorthand notation for the quotient \( p_{(0)}/\varepsilon_{(0)} \), it does not mean that the equation of state is barotropic. An over-dot denotes differentiation with respect to \( ct \) and the sub-index \( (0) \) refers to the background, i.e., unperturbed, quantities. Furthermore, \( H \equiv \dot{a}/a \) is the Hubble function which is equal to \( H = \frac{1}{3}\theta_{(0)} \), where \( \theta_{(0)} \) is the background value of the expansion scalar \( \theta \equiv u^\mu;_\mu \) with \( u^\mu \equiv c^{-1}U^\mu \) the fluid four-velocity, normalized to unity \( (u^\mu u_\mu = 1) \). A semicolon denotes covariant differentiation with respect to the background metric \( g_{(0) \mu \nu} \). The spatial part of the background Ricci curvature tensor \( R^i_{(0)ij} \) and its trace \( R_{(0)} \) are given by

\[ R^i_{(0)ij} = -\frac{2K}{a^2}\delta^i_j, \quad R_{(0)} = -\frac{6K}{a^2}, \quad K = -1, 0, +1, \]  

(2.8)
where \( R_{(0)} \) is the global spatial curvature. The quantity \( \vartheta_{(0)} \) is the three-divergence of the spatial part of the fluid four-velocity \( u^\mu_{(0)} \). For an isotropically expanding universe the fluid four-velocity is \( u^\mu_{(0)} = \delta^\mu_0 \), so that \( \vartheta_{(0)} = 0 \), expressing the fact that there is no local fluid flow. In the background the coordinate system is essentially co-moving.

From the system (2.7) one may infer that the evolution of an unperturbed FLRW universe is determined by exactly three independent scalars, namely

\[ \varepsilon = T^{\mu \nu}u_\mu u_\nu, \quad n = N^\mu u_\mu, \quad \theta = u^{\mu} u_\mu, \]  

(2.9)
where \( N^\mu \equiv nu^\mu \) is the cosmological particle current four-vector, which satisfies the particle number conservation law \( N^{\mu} u_\mu = 0, \) (2.7e), see Weinberg [16], Appendix B. As will become clear in Section 2.6, the quantities (2.9) and their first-order counterparts play a key role in the evolution of cosmological density perturbations.
2.3.2 Perturbation Equations

The complete set of first-order Einstein equations and conservation laws for the open, flat or closed FLRW universe is, in synchronous coordinates, given by

\[
\begin{align*}
H \dot{h}^k{}_k + \frac{1}{3} \dot{R}^{(1)} = & -\kappa \varepsilon^{(1)}, \\
\dot{h}^k{}_l - \dot{h}^k{}_{ij} = & 2\kappa (\varepsilon^{(0)} + p^{(0)}) u^{(1)i}, \\
\dot{h}^i{}_{jj} + 3H \dot{h}^i{}_{jj} + \delta^i{}_{J} H \dot{h}^k{}_{k} + 2R^{i}{}_{(1)j} = & -\kappa \delta^i{}_{J}(\varepsilon^{(1)} - p^{(1)}), \\
\dot{\varepsilon}^{(1)} + 3H(\varepsilon^{(1)} + p^{(1)}) + (\varepsilon^{(0)} + p^{(0)}) \theta^{(1)} = & 0, \\
\frac{1}{c} \frac{d}{dt} \left[ (\varepsilon^{(0)} + p^{(0)}) u^i_{(1)} \right] - g^{ik}_{(0)} p^{(1)k} + 5H(\varepsilon^{(0)} + p^{(0)}) u^i_{(1)} = & 0, \\
\dot{n}^{(1)} + 3H n^{(1)} + n^{(0)} \theta^{(1)} = & 0, 
\end{align*}
\]

where \( h_{\mu\nu} \equiv -g_{(1)\mu\nu} \) and \( h^{\mu\nu} = +g^{(0)\mu\nu} \) with \( h_{00} = 0, h_{0i} = 0 \) is the perturbed metric, \( h^{i}{}_j = g^{ik}_{(0)} h_{kj} \), and \( g^{(0)}_{ij} \) is the unperturbed background metric (2.4) for an open, flat or closed FLRW universe. Quantities with a sub-index \( (1) \) are the first-order counterparts of the background quantities with a sub-index \( (0) \). A vertical bar denotes covariant differentiation with respect to \( g_{(0)ij} \).

The first-order perturbation to the pressure is given by the perturbed equation of state

\[
p^{(1)} = p_n n^{(1)} + p_\varepsilon \varepsilon^{(1)}, \quad p_n \equiv \left( \frac{\partial p}{\partial n} \right)_\varepsilon, \quad p_\varepsilon \equiv \left( \frac{\partial p}{\partial \varepsilon} \right)_n,
\]

where \( p_n(n, \varepsilon) \) and \( p_\varepsilon(n, \varepsilon) \) are the partial derivatives of the equation of state \( p(n, \varepsilon) \).

The first-order perturbation to the spatial part of the Ricci tensor (2.8) reads

\[
R^{i}{}_{(1)j} = \left( g^{ik} R_{k}{}_{j} \right)^{(1)} = g^{ik}_{(0)} R^{(1)k}{}_{j} + \frac{1}{2} R^{(0)} h^{i}{}_{j}.
\]

Using Lifshitz’ formula, see Lifshitz and Khalatnikov [2], equation (I.3) and Weinberg [17], equation (10.9.1),

\[
\Gamma^{k}{}_{(1)ij} = -\frac{1}{2} g^{kl}_{(0)} (h_{lij} + h_{lj}{}_{i} - h_{ij}{}_{l}),
\]

and the contracted Palatini identities, see [2], equation (I.5) and [17], equation (10.9.2),

\[
R^{(1)ij} = \Gamma^{k}{}_{(1)ij}{}_{k} - \Gamma^{k}{}_{(1)ikk}{}_{j},
\]

one finds, using that \( g^{ij}_{(0)} h^{k}{}_{ij}{}_{k} = g^{ij}_{(0)} h^{k}{}_{ij}{}_{k} \), for the trace of (2.12)

\[
R^{(1)} = g^{ij}_{(0)} (h^{k}{}_{ik}{}_{j} - h^{k}{}_{ij}{}_{k}) + \frac{1}{2} R^{(0)} h^{k}{}_{kk}.
\]

Expression (2.15) is the local perturbation to the global spatial curvature \( R^{(0)} \) due to a local density perturbation.

Finally, \( \theta^{(1)} \) is the first-order perturbation to the expansion scalar \( \theta \equiv u^{\mu}{}_{;\mu} \). Using that \( u^{(0)}_{\mu} = \delta^{\mu}{}_{0} \), one gets

\[
\theta^{(1)} = \vartheta^{(1)} - \frac{1}{2} h^{k}{}_{k}, \quad \vartheta^{(1)} \equiv u^{k}{}_{(1)k},
\]

where \( \vartheta^{(1)} \) is the divergence of the spatial part of the perturbed fluid four-velocity \( u^{(1)}_{\mu} \). The quantities (2.15) and (2.16) play an important role in the derivation of the evolution equations (2.44). Since \( p^{(1)} \neq 0 \), one has \( u^{(1)}_{\mu} \neq 0 \), so that the coordinate system can not be co-moving in the perturbed universe, see the textbook of Landau and Lifshitz [14], §97 and §115. The case \( u^{(1)}_{\mu} \rightarrow 0 \) will be considered in Section 4 on the non-relativistic limit.
2.4 Decomposition of the Spatial Metric, the Spatial Ricci Tensor and the Spatial Fluid Velocity

The decomposition theorems of York, Stewart, and Walker will be used in Section 2.5 to rewrite the full set of Einstein equations (2.10) into a substantially simpler set of equations (2.29) which describe exclusively scalar perturbations. This set, together with the background equations (2.7), form the basis of the evolution equations (2.44).

York [12], Stewart and Walker [13] and Stewart [4] showed that any symmetric second-rank three-tensor, and hence the perturbation tensor \( h_{ij} \) and the perturbed Ricci tensor \( R^{(1)ij} \), can uniquely be decomposed into three parts. For the perturbed metric, one has

\[
h^i_j = h^i_{\parallel,j} + h^i_{\perp,j} + h^i_{*,j},
\]

(2.17)

where the scalar, vector and tensor perturbations are denoted by \( \parallel \), \( \perp \) and \( * \), respectively. The constituents have the properties

\[
h^k_{\perp k} = 0, \quad h^k_{*k} = 0, \quad h^k_{*|k} = 0.
\]

(2.18)

Moreover, York and Stewart demonstrated that the components \( h^i_{\parallel j} \) can be written in terms of two independent potentials \( \phi(t, x) \) and \( \zeta(t, x) \), namely

\[
h^i_{\parallel j} = \frac{2}{c^2}(\phi \delta^i_j + \zeta^i_{|j}).
\]

(2.19)

In Section 4 it will be shown that, for a flat FLRW universe in the non-relativistic limit, the potential \( \phi \) becomes independent of time, the Newtonian potential is \( \varphi(x) \equiv \phi(x)/a^2(t_0) \) and the potential \( \zeta(x, t) \) does not play a role anymore.

Finally, Stewart also proved that the spatial part \( u^{(1)} \) of the perturbed fluid four-velocity can uniquely be decomposed into two parts

\[
u^{(1)} = u^{(1)\parallel} + u^{(1)\perp},
\]

(2.20)

where the constituents have the properties

\[
\tilde{\nabla} \cdot u^{(1)} = \tilde{\nabla} \cdot u^{(1)\parallel}, \quad \tilde{\nabla} \times u^{(1)} = \tilde{\nabla} \times u^{(1)\perp},
\]

(2.21)

and where \( \tilde{\nabla} \) is the generalized vector differential operator, with components \( \tilde{\nabla}^i \equiv \tilde{g}^{ij} \partial_j \).

The perturbed Ricci tensor, \( R^{(1)ij} \), being a symmetric second-rank three-tensor, should obey also the decomposition (2.17) with the properties (2.18), with \( h^i_j \) replaced by \( R^{(1)ij} \). It follows from \( R^{(1)\perp k} = 0 \) and (2.15) and (2.18) that \( h^i_{\perp} \) must obey

\[
h^i_{\perp|k|} = 0,
\]

(2.22)

in addition to \( h^k_{\perp k} = 0 \), whereas \( R^{(1)*} = 0 \) and \( R^{(1)*|k} = 0 \) are automatically fulfilled due to the properties (2.18) of \( h^i_{*} \), see [15], Section IV.

The three different kinds of perturbations will now be considered according to the decompositions (2.17) and (2.20), and their properties (2.18) and (2.21).
Using that $R_{(1)\ast} = 0$ and (2.18), equations (2.10a)–(2.10c) imply that tensor perturbations are not coupled to $\varepsilon_{(1)}$, $p_{(1)}$ and $u_{(1)}$.

From (2.10a), the trace of (2.10c), $R_{(1)\perp k} = 0$ and (2.18) it follows that vector perturbations are not coupled to $\varepsilon_{(1)}$ and $p_{(1)}$. Raising the index $i$ of equations (2.10b) with $g^i_{(0)}$, one finds that these equations read for vector perturbations

$$h^{kj}_{\perp |k|} + 2Hh^{kj}_{\perp |k|} = 2\kappa(\varepsilon_{(0)} + p_{(0)})u^j_{(1)},$$  \hspace{1cm} (2.23)

where it is used that $\dot{g}^i_{(0)} = -2Hg^i_{(0)}$. Taking the covariant derivative of (2.23) with respect to the index $j$ one finds, using also (2.22), that equations (2.23) reduce to $u^j_{(1)j} = 0$, implying with (2.21) that the rotational part $u_{(1)\perp}$ is coupled to vector perturbations.

Finally, since both $h^k_{\mid k} \neq 0$ and $R_{(1)\parallel} \neq 0$, scalar perturbations are coupled to $\varepsilon_{(1)}$ and $p_{(1)}$. It will now be demonstrated that $u_{(1)\parallel}$ is coupled to scalar perturbations, by showing that equations (2.10b) require that the rotation of $u_{(1)}$ vanishes, if the metric is of the form (2.19). Differentiating (2.10b) covariantly with respect to the index $j$ and subsequently substituting expression (2.19) yields

$$2\dot{g}^i_{[ij]} + \zeta^{[i|k][i]|j]} - \dot{\zeta}^{[i|k][i]|j]} = \kappa\epsilon^2(\varepsilon_{(0)} + p_{(0)})u_{(1)ij}.$$  \hspace{1cm} (2.24)

Interchanging $i$ and $j$ and subtracting the result from (2.24) one gets

$$\dot{\zeta}^{[i|k][i]|j]} - \dot{\zeta}^{[i|k][j]|i]} = \kappa\epsilon^2(\varepsilon_{(0)} + p_{(0)})(u_{(1)ij} - u_{(1)ji}).$$  \hspace{1cm} (2.25)

By rearranging the covariant derivatives, (2.25) can be cast in the form

$$\begin{align*}
(\dot{\zeta}^{[i|k][i]|j]} - \dot{\zeta}^{[i|k][j]|i]} - (\dot{\zeta}^{[i|k][j]|i]} - \dot{\zeta}^{[i|k][i]|j]}) \\
+ (\dot{\zeta}^{[i|j]} - \dot{\zeta}^{[i|j]})k = \kappa\epsilon^2(\varepsilon_{(0)} + p_{(0)})(u_{(1)ij} - u_{(1)ji}).
\end{align*}$$  \hspace{1cm} (2.26)

Using the expressions for the commutator of second order covariant derivatives (Weinberg [17], Chapter 6, Section 5)

$$\begin{align*}
A^{i}_{p[j|q]} - A^{i}_{j|p|} &= A^{i}_{k}R_{(0)jkpq} - A^{i}_{k}R_{(0)kjqp} \\
B^{i}_{p[q]} - B^{i}_{q|p} &= B^{k}R_{(0)kqp} \\
\end{align*}$$  \hspace{1cm} (2.27a, 2.27b)

and substituting the background Riemann tensor for the three-spaces of constant time

$$R^{i}_{(0)jkl} = K(\delta^{i}_{k}\tilde{g}_{jl} - \delta^{i}_{l}\tilde{g}_{jk}), \quad K = -1, 0, +1,$$

one finds that the left-hand sides of equations (2.26) vanish identically, implying that the rotation of $u_{(1)}$ is zero. Therefore, only $u_{(1)\parallel}$ is coupled to scalar perturbations.

### 2.5 First-order Equations for Scalar Perturbations

The results of Section 2.4 will now be used to derive a set of equations that describe exclusively scalar perturbations from the system of equations (2.10).

Since scalar perturbations, i.e., perturbations in $\varepsilon_{(1)}$ and $p_{(1)}$, are only coupled to $h^{i}_{\parallel j}$ and $u^{i}_{(1)\parallel j}$, one may replace in (2.10)–(2.16) $h^{i}_{\parallel j}$ by $h^{i}_{\parallel j}$ and $u^{i}_{(1)\parallel j}$ by $u^{i}_{(1)\parallel j}$ to obtain perturbation equations which exclusively describe the evolution of scalar perturbations. From now on, only scalar
perturbations are considered, and the subscript \( \parallel \) will be omitted. Using the decompositions \((2.17)\) and \((2.20)\) and the properties \((2.18)\) and \((2.21)\), it will now be shown that the evolution equations for scalar perturbations can be written in the form

\[
2H(\theta_{(1)} - \vartheta_{(1)}) = \frac{1}{2}R_{(1)} + \kappa \varepsilon_{(1)},
\]

\[
\dot{R}_{(1)} = -2HR_{(1)} + 2\kappa \varepsilon_{(0)} (1 + w) \vartheta_{(1)} - \frac{2}{3}R_{(0)} (\theta_{(1)} - \vartheta_{(1)}),
\]

\[
\dot{\varepsilon}_{(1)} = -3H(\varepsilon_{(1)} + p_{(1)}) - \varepsilon_{(0)} (1 + w) \theta_{(1)},
\]

\[
\dot{\vartheta}_{(1)} = -H(2 - 3\beta^2) \vartheta_{(1)} - \frac{1}{\varepsilon_{(0)} (1 + w)} \nabla^2 p_{(1)} ,
\]

\[
\dot{n}_{(1)} = -3Hn_{(1)} - n_{(0)} \theta_{(1)}.
\]

Note the remarkable similarity between the systems \((2.7)\) and \((2.29)\). In fact, the set \((2.29)\) is precisely the perturbed counterpart of the set \((2.7)\). Just as in the unperturbed case \((2.7)\), the perturbation equations consist of one algebraic equation \((2.29a)\) and four ordinary differential equations \((2.29b)-(2.29e)\) for the five unknown quantities \(\varepsilon_{(1)}, n_{(1)}, \vartheta_{(1)}, R_{(1)}\) and \(\theta_{(1)}\). The systems \((2.7)\) and \((2.29)\) form the basis of the evolution equations \((2.44)\) for density perturbations in closed, flat and open FLRW universes, which will be derived from these two sets of equations in Section 2.7.

The quantity \(\beta(t)\) is defined by \(\beta^2 \equiv \dot{p}_{(0)} / \dot{\varepsilon}_{(0)}\). Using that \(\dot{p}_{(0)} = p_n \dot{n}_{(0)} + p_\varepsilon \dot{\varepsilon}_{(0)}\) and the conservation laws \((2.7c)\) and \((2.7e)\) one gets

\[
\beta^2 = p_\varepsilon + \frac{n_{(0)} p_n}{\varepsilon_{(0)} (1 + w)}.
\]

Finally, the symbol \(\nabla^2\) denotes the generalized Laplace operator with respect to the three-space metric \(\tilde{g}_{ij}\), defined by \(\tilde{\nabla}^2 f \equiv \tilde{g}^{ij} f_{;ij}\).

The derivation of the evolution equations \((2.29)\) for scalar perturbations will now be given. Eliminating \(\dot{h}^k_k\) from \((2.10a)\) with the help of \((2.16)\) yields the algebraic equation \((2.29a)\).

Multiplying both sides of equations \((2.10b)\) by \(g_{(0)}^{ij}\) and taking the covariant derivative with respect to the index \(j\), one finds

\[
g_{(0)}^{ij} \left( \dot{h}^k_{;k[ij]} - \dot{h}^k_{;i[kj]} \right) = 2\kappa (\varepsilon_{(0)} + p_{(0)}) \vartheta_{(1)},
\]

where also \((2.16)\) has been used. The left-hand side of \((2.31)\) will turn up as a part of the time-derivative of the curvature \(R_{(1)}\). In fact, differentiating \((2.15)\) with respect to time and recalling that the background connection coefficients \(\Gamma^k_{(0)ij}\) are for FLRW metrics \((2.4)\) independent of time, one gets, using also \(g_{(0)}^{ij} = -2H g_{(0)}^{ij}\) and \((2.7b)\),

\[
\dot{R}_{(1)} = -2HR_{(1)} + g_{(0)}^{ij} (\dot{h}^k_{;k[ij]} - \dot{h}^k_{;i[kj]}) + \frac{1}{3}R_{(0)} \dot{h}^k_k.
\]

Combining \((2.31)\) and \((2.32)\) and using \((2.16)\) to eliminate \(\dot{h}^k_k\) yields \((2.29b)\). Thus, the \(G_{(1)ij}\) momentum constraint equations \((2.10b)\) have been recast in one equation \((2.29b)\) for the local spatial curvature due to a density perturbation.

It will now be shown that, for scalar perturbations, the dynamical equations \((2.10c)\) need not be considered. For \(i \neq j\) equations \((2.10c)\) are not coupled to scalar perturbations. Taking
the trace of (2.10c) and eliminating the quantity $\dot{h}^k_k$ with the help of (2.16), one arrives at
\[
\dot{\theta}_{(1)} - \dot{\vartheta}_{(1)} + 6H(\theta_{(1)} - \vartheta_{(1)}) - R_{(1)} = \frac{3}{2}\kappa(\varepsilon_{(1)} - p_{(1)}).
\] (2.33)

Using (2.29a) to eliminate the second term of (2.33) yields for the trace of (2.10c)
\[
\dot{\theta}_{(1)} - \dot{\vartheta}_{(1)} = -\frac{3}{2}\kappa(\varepsilon_{(1)} + p_{(1)}).
\] (2.34)

This equation is identical to the time-derivative of the constraint equation (2.29a). Differentiation of equation (2.29a) with respect to time yields
\[
2\dot{H}(\theta_{(1)} - \vartheta_{(1)}) + 2H(\dot{\theta}_{(1)} - \dot{\vartheta}_{(1)}) = \frac{1}{2}\dot{R}_{(1)} + \kappa\dot{\varepsilon}_{(1)}.
\] (2.35)

Eliminating the time-derivatives $\dot{H}$, $\dot{R}_{(1)}$ and $\dot{\varepsilon}_{(1)}$ with the help of (2.7a)–(2.7c), (2.29b) and (2.29c), respectively, yields the dynamical equation (2.34). Consequently, for scalar perturbations the dynamical equations (2.10c) need not be considered.

Finally, taking the covariant derivative of (2.10c) with respect to the metric $g^{(0)ij}$ and using (2.16), one gets
\[
\frac{1}{c}\frac{d}{dt}\left[(\varepsilon_{(0)} + p_{(0)})\theta_{(1)} - g^{(k)(0)}_{ij}(\varepsilon_{(0)} + p_{(0)})\partial_{(1)} - 5H(\varepsilon_{(0)} + p_{(0)})\vartheta_{(1)} = 0,
\] (2.36)

where it is used that the operations of taking the time-derivative and the covariant derivative commute, since the background connection coefficients $\Gamma^k_{ij}$ are independent of time for FLRW metrics. With (2.4), $\tilde{\nabla}^2 f \equiv \tilde{g}^{ij} f_{ij}$ and (2.7c) one can rewrite (2.36) in the form
\[
\dot{\vartheta}_{(1)} + H\left(2 - \frac{3}{\varepsilon_{(0)}}\right)\vartheta_{(1)} + \frac{1}{\varepsilon_{(0)} + p_{(0)}}\frac{\tilde{\nabla}^2 p_{(1)}}{a^2} = 0.
\] (2.37)

Using the definitions $w \equiv p_{(0)}/\varepsilon_{(0)}$ and $\beta^2 \equiv \hat{p}_{(0)}/\dot{\varepsilon}_{(0)}$ one arrives at equation (2.29d).

This concludes the derivation of the system (2.29). As follows from its derivation, this system is, for scalar perturbations, equivalent to the full set of first-order Einstein equations and conservation laws (2.10).

### 2.6 Unique Gauge-invariant Cosmological Density Perturbations

The background equations (2.7) and the perturbation equations (2.29) are both written with respect to the same system of reference. Therefore, these two sets can be combined to describe the evolution of scalar perturbations. Now, it is of importance to remark that the five background quantities $\theta_{(0)} = 3H$, $R_{(0)}$, $\varepsilon_{(0)}$, $\vartheta_{(0)} = 0$ and $n_{(0)}$ have precisely five first-order counterparts $\theta_{(1)}$, $R_{(1)}$, $\varepsilon_{(1)}$, $\vartheta_{(1)}$ and $n_{(1)}$. Of these five quantities, precisely three, namely $\theta$, $\varepsilon$ and $n$ (2.9) are scalars. Consequently, the evolution of cosmological density perturbations is described by the three independent scalars (2.9) and their first-order perturbations. A complicating factor is that the first-order quantities $\varepsilon_{(1)}$ and $n_{(1)}$, which are supposed to describe the energy density and the particle number density perturbations, have no physical significance, as will now be established.

A first-order perturbation to one of the scalars (2.9) transforms under a general infinitesimal coordinate transformation (2.3) as
\[
S_{(1)}(t, x) \rightarrow S_{(1)}(t, x) + \xi^{0}(t, x)\dot{S}_{(0)}(t),
\] (2.38)
where \( S_0 \) and \( S_1 \) are the background and first-order perturbation of one of the three scalars
\( S = \varepsilon, n, \theta \). In (2.38) the term \( \hat{S} = \xi^0 \dot{S}_0 \) is the so-called gauge mode. The complete set of gauge modes for the system of equations (2.29) is given by
\[
\xi_0(0) = \varepsilon(0), \quad \dot{\eta}(0) = \dot{n}(0), \quad \dot{\theta}(0) = \dot{\theta}(0), \quad \dot{R}(0) = 4H \left[ \frac{\nabla^2 \dot{\psi}}{a^2} - \frac{1}{2} R(0) \right],
\]
(2.39a)
\[
\hat{\varepsilon}(1) = \dot{\psi} \varepsilon(0), \quad \hat{n}(1) = \dot{\psi} \dot{n}(0), \quad \hat{\theta}(1) = \dot{\psi} \dot{\theta}(0), \quad \hat{R}(1) = 4H \left[ \frac{\nabla^2 \dot{\psi}}{a^2} - \frac{1}{2} R(0) \dot{\psi} \right],
\]
(2.39b)
where \( \xi^0 = \psi(x) \) in synchronous coordinates, see (2.5). The quantities (2.39) are mere coordinate artifacts, which have no physical meaning, since the gauge function \( \psi(x) \) is an arbitrary (infinitesimal) function. Equations (2.29) are invariant under coordinate transformations (2.3) combined with (2.5), i.e., the gauge modes (2.39) are solutions of the set (2.29). This property combined with the linearity of the perturbation equations, implies that a solution set \( \{ \varepsilon(1), n(1), \theta(1), R(1) \} \) can be augmented with the corresponding gauge modes (2.39) to obtain a new solution set. Therefore, the solution set \( \{ \varepsilon(1), n(1), \theta(1), R(1) \} \) has no physical significance, since the general solution of the system (2.29) can be modified by an infinitesimal coordinate transformation. This is the notorious gauge problem of cosmology.

In this article, the cosmological gauge problem has been solved as follows. The perturbation equations (2.10) have first been rewritten as (2.29) in order to isolate the scalar perturbations from the vector and tensor perturbations. The fact that the system of equations (2.29) describes exclusively the evolution of scalar perturbations has the following important consequence:

- From the background equations (2.7) and the perturbation equations (2.29) for scalar perturbations it follows that only the three independent scalars (2.9) and their first-order perturbations play a role in the evolution of density perturbations.

This fact reduces the number of possible gauge-invariant quantities substantially, since one needs to consider only the three independent scalars (2.9). Since scalar perturbations transform under the general infinitesimal transformation (2.3) according to (2.38), one can combine two independent scalars to eliminate the gauge function \( \xi^0(t, x) \). With the three independent scalars (2.9), one can make \( \binom{3}{2} = 3 \) different sets of three gauge-invariant quantities. In each of these sets exactly one gauge-invariant quantity vanishes. One of the three possible choices is
\[
\varepsilon_{gi}(1) = \varepsilon(1) - \frac{\dot{\varepsilon}(0)}{\dot{\theta}(0)} \theta(1), \quad n_{gi}(1) = n(1) - \frac{\dot{n}(0)}{\dot{\theta}(0)} \theta(1), \quad \theta_{gi}(1) = \theta(1) - \frac{\dot{\theta}(0)}{\dot{\theta}(0)} \theta(1) \equiv 0.
\]
(2.40a)
\[
\varepsilon_{gi}(1) = \varepsilon(1) - \frac{\dot{\varepsilon}(0)}{\dot{\theta}(0)} \theta(1), \quad n_{gi}(1) = n(1) - \frac{\dot{n}(0)}{\dot{\theta}(0)} \theta(1), \quad \theta_{gi}(1) = \theta(1) - \frac{\dot{\theta}(0)}{\dot{\theta}(0)} \theta(1) \equiv 0.
\]
(2.40b)
It follows from the general transformation rule (2.38) that the quantities (2.40) are invariant under the general infinitesimal transformation (2.3), i.e., they are gauge-invariant, hence the superscript ‘\( gi \)’.

In Section 4, it will be shown that the relativistic perturbation equations (2.29) become, in the non-relativistic limit, equal to the Poisson equation (4.8) of the Newtonian Theory of Gravity if and only if one uses the definitions (2.40a). This implies that \( \varepsilon_{gi}(1) \) is the real, measurable, energy density perturbation and \( n_{gi}(1) \) is the real, measurable, particle number density perturbation. The
physical interpretation of (2.40b) is that, in first-order, the global expansion \( \theta_{(0)} = 3H \) is not affected by a local density perturbation.

The quantities (2.40a) are different from the gauge-invariant quantities used in former perturbation theories [3, 5–10, 18–21]. They have two essential properties. Firstly, due to the quotients \( \varepsilon_{(0)}/\dot{\theta}_{(0)} \) and \( n_{(0)}/\dot{\theta}_{(0)} \) of time derivatives, the quantities \( \varepsilon_{(1)}^{\text{gi}} \) and \( n_{(1)}^{\text{gi}} \) are independent of the definition of time. As a consequence, the evolution of \( \varepsilon_{(1)}^{\text{gi}} \) and \( n_{(1)}^{\text{gi}} \) is only determined by their propagation equations. Secondly, the quantities (2.40a) do not contain spatial derivatives, so that unnecessary gradient terms [22] do not occur in the evolution equations (2.44).

The quantities (2.40a) are completely determined by the background equations (2.7) and their perturbed counterparts (2.29). In principle, these two sets can be used to study the evolution of density perturbations in FLRW universes. However, the set (2.29) is still too complicated, since it also admits the non-physical solutions (2.39). The aim will be a system of evolution equations for \( \varepsilon_{(1)}^{\text{gi}} \) and \( n_{(1)}^{\text{gi}} \) that do not have the gauge modes (2.39) as solution. In other words, evolution equations will be derived which are invariant under general infinitesimal coordinate transformations (2.3), as well as their solutions. The derivation of these equations will be the subject of the next subsection.

### 2.7 Evolution Equations for the Density Contrast Functions

In this section the derivation of the evolution equations is given. Firstly, it is observed that the gauge-dependent variable \( \theta_{(1)} \) is not needed in the calculations, since its gauge-invariant counterpart \( \theta_{(1)}^{\text{gi}} \), (2.40b), vanishes identically. Eliminating \( \theta_{(1)} \) from the differential equations (2.29b)–(2.29e) with the help of the (algebraic) constraint equation (2.29a) yields the set of four first-order ordinary differential equations

\[
\begin{align*}
\dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + \varepsilon_{(0)}(1 + w) \left[ \dot{n}_{(1)} + \frac{1}{2H} (\kappa \varepsilon_{(1)} + \frac{1}{2} R_{(1)}) \right] &= 0, \tag{2.41a} \\
\dot{n}_{(1)} + 3Hn_{(1)} + n_{(0)} \left[ \dot{\theta}_{(1)} + \frac{1}{2H} (\kappa \varepsilon_{(1)} + \frac{1}{2} R_{(1)}) \right] &= 0, \tag{2.41b} \\
\dot{\theta}_{(1)} + H(2 - 3\beta^2) \dot{\vartheta}_{(1)} + \frac{1}{\varepsilon_{(0)}(1 + w)} \bar{\nabla}^2 p_{(1)} &= 0, \tag{2.41c} \\
\dot{R}_{(1)} + 2HR_{(1)} - 2\kappa \varepsilon_{(0)}(1 + w) \vartheta_{(1)} + \frac{R_{(0)}}{3H} (\kappa \varepsilon_{(1)} + \frac{1}{2} R_{(1)}) &= 0, \tag{2.41d}
\end{align*}
\]

for the four quantities \( \varepsilon_{(1)} \), \( n_{(1)} \), \( \theta_{(1)} \) and \( R_{(1)} \).

Using the background equations (2.7) to eliminate all time-derivatives and the first-order constraint equation (2.29a) to eliminate \( \theta_{(1)} \), the quantities (2.40a) become

\[
\begin{align*}
\varepsilon_{(1)}^{\text{gi}} &= \frac{\varepsilon_{(1)}R_{(0)} - 3\varepsilon_{(0)}(1 + w)(2H\vartheta_{(1)} + \frac{1}{2} R_{(1)})}{R_{(0)} + 3\kappa \varepsilon_{(0)}(1 + w)}, \tag{2.42a} \\
n_{(1)}^{\text{gi}} &= n_{(1)} - \frac{3n_{(0)}(\kappa \varepsilon_{(1)} + 2H \vartheta_{(1)} + \frac{1}{2} R_{(1)})}{R_{(0)} + 3\kappa \varepsilon_{(0)}(1 + w)}. \tag{2.42b}
\end{align*}
\]

These quantities are completely determined by the background equations (2.7) and the first-order equations (2.41). In the study of the evolution of density perturbations, it is convenient not to
use \( \varepsilon_{(i)}^g \) and \( n_{(i)}^g \) directly, but instead their corresponding contrast functions \( \delta_e \) and \( \delta_n \)

\[
\delta_e(t, x) \equiv \frac{\varepsilon_{(i)}^g(t, x)}{\varepsilon_{(0)}}, \quad \delta_n(t, x) \equiv \frac{n_{(i)}^g(t, x)}{n_{(0)}},
\]

(2.43)

The system of equations (2.41) for the four independent quantities \( \varepsilon_{(i)}, n_{(i)}, \theta_{(i)} \) and \( R_{(i)} \) will now be recast, using the procedure given in the Appendix, in a new system of equations for the two independent quantities \( \delta_e \) and \( \delta_n \). In this procedure it is explicitly assumed that \( p \neq 0 \), i.e., the pressure does not vanish identically. The case \( p \to 0 \) will be considered in Section 4 on the non-relativistic limit. The final results are the evolution equations for the relative density perturbations (2.43) in FLRW universes:

\[
\ddot{\delta}_e + b_1 \dot{\delta}_e + b_2 \delta_e = b_3 \left[ \delta_n - \frac{\delta_e}{1 + w} \right],
\]

(2.44a)

\[
\frac{1}{c} \frac{d}{dt} \left[ \frac{\delta_n - \delta_e}{1 + w} \right] = \frac{3Hn_{(i)}p_n}{\varepsilon_{(i)}(1 + w)} \left[ \delta_n - \frac{\delta_e}{1 + w} \right],
\]

(2.44b)

where the coefficients \( b_1, b_2 \) and \( b_3 \) are given by

\[
b_1 = \frac{\kappa \varepsilon_{(0)}(1 + w)}{H} - 2 \frac{\dot{\beta}}{\beta} - H(2 + 6w + 3\beta^2) + R_{(0)} \left[ \frac{1}{3H} + \frac{2H(1 + 3\beta^2)}{R_{(0)} + 3 \kappa \varepsilon_{(0)}(1 + w)} \right],
\]

(2.45a)

\[
b_2 = - \frac{1}{2} \kappa \varepsilon_{(0)}(1 + w)(1 + 3w)
\]

\[
+ H^2(1 - 3w + 6\beta^2(2 + 3w)) + 6H \frac{\dot{\beta}}{\beta} \left[ w + \frac{\kappa \varepsilon_{(0)}(1 + w)}{R_{(0)} + 3 \kappa \varepsilon_{(0)}(1 + w)} \right]
\]

\[
- R_{(0)} \left( \frac{1}{2}w + \frac{H^2(1 + 6w)(1 + 3\beta^2)}{R_{(0)} + 3 \kappa \varepsilon_{(0)}(1 + w)} \right) - \beta^2 \left( \frac{\nabla^2}{a^2} - \frac{1}{2}R_{(0)} \right),
\]

(2.45b)

\[
b_3 = \left\{ \begin{array}{l}
- \frac{18H^2}{R_{(0)} + 3 \kappa \varepsilon_{(0)}(1 + w)} \varepsilon_{(0)}p_{en}(1 + w)
+ \frac{2p_n \beta}{3H} + p_n(p_e - \beta^2) + n_{(0)}p_{nn} \right\} + 
\frac{p_n}{\varepsilon_{(0)}} \left( \frac{\nabla^2}{a^2} - \frac{1}{2}R_{(0)} \right).
\]

(2.45c)

In these expressions the partial derivatives of the pressure, \( p_e \) and \( p_n \), are defined by (2.11) and \( p_{nn} \equiv \partial^2 p / \partial n^2 \) and \( p_{en} \equiv \partial^2 p / \partial \varepsilon \partial n \). In the derivation of the coefficients (2.45) it is used that the time-derivative of \( w \) is

\[
\dot{w} = 3H(1 + w)(w - \beta^2).
\]

(2.46)

This relation follows from the definitions \( w \equiv p_{(0)}/\varepsilon_{(0)} \) and \( \beta^2 \equiv \dot{\varepsilon}_{(0)}/\varepsilon_{(0)} \) using only the energy conservation law (2.7c), and is, therefore, independent of the equation of state.

The equations (2.44) have been checked (see the attached MAXIMA file) using a computer algebra system [23], as follows. Substituting the contrast functions (2.43) into equations (2.44), where \( \varepsilon_{(i)}^g \) and \( n_{(i)}^g \) are given by (2.42), and subsequently eliminating the time-derivatives of \( \varepsilon_{(0)}, n_{(0)}, H, R_{(0)} \) and \( \varepsilon_{(i)}, n_{(i)}, \theta_{(i)}, R_{(i)} \) with the help of equations (2.7) and (2.41), respectively, yields two identities for each of the two equations (2.44).

The system of equations (2.44) is equivalent to a system of three first-order differential equations, whereas the original set (2.41) is a fourth-order system. This difference is due to the
fact that the gauge modes (2.39), which are solutions of the set (2.41), are completely removed from the solution set of (2.44): one degree of freedom, namely the unknown gauge function $\xi^0(t, x)$ in (2.38), which can in no way be determined, has disappeared altogether. In other words, the general solution of the system (2.44) is not contaminated with the non-physical gauge modes $\hat{\delta}_\varepsilon \equiv \xi^0 \dot{\varepsilon}_{(0)} / \varepsilon_{(0)}$ and $\hat{\delta}_n \equiv \xi^0 \dot{n}_{(0)} / n_{(0)}$, (2.38):

$$\hat{\delta}_\varepsilon(t, x) = -3H(t)[1 + w(t)]\xi^0(t, x), \quad \hat{\delta}_n(t, x) = -3H(t)\xi^0(t, x),$$  \hspace{1cm} (2.47)

where the background equations (2.7c) and (2.7e) have been used. This implies that one can impose initial values $\delta_{\varepsilon}(t_0, x)$, $\delta_{\varepsilon}(t_0, x)$ and $\delta_n(t_0, x)$ which can, in principle, be obtained from observation and, subsequently, calculate the evolution of the physical, i.e., measurable, density contrast functions $\delta_{\varepsilon}(t, x)$ and $\delta_n(t, x)$.

The background equations (2.7) and the new perturbation equations (2.44) constitute a set of equations which enables one to study the evolution of small fluctuations in the energy density $\delta_{\varepsilon}$ and the particle number density $\delta_n$ in an open, flat or closed FLRW universe with $\Lambda \neq 0$ and filled with a perfect fluid described by a non-barotropic equation of state for the pressure $p = p(n, \varepsilon)$.

Without solving equations (2.44), one can draw the following conclusion from equation (2.44b): perturbations in the particle number density are coupled by gravitation to perturbations in the total energy density if $p_n \equiv (\partial p / \partial n)_\varepsilon \leq 0$, or, equivalently, $p_\varepsilon \equiv (\partial p / \partial \varepsilon)_n \geq \beta^2$, see (2.30). This coupling is independent of the nature of the particles, i.e., it holds true for all kinds of matter which interact through gravitation, in particular ordinary matter and Cold Dark Matter.

### 3 Thermodynamics

In this section expressions for the gauge-invariant pressure and temperature perturbations are derived. It is shown that density perturbations are adiabatic if and only if the particle number density does not contribute to the pressure.

#### 3.1 Gauge-invariant Pressure and Temperature Perturbations

The gauge-invariant pressure and temperature perturbations, which are needed in the forthcoming sections, will now be derived.

From the equation of state for the pressure $p = p(n, \varepsilon)$, one has $p_{(0)} = p_n \dot{n}_{(0)} + p_\varepsilon \dot{\varepsilon}_{(0)}$. Multiplying both sides of this expression by $\theta_{(1)} / \theta_{(0)}$ and subtracting the result from $p_{(1)}$ given by (2.11), one gets, using also (2.40a),

$$p_{(1)} - \frac{\dot{p}_{(0)}}{\theta_{(0)}} \theta_{(1)} = p_n \theta_{(1)}^{gi} + p_\varepsilon \varepsilon_{(1)}^{gi}. \quad (3.1)$$

Hence, the quantity defined by

$$p_{(1)}^{gi} \equiv p_{(1)} - \frac{\dot{p}_{(0)}}{\theta_{(0)}} \theta_{(1)}^{gi}, \quad (3.2)$$

is the gauge-invariant pressure perturbation. Combining (3.1) and (3.2) and eliminating $p_\varepsilon$ with the help of (2.30), one arrives at

$$p_{(1)}^{gi} = \beta^2 \varepsilon_{(0)} \delta_{\varepsilon} + n_{(0)} p_n \left[ \delta_n - \frac{\delta_{\varepsilon}}{1 + w} \right], \quad (3.3)$$
where also (2.43) has been used. The first term in this expression is the *adiabatic* part of the pressure perturbation and the second term is the *diabatic* part, as follows from (3.11).

From the equation of state (2.1) for the energy density $\varepsilon = \varepsilon(n, T)$ it follows that

$$\dot{\varepsilon}_{(0)} = \left( \frac{\partial \varepsilon}{\partial n} \right)_T \dot{n}_{(0)} + \left( \frac{\partial \varepsilon}{\partial T} \right)_n \dot{T}_{(0)}, \quad \varepsilon_{(1)} = \left( \frac{\partial \varepsilon}{\partial n} \right)_T n_{(1)} + \left( \frac{\partial \varepsilon}{\partial T} \right)_n T_{(1)}. \quad (3.4)$$

Multiplying $\dot{\varepsilon}_{(0)}$ by $\theta_{(1)} / \dot{\theta}_{(0)}$ and subtracting the result from $\varepsilon_{(1)}$, one finds, using (2.40a),

$$\varepsilon_{g1}^{(1)} = \left( \frac{\partial \varepsilon}{\partial n} \right)_T n_{g1}^{(1)} + \left. \left( \frac{\partial \varepsilon}{\partial T} \right)_n \right|_{T_{(1)}}^{T_{(0)}} \left[ T_{(1)} - \frac{\dot{T}_{(0)}}{\dot{\theta}_{(0)}} \theta_{(1)} \right], \quad (3.5)$$

implying that the quantity defined by

$$T_{(1)}^{g1} \equiv T_{(1)} - \frac{\dot{T}_{(0)}}{\dot{\theta}_{(0)}} \theta_{(1)}, \quad (3.6)$$

is the gauge-invariant temperature perturbation. The expressions (3.2) and (3.6) are both of the form (2.40).

### 3.2 Diabatic Density Perturbations

In this section equations (2.44) will be linked to thermodynamics and it will be shown that, in general, density perturbations evolve diabatically, i.e., they exchange heat with their environment during their evolution.

The combined First and Second Law of Thermodynamics is given by

$$dE = TdS - pdV + \mu dN, \quad (3.7)$$

where $E$, $S$ and $N$ are the energy, the entropy and the number of particles of a system with volume $V$ and pressure $p$, and where $\mu$, the thermal — or chemical — potential, is the energy needed to add one particle to the system. In terms of the particle number density $n = N/V$, the energy per particle $E/N = \varepsilon/n$ and the entropy per particle $s = S/N$ the law (3.7) can be rewritten as

$$d \left( \frac{\varepsilon}{n} N \right) = Td(sN) - pd \left( \frac{N}{n} \right) + \mu dN, \quad (3.8)$$

where $\varepsilon$ is the energy density. The system is *extensive*, i.e., $S(\alpha E, \alpha V, \alpha N) = \alpha S(E, V, N)$, implying that the entropy of the gas is $S = (E + pV - \mu N)/T$. Dividing this relation by $N$ one gets the so-called Euler relation

$$\mu = \frac{\varepsilon + p}{n} - Ts. \quad (3.9)$$

Eliminating $\mu$ in (3.8) with the help of (3.9), one finds that the combined First and Second Law of Thermodynamics (3.7) can be cast in a form without $\mu$ and $N$, i.e.,

$$Tds = d \left( \frac{\varepsilon}{n} \right) + pd \left( \frac{1}{n} \right). \quad (3.10)$$

From the background equations (2.7) and the thermodynamic law (3.10) it follows that $\dot{s}_{(0)} = 0$, implying that, in zeroth-order, the expansion of the universe takes place without generating entropy. Using (2.38) one finds that $s_{(1)} = s_{g1}^{(1)}$ is automatically gauge-invariant.
The thermodynamic relation (3.10) can, using (2.43), be rewritten as
\[ T_{(0)}s_{(1)}^{\text{gi}} = -\frac{\varepsilon_{(0)}(1 + w)}{n_{(0)}} \left[ \delta_n - \frac{\delta \varepsilon}{1 + w} \right]. \] (3.11)

Thus, the right-hand side of (2.44a) is related to local perturbations in the entropy, and (2.44b) can be considered as an evolution equation for entropy perturbations.

The condition for adiabaticity will now be derived. Adiabatic perturbations do not exchange heat with their environment, i.e., \( T_{(0)}s_{(1)}^{\text{gi}} = 0 \). This implies with (3.11) that \((1 + w)\delta_n - \delta \varepsilon = 0\). Multiplying this expression by \( 3H\varepsilon_{(0)}n_{(0)} \) and substituting (2.43) one finds from the background conservation laws (2.7c) and (2.7e) that the adiabatic condition \( s_{(1)}^{\text{gi}} = 0 \) reads \( \dot{n}_{(0)}\varepsilon_{(1)} - \dot{\varepsilon}_{(0)}n_{(1)} = 0 \). Using that \( \varepsilon = \varepsilon(n, T) \) the latter expression becomes
\[ \left( \frac{\partial \varepsilon}{\partial T} \right)_n \left[ \dot{n}_{(0)}T_{(1)}^{\text{gi}} - n_{(1)}^{\text{gi}}\dot{T}_{(0)} \right] = 0. \] (3.12)

Since \( n \) and \( T \) are independent quantities and since in a non-static universe one has \( \dot{n}_{(0)} \neq 0 \) and \( \dot{T}_{(0)} \neq 0 \), the adiabatic condition (3.12) is satisfied if, and only if,
\[ \left( \frac{\partial \varepsilon}{\partial T} \right)_n = 0, \] (3.13)

implying that \( \varepsilon = \varepsilon(n) \). In particular, in the non-relativistic limit, where \( \varepsilon = nmc^2 \) and \( p = 0 \), density perturbations are adiabatic. This is in accordance with the fact that in the non-relativistic limit, which will be elaborated in the next section, density perturbations do not evolve, i.e., \( \dot{\varepsilon}_{(1)} = 0 \). In all other cases where \( p = p(n, \varepsilon) \) local density perturbations evolve diabatically.

Finally, in the limiting case that the particle number density does not contribute to the pressure, i.e., \( p_n \approx 0 \), it follows from (3.3) that the pressure perturbation is adiabatic. In this case the equation of state is barotropic, i.e., \( p \approx p(\varepsilon) \), implying that the coefficient \( b_3 \), (2.45c), vanishes so that the evolution equation for energy density perturbations (2.44a) is homogeneous and is decoupled from the evolution equation for entropy perturbations (2.44b). Consequently, for barotropic equations of state density perturbations evolve adiabatically.

4 Non-relativistic Limit

In Section 2.6 it has been shown that the two gauge-invariant quantities \( \varepsilon_{(1)}^{\text{gi}} \) and \( n_{(1)}^{\text{gi}} \) are unique. It will be demonstrated that in the non-relativistic limit equations (2.29) combined with (2.40) reduce to the results (4.8) and (4.9) of the Newtonian Theory of Gravity and that the quantities \( \varepsilon_{(1)}^{\text{gi}} \) and \( n_{(1)}^{\text{gi}} \) become equal to their Newtonian counterparts.

The non-relativistic limit is defined by three requirements, Carroll [24], page 153:

- The gravitational field should be weak, i.e., can be considered as a perturbation of flat space.
- The particles are moving slowly with respect to the speed of light.
- The gravitational field of a density perturbation should be static, i.e., it does not change with time.
This definition of the non-relativistic limit, which is essential to put an accurate physical interpretation on the quantities $\varepsilon^{(1)}_1$ and $n^{(1)}_1$, has not been used in former perturbation theories [3, 5–10, 18–21] to explain the meaning of the gauge-invariant quantities.

For first-order cosmological perturbations the gravitational field is already weak. In order to meet the first requirement, a flat ($R_{(0)} = 0$) FLRW universe is considered. Using (2.19) and the fact that spatial covariant derivatives become, in flat three-space, ordinary derivatives with respect to the spatial coordinates, the local perturbation to the spatial curvature, (2.15), reduces for a flat FLRW universe to

$$R_{(1)} = \frac{4}{c^2} \phi |_{k} = -\frac{4}{c^2} \nabla^2 \phi,$$

where $\nabla^2$ is the usual Laplace operator. Substituting this expression into the perturbation equations (2.29), one gets

$$H(\theta_{(1)} - \theta_{(1)}) = -\frac{1}{c^2} \frac{\nabla^2 \dot{\phi}}{a^2} + \frac{4\pi G N}{c^4} \left[ \varepsilon^{(1)}_1 + \frac{\varepsilon^{(0)}_1}{\theta^{(0)}_1} \right],$$

$$\frac{\nabla^2 \dot{\phi}}{a^2} = -\frac{4\pi G N}{c^2} \varepsilon^{(0)}_1 (1 + w) \theta_{(1)},$$

$$\dot{\varepsilon}_{(1)} = -3H(\varepsilon_{(1)} + p_{(1)}) - \varepsilon^{(0)}_1 (1 + w) \theta_{(1)},$$

$$\dot{\theta}_{(1)} = -H(2 - 3\beta^2) \dot{\theta}_{(1)} - \frac{1}{\varepsilon^{(0)}_1 (1 + w)} \frac{\nabla^2 p_{(1)}}{a^2},$$

$$\dot{n}_{(1)} = -3Hn_{(1)} - n^{(0)}_1 \theta_{(1)},$$

where (2.40a) has been used to eliminate $\varepsilon^{(1)}_1$ from the constraint equation (2.29a).

Next, the second requirement will be implemented. Since the spatial parts $u^{(1)}_i$ of the fluid four-velocity are gauge-dependent with physical components and non-physical gauge parts, the second requirement must be defined by

$$u^{(1)}_i \text{physical} \equiv \frac{U^{(1)}_i \text{physical}}{c} \rightarrow 0,$$

i.e., the physical parts of the spatial parts of the fluid four-velocity are negligible with respect to the speed of light. In this limit, the mean kinetic energy per particle $\frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} k_B T \rightarrow 0$ is very small compared to the rest energy $mc^2$ per particle. This implies that the pressure $p = nk_B T \rightarrow 0$ ($n \neq 0$) is vanishingly small with respect to the rest energy density $nmc^2$. Therefore, one must take the limits $p_{(0)} \rightarrow 0$ in the background and $p_{(1)}^{(1)} \rightarrow 0$ in the perturbed universe to arrive at the non-relativistic limit. With (3.2) it follows that also $p_{(1)} \rightarrow 0$. Substituting $p_{(0)} = 0$ and $p_{(1)} = 0$ into the momentum conservation laws (2.10e) yields, using also the background equation (2.7c) with $w \equiv p_{(0)}/\varepsilon_{(0)} \rightarrow 0$,

$$\dot{u}^{i}_{(1)} = -2Hu^{i}_{(1)}.$$

Since the physical parts of $u^{i}_{(1)}$ vanish in the non-relativistic limit, the general solutions of equations (4.4) are exactly equal to the gauge modes, see (2.39b),

$$\ddot{u}^{i}_{(1)}(t, x) = -\frac{1}{a^2(t)} g^{ik}(x) \partial_k \psi(x),$$

\text{(recall that $u^{i}_{(1)}$ is the irrotational part of the three-space fluid velocity. The rotational part of $u^{i}_{(1)}$ is not coupled to density perturbations and need, therefore, not be considered. See Section 2.4.)}
where it is used that \( H \equiv \dot{a}/a \). Thus, in the limit (4.3) one is left with the gauge modes (4.5) only. Consequently, one may, without losing any physical information, put the gauge modes \( \vec{u}_{(i)} \) equal to zero, implying that \( \partial_k \psi = 0 \), so that \( \psi = C \) is an arbitrary constant in the non-relativistic limit. Substituting \( \psi = C \) into (2.5) one finds that the relativistic transformation (2.3) between synchronous coordinates reduces in the limit (4.3) to the (infinitesimal) transformation

\[
x^0 \to x^0 - C, \quad x^i \to x^i - \chi^i(x),
\]

where \( C \) is an arbitrary constant and \( \chi^i(x) \) are three arbitrary functions of the spatial coordinates. Thus, in the non-relativistic limit time and space transformations are decoupled: time coordinates may be shifted and spatial coordinates may be chosen arbitrarily. The residual gauge freedoms \( C \) and \( \chi^i(x) \) in the non-relativistic limit express the fact that the Newtonian Theory of Gravity is invariant under the ‘gauge transformation’ (4.6).

Substituting \( \vartheta_{(1)} = 0 \) and \( p = 0 \) (i.e., \( p(0) = 0 \) and \( p_{(1)} = 0 \)) into the system (4.2), one arrives at the Einstein equations and conservation laws in the non-relativistic limit:

\[
\begin{align*}
\nabla^2 \phi &= \frac{4\pi G N}{c^2} a^2 \varepsilon^{\text{gi}}_{(1)}, \\
\nabla^2 \dot{\phi} &= 0, \\
\dot{\varepsilon}_{(1)} &= -3H \varepsilon_{(1)} - \varepsilon(0) \vartheta_{(1)}, \\
\dot{n}_{(1)} &= -3H n_{(1)} - n(0) \vartheta_{(1)}.
\end{align*}
\]  

(4.7a)–(4.7d)

The constraint equation (4.7a) can be found by subtracting \( \frac{1}{c} \dot{\vartheta}_{(1)} \dot{H} \) times the time-derivative of the background constraint equation (2.7a) with \( R_{(0)} = 0 \) from the constraint equation (4.2a) and using that \( \vartheta(0) = 3H \). Note that the cosmological constant \( \Lambda \) need not be zero.

Since \( \vartheta_{(1)} = 0 \) there is no fluid flow so that density perturbations do not evolve. This implies the basic fact of the Newtonian Theory of Gravity, namely that the gravitational field is static (4.7b). Consequently, \( a^2(t)\varepsilon_{(1)}^{\text{gi}}(t, x) \) in (4.7a) should be replaced by \( a^2(t_0)\varepsilon_{(1)}^{\text{gi}}(x) \). Defining the potential \( \varphi(x) \equiv \phi(x)/a^2(t_0) \), equations (4.7a) and (4.7b) imply

\[
\nabla^2 \varphi(x) = 4\pi G N \rho_{(1)}(x), \quad \rho_{(1)}(x) \equiv \frac{\varepsilon^{\text{gi}}_{(1)}(x)}{c^2},
\]

(4.8)

which is the Poisson equation of the Newtonian Theory of Gravity. With (4.8) the third requirement for the non-relativistic limit, i.e., a static gravitational field, has been satisfied.

The expression (2.42a) reduces in the non-relativistic limit to \( \varepsilon^{\text{gi}}_{(1)} = -R_{(1)}/(2\kappa) \), which is, with (4.1) and (4.7b), equivalent to the Poisson equation (4.8). Using that \( \varepsilon(0) = n(0)mc^2 \) and \( \varepsilon_{(1)} = n_{(1)}mc^2 \), expression (2.42b) reduces in the non-relativistic limit to the well-known result

\[
\begin{align*}
n^{\text{gi}}_{(1)}(x) &= \frac{\varepsilon^{\text{gi}}_{(1)}(x)}{mc^2}, \\
\rho_{(1)}(x) &= mn^{\text{gi}}_{(1)}(x).
\end{align*}
\]

(4.9)

where it has been used that in the non-relativistic limit \( R_{(1)} = -2\kappa \varepsilon^{\text{gi}}_{(1)} \).

The universe is in the non-relativistic limit not static, since \( H \neq 0 \) and \( \dot{H} \neq 0 \), as follows from the background equations (2.7) with \( w = 0 \) and \( R_{(0)} = 0 \). Equations (4.7) imply that in the non-relativistic limit a local density perturbation does not follow the global expansion of
the universe and the system of reference has become co-moving. Since density perturbations do not evolve in the non-relativistic limit, they are essentially adiabatic, in accordance with the conclusion at the end of Section 3.2.

The gauge modes $\dot{\epsilon}_{(1)}$, $\dot{n}_{(1)}$ and $\dot{\theta}_{(1)}$ (2.39a) do not vanish, since $\psi = C$ is an arbitrary constant which cannot be fixed. As a consequence, the gauge-dependent quantities $\varepsilon_{(1)}$, $n_{(1)}$ and $\theta_{(1)}$ do not become gauge-invariant in the non-relativistic limit. In fact, the gauge modes $\dot{\epsilon}_{(1)}$, $\dot{n}_{(1)}$ and $\dot{\theta}_{(1)}$ are solutions of (4.7c) and (4.7d). Since these equations are decoupled from the physical equations (4.7a) and (4.7b) they are not part of the Newtonian Theory of Gravity and need not be considered. Thus, in the non-relativistic limit, one is left with the Poisson equation (4.8) and the well-known energy-mass relation (4.9).

Finally, the potential $\zeta$ which occurs by (2.19) in $R_{(1)}$, (2.15), and $\theta_{(1)}$, (2.16), in the general relativistic case, drops from the evolution equations in the non-relativistic limit. Consequently, one is left with one potential $\varphi(x)$ only.

It has been shown that equations (2.29) combined with the unique gauge-invariant quantities (2.40) reduce in the non-relativistic limit to the Newtonian results (4.8) and (4.9). Consequently, $\varepsilon_{(1)}^{\tilde{g}}$ and $n_{(1)}^{\tilde{g}}$ are the real, physical, i.e., measurable, perturbations to the energy density and particle number density, respectively.

A Derivation of the Evolution Equations using Computer Algebra

In this Appendix the perturbation equations (2.44) of the main text will be derived from the basic perturbation equations (2.41) and the definitions (2.43). This will be done by first deriving the evolution equations for the gauge-invariant quantities $\varepsilon_{(1)}^{\tilde{g}}$ and $n_{(1)}^{\tilde{g}}$ (2.40a), or, equivalently, (2.42):

\begin{equation}
\frac{1}{c^2} \frac{d}{dt} \left( n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)} \varepsilon_{(1)} \right) = -3H \left( 1 - \frac{n_{(0)}p_n}{\varepsilon_{(0)}(1+w)} \right) \left( \frac{n_{(1)}}{\varepsilon_{(0)}} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)} \varepsilon_{(1)} \right). \tag{A.1b}
\end{equation}

The coefficients $a_1$, $a_2$ and $a_3$ occurring in equation (A.1a) are given by

\begin{equation}
a_1 = \frac{\kappa \varepsilon_{(0)}(1+w)}{H} - 2\beta \frac{\dot{\beta}}{\beta} + H(4-3\beta^2) + R_{(0)} \left( \frac{1}{3H} + \frac{2H(1+3\beta^2)}{R_{(0)} + 3\kappa \varepsilon_{(0)}(1+w)} \right), \tag{A.2a}
\end{equation}

\begin{equation}
a_2 = \kappa \varepsilon_{(0)}(1+w) - 4H \frac{\dot{\beta}}{\beta} + 2H^2(2-3\beta^2)
+ R_{(0)} \left( \frac{1}{2} + \frac{5H^2(1+3\beta^2) - 2H \dot{\beta}}{R_{(0)} + 3\kappa \varepsilon_{(0)}(1+w)} \right) - \beta^2 \left( \frac{\nabla^2}{a^2} - \frac{1}{2} R_{(0)} \right), \tag{A.2b}
\end{equation}

\begin{equation}
a_3 = \begin{cases}
-18H^2 / \left( R_{(0)} + 3\kappa \varepsilon_{(0)}(1+w) \right) \left( \varepsilon_{(0)}p_{en}(1+w) \\
+ 2p_n \beta + p_n(p_e - \beta^2) + n_{(0)}p_{nn} \right) + p_n \left( \frac{\nabla^2}{a^2} - \frac{1}{2} R_{(0)} \right), \tag{A.2c}
\end{cases}
\end{equation}

}\]
\[
\begin{array}{|c|c|c|c|c|}
\hline
 & \dot{\varepsilon}_{(1)} & \dot{n}_{(1)} & \dot{\vartheta}_{(1)} & \dot{R}_{(1)} \\
\hline
\varepsilon_{(1)} & 3H(1+p_\varepsilon) + \frac{\kappa \varepsilon_{(0)}(1+w)}{2H} & 3Hp_n & \varepsilon_{(0)}(1+w) & \frac{\varepsilon_{(0)}(1+w)}{4H} \\
\dot{n}_{(1)} & \frac{\kappa n_{(0)}}{2H} & 3H & n_{(0)} & \frac{n_{(0)}}{4H} \\
\dot{\vartheta}_{(1)} & \frac{p_\varepsilon}{\varepsilon_{(0)}(1+w) a^2} & \frac{p_n}{\varepsilon_{(0)}(1+w) a^2} & H(2-3\beta^2) & 0 \\
\dot{R}_{(1)} & \frac{\kappa R_{(0)}}{3H} & 0 & -2\kappa \varepsilon_{(0)}(1+w) & 2H + \frac{R_{(0)}}{6H} \\
\varepsilon_{(1)} & -\frac{R_{(0)}}{R_{(0)} + 3\kappa \varepsilon_{(0)}(1+w)} & 0 & \frac{6\varepsilon_{(0)}H(1+w)}{R_{(0)} + 3\kappa \varepsilon_{(0)}(1+w)} & \frac{6\varepsilon_{(0)}(1+w)}{R_{(0)} + 3\kappa \varepsilon_{(0)}(1+w)} \\
\hline
\end{array}
\]

Table 1: The coefficients \(\alpha_{ij}\) figuring in the equations (A.3).

In calculating the coefficients \(a_1, a_2\) and \(a_3\), (A.2), it is used that the time derivative of the quotient \(w \equiv \frac{p_{(0)}}{\varepsilon_{(0)}}\) is given by (2.46). Moreover, it is convenient not to expand the function \(\beta^2 \equiv \frac{\dot{p}_{(0)}}{\dot{\varepsilon}_{(0)}}\) since this will considerably complicate the expressions for the coefficients \(a_1, a_2\) and \(a_3\).

A.1 Derivation of the Evolution Equation for the Energy Density Perturbation

In order to derive equation (A.1a), the system (2.41) and expression (2.42a) will be rewritten, using (2.11), in the form

\[
\begin{align*}
\dot{\varepsilon}_{(1)} &= \alpha_{11}\varepsilon_{(1)} + \alpha_{12}n_{(1)} + \alpha_{13}\vartheta_{(1)} + \alpha_{14}R_{(1)} = 0, \\
\dot{n}_{(1)} &= \alpha_{21}\varepsilon_{(1)} + \alpha_{22}n_{(1)} + \alpha_{23}\vartheta_{(1)} + \alpha_{24}R_{(1)} = 0, \\
\dot{\vartheta}_{(1)} &= \alpha_{31}\varepsilon_{(1)} + \alpha_{32}n_{(1)} + \alpha_{33}\vartheta_{(1)} + \alpha_{34}R_{(1)} = 0, \\
\dot{R}_{(1)} &= \alpha_{41}\varepsilon_{(1)} + \alpha_{42}n_{(1)} + \alpha_{43}\vartheta_{(1)} + \alpha_{44}R_{(1)} = 0, \\
\varepsilon_{(1)} &= \alpha_{51}\varepsilon_{(1)} + \alpha_{52}n_{(1)} + \alpha_{53}\vartheta_{(1)} + \alpha_{54}R_{(1)} = 0,
\end{align*}
\]

where the coefficients \(\alpha_{ij}\) are given in Table 1.

**Step 1.** First the quantity \(R_{(1)}\) will be eliminated from equations (A.3). Differentiating equation (A.3e) with respect to time and eliminating the time derivatives \(\dot{\varepsilon}_{(1)}, \dot{n}_{(1)}, \dot{\vartheta}_{(1)}\) and \(\dot{R}_{(1)}\) with the help of equations (A.3a)--(A.3d), one arrives at the equation

\[
\dot{\varepsilon}_{(1)} + p_1\varepsilon_{(1)} + p_2n_{(1)} + p_3\vartheta_{(1)} + p_4R_{(1)} = 0,
\]

where the coefficients \(p_1, \ldots, p_4\) are given by

\[
p_i = \dot{\alpha}_{5i} - \alpha_{51}\alpha_{1i} - \alpha_{52}\alpha_{2i} - \alpha_{53}\alpha_{3i} - \alpha_{54}\alpha_{4i}.
\]
From equation (A.4) it follows that

\[ R_{(1)} = -\frac{1}{p_4} \dot{\varepsilon}_{(1)}^{i} - \frac{p_1}{p_4} \varepsilon_{(1)} - \frac{p_2}{p_4} n_{(1)} - \frac{p_3}{p_4} \dot{\vartheta}_{(1)}. \]  

(A.6)

In this way the quantity \( R_{(1)} \) has been expressed as a linear combination of the quantities \( \varepsilon_{(1)}^{i} \), \( n_{(1)} \) and \( \vartheta_{(1)} \). Upon replacing \( R_{(1)} \) in equations (A.3) by the right-hand side of (A.6), one arrives at the system of equations

\[ \dot{\varepsilon}_{(1)} + q_1 \varepsilon_{(1)}^{i} + \gamma_{11} \varepsilon_{(1)} + \gamma_{12} n_{(1)} + \gamma_{13} \dot{\vartheta}_{(1)} = 0, \]  

(A.7a)

\[ \dot{n}_{(1)} + q_2 \varepsilon_{(1)}^{i} + \gamma_{21} \varepsilon_{(1)} + \gamma_{22} n_{(1)} + \gamma_{23} \dot{\vartheta}_{(1)} = 0, \]  

(A.7b)

\[ \dot{\vartheta}_{(1)} + q_3 \varepsilon_{(1)}^{i} + \gamma_{31} \varepsilon_{(1)} + \gamma_{32} n_{(1)} + \gamma_{33} \dot{\vartheta}_{(1)} = 0, \]  

(A.7c)

\[ \dot{R}_{(1)} + q_4 \varepsilon_{(1)}^{i} + \gamma_{41} \varepsilon_{(1)} + \gamma_{42} n_{(1)} + \gamma_{43} \dot{\vartheta}_{(1)} = 0, \]  

(A.7d)

\[ \varepsilon_{(1)}^{i} + q_5 \varepsilon_{(1)}^{i} + \gamma_{51} \varepsilon_{(1)} + \gamma_{52} n_{(1)} + \gamma_{53} \dot{\vartheta}_{(1)} = 0, \]  

(A.7e)

where the coefficients \( q_i \) and \( \gamma_{ij} \) are given by

\[ q_i = -\frac{\alpha_{i4}}{p_4}, \quad \gamma_{ij} = \alpha_{ij} + q_i p_j. \]  

(A.8)

It has now been achieved that the quantity \( R_{(1)} \) occurs explicitly only in equation (A.7d), whereas \( R_{(1)} \) occurs implicitly in the remaining equations. Therefore, equation (A.7d) is not needed anymore. Equations (A.7a)-(A.7c) and (A.7e) are four ordinary differential equations for the four unknown quantities \( \varepsilon_{(1)}, n_{(1)}, \dot{\vartheta}_{(1)} \) and \( \varepsilon_{(1)}^{i} \).

**Step 2.** In the same way as in Step 1, the explicit occurrence of the quantity \( \dot{\vartheta}_{(1)} \) will be eliminated from the system of equations (A.7). Differentiating equation (A.7e) with respect to time and eliminating the time derivatives \( \dot{\varepsilon}_{(1)}, \dot{n}_{(1)} \) and \( \dot{\vartheta}_{(1)} \) with the help of equations (A.7a)-(A.7c), one arrives at

\[ q_5 \varepsilon_{(1)}^{i} + r \varepsilon_{(1)}^{i} + s_1 \varepsilon_{(1)} + s_2 n_{(1)} + s_3 \dot{\vartheta}_{(1)} = 0, \]  

(A.9)

where the coefficients \( s_i \) and \( r \) are given by

\[ s_i = \gamma_{5i} - \gamma_{51} \gamma_{1i} - \gamma_{52} \gamma_{2i} - \gamma_{53} \gamma_{3i}, \]  

(A.10a)

\[ r = 1 + \dot{\gamma}_5 - \gamma_{51} q_1 - \gamma_{52} q_2 - \gamma_{53} q_3. \]  

(A.10b)

From equation (A.9) it follows that

\[ \dot{\vartheta}_{(1)} = -\frac{q_5}{s_3} \varepsilon_{(1)}^{i} - \frac{r}{s_3} \varepsilon_{(1)}^{i} - \frac{s_1}{s_3} \varepsilon_{(1)} - \frac{s_2}{s_3} n_{(1)}. \]  

(A.11)

In this way the quantity \( \dot{\vartheta}_{(1)} \) is expressed as a linear combination of the quantities \( \varepsilon_{(1)}^{i}, \varepsilon_{(1)}^{i}, \varepsilon_{(1)} \) and \( n_{(1)} \). Upon replacing \( \dot{\vartheta}_{(1)} \) in equations (A.7) by the right-hand side of (A.11), one arrives at
the system of equations

\[
\begin{align*}
\dot{\varepsilon}_{(1)} &= -\gamma_{13} q_5 \varepsilon_{(1)}^{21} + \left( q_1 - \gamma_{13} \frac{r}{s_3} \right) \varepsilon_{(1)}^{21} + \left( \gamma_{11} - \gamma_{13} \frac{s_1}{s_3} \right) \varepsilon_{(1)} + \left( \gamma_{12} - \gamma_{13} \frac{s_2}{s_3} \right) n_{(1)} = 0, \\
\dot{n}_{(1)} &= -\gamma_{23} q_5 \varepsilon_{(1)}^{21} + \left( q_2 - \gamma_{23} \frac{r}{s_3} \right) \varepsilon_{(1)}^{21} + \left( \gamma_{21} - \gamma_{23} \frac{s_1}{s_3} \right) \varepsilon_{(1)} + \left( \gamma_{22} - \gamma_{23} \frac{s_2}{s_3} \right) n_{(1)} = 0, \\
\dot{\gamma}_{(1)} &= -\gamma_{33} q_5 \varepsilon_{(1)}^{21} + \left( q_3 - \gamma_{33} \frac{r}{s_3} \right) \varepsilon_{(1)}^{21} + \left( \gamma_{31} - \gamma_{33} \frac{s_1}{s_3} \right) \varepsilon_{(1)} + \left( \gamma_{32} - \gamma_{33} \frac{s_2}{s_3} \right) n_{(1)} = 0, \\
\dot{R}_{(1)} &= -\gamma_{43} q_5 \varepsilon_{(1)}^{21} + \left( q_4 - \gamma_{43} \frac{r}{s_3} \right) \varepsilon_{(1)}^{21} + \left( \gamma_{41} - \gamma_{43} \frac{s_1}{s_3} \right) \varepsilon_{(1)} + \left( \gamma_{42} - \gamma_{43} \frac{s_2}{s_3} \right) n_{(1)} = 0.
\end{align*}
\]

It has now been achieved that the quantities \(\dot{\vartheta}_{(1)}\) and \(R_{(1)}\) occur explicitly only in equations (A.12c) and (A.12d), whereas they occur implicitly in the remaining equations. Therefore, equations (A.12c) and (A.12d) are not needed anymore. Equations (A.12a), (A.12b) and (A.12e) are three ordinary differential equations for the three unknown quantities \(\varepsilon_{(1)}, n_{(1)}\) and \(\varepsilon_{(1)}^{21}\).

**Step 3.** At first sight, the next steps would be to eliminate, successively, the quantities \(\varepsilon_{(1)}\) and \(n_{(1)}\) from equation (A.12e) with the help of equations (A.12a) and (A.12b). One would then end up with a fourth-order differential equation for the unknown quantity \(\varepsilon_{(1)}^{21}\). This, however, is impossible, since the gauge-dependent quantities \(\varepsilon_{(1)}\) and \(n_{(1)}\) do not occur explicitly in equation (A.12e), as will now be shown. Firstly, it is observed that equation (A.12e) can be rewritten as

\[
\varepsilon_{(1)}^{21} + a_1 \varepsilon_{(1)} + a_2 \varepsilon_{(1)}^{21} = a_3 \left( n_{(1)} + \frac{\gamma_{51} s_3 - \gamma_{53} s_1}{\gamma_{52} s_3 - \gamma_{53} s_2} \varepsilon_{(1)} \right),
\]

where the coefficients \(a_1, a_2\) and \(a_3\) are given by

\[
a_1 = -\frac{s_3}{\gamma_{53}} + \frac{r}{q_5}, \quad a_2 = -\frac{s_3}{\gamma_{53} q_5}, \quad a_3 = \frac{\gamma_{52} s_3}{\gamma_{53} q_5} - \frac{s_2}{q_5}.
\]

These are precisely the coefficients (A.2). Secondly, one finds

\[
\frac{\gamma_{51} s_3 - \gamma_{53} s_1}{\gamma_{52} s_3 - \gamma_{53} s_2} = -\frac{n_{(0)}}{\varepsilon_{(0)}(1+w)}.
\]

Finally, using the definitions (2.40a) and the conservation laws (2.7c) and (2.7e), it is found that

\[
n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)} \varepsilon_{(1)} = n_{(1)}^{gi} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)} \varepsilon_{(1)}^{gi}.
\]

Thus, the right-hand side of (A.13) does not explicitly contain the gauge-dependent quantities \(\varepsilon_{(1)}\) and \(n_{(1)}\). With the help of expression (A.16) one can rewrite equation (A.13) in the form (A.1a).

The derivation of the coefficients (A.2) from (A.14) and the proof of the equality (A.15) is straightforward, but extremely complicated. The computer algebra system Maxima [23] has been used to perform this algebraic task, see the attached file.
A.2 Derivation of the Evolution Equation for the Entropy Perturbation

The basic set of equations (2.41) from which the evolution equations are derived is of fourth-order. From this system a second-order equation (A.1a) for \( \varepsilon_{(1)}^{(1)} \) has been extracted. Therefore, the remaining system from which an evolution equation for \( n_{(1)}^{(1)} \) can be derived is at most of second order. Since gauge-invariant quantities \( \varepsilon_{(1)}^{(1)} \) and \( n_{(1)}^{(1)} \) have been used, one degree of freedom, namely the gauge function \( \xi^0(t, \mathbf{x}) \) in (2.38) has disappeared. As a consequence, only a first-order evolution equation for \( n_{(1)}^{(1)} \) can be derived. Instead of deriving an equation for \( n_{(1)}^{(1)} \), an evolution equation (A.1b) for the entropy perturbation, which contains \( n_{(1)}^{(1)} \), will be derived.

From the combined First and Second Law of Thermodynamics (3.10) it follows that

\[
T_{(0)} s_{(1)} = -\frac{\varepsilon_{(0)}(1 + w)}{n_{(0)}^2} \left[ n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1 + w)} \varepsilon_{(1)} \right],
\]

(A.17)

where the right-hand side is gauge-invariant by virtue of (A.16), so that \( s_{(1)} = s_{(1)}^{gi} \) is gauge-invariant, in accordance with the remark below (3.10). Differentiating the term between square brackets in (A.17) with respect to time and using the background equations (2.7c) and (2.7e), the first-order equations (2.41a) and (2.41b) and the definitions \( w = p_{(0)}/\varepsilon_{(0)} \) and \( \beta^2 = \dot{p}_{(0)}/\dot{\varepsilon}_{(0)} \), one finds

\[
\frac{1}{c} \frac{d}{dt} \left( n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1 + w)} \varepsilon_{(1)} \right) = -3H \left( 1 - \frac{n_{(0)} p_{n}}{\varepsilon_{(0)}(1 + w)} \right) \left( n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1 + w)} \varepsilon_{(1)} \right),
\]

(A.18)

where MAXIMA [23] has been used to perform the algebraic task. By virtue of (A.16), one may in this equation replace \( n_{(1)} \) and \( \varepsilon_{(1)} \) by \( n_{(1)}^{gi} \) and \( \varepsilon_{(1)}^{gi} \), respectively, thus obtaining equation (A.1b).

A.3 Evolution Equations for the Contrast Functions

First the entropy equation (2.44b) will be derived. From the definitions (2.43) it follows that

\[
n_{(1)}^{gi} - \frac{n_{(0)}}{\varepsilon_{(0)}(1 + w)} \varepsilon_{(1)}^{gi} = n_{(0)} \left( \delta_n - \frac{\delta_\varepsilon}{1 + w} \right).
\]

(A.19)

Differentiating this expression with respect to \( c t \) yields

\[
\frac{1}{c} \frac{d}{dt} \left( n_{(1)}^{gi} - \frac{n_{(0)}}{\varepsilon_{(0)}(1 + w)} \varepsilon_{(1)}^{gi} \right) = \dot{n}_{(0)} \left( \delta_n - \frac{\delta_\varepsilon}{1 + w} \right) + n_{(0)} \frac{1}{c} \frac{d}{dt} \left( \delta_n - \frac{\delta_\varepsilon}{1 + w} \right).
\]

(A.20)

Using equations (2.7e) and (A.18), one arrives at equation (2.44b) of the main text.

Finally, equation (2.44a) is derived. Upon substituting the expression

\[
\varepsilon_{(1)}^{gi} = \varepsilon_{(0)} \delta_\varepsilon,
\]

(A.21)

into equation (A.1a), and dividing by \( \varepsilon_{(0)} \), one finds

\[
b_1 = 2 \frac{\dot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} + a_1, \quad b_2 = \frac{\dot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} + a_1 \frac{\dot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} + a_2, \quad b_3 = a_3 \frac{n_{(0)}}{\varepsilon_{(0)}},
\]

(A.22)

where also (A.19) has been used. Using MAXIMA [23], one arrives at the coefficients (2.45) of the main text.
Acknowledgments
The author is indebted to Willem van Leeuwen for critically reading the manuscript and helpful discussions.

References

[1] E. M. Lifshitz. On the Gravitational Stability of the Expanding Universe. *J. Phys.*, X(2):116–129, 1946.

[2] E. M. Lifshitz and I. M. Khalatnikov. Investigations in Relativistic Cosmology. *Adv. Phys.*, 12:185–249, 1963. doi: 10.1080/00018736300101283.

[3] J. M. Bardeen. Gauge-invariant Cosmological Perturbations. *Phys. Rev.*, D22:1882–1905, 1980. doi: 10.1103/PhysRevD.22.1882.

[4] J. M. Stewart. Perturbations of Friedmann-Robertson-Walker Cosmological Models. *Class. Quantum Grav.*, 7:1169–1180, 1990. doi: 10.1088/0264-9381/7/7/013.

[5] H. Kodama and M. Sasaki. Cosmological Perturbation Theory. *Progress of Theoretical Physics Supplements*, 78:1–166, 1984. doi: 10.1143/PTPS.78.1.

[6] G. F. R. Ellis and M. Bruni. Covariant and Gauge-invariant Approach to Cosmological Density Fluctuations. *Phys. Rev. D*, 40(6):1804–1818, 1989. doi: 10.1103/PhysRevD.40.1804

[7] G. F. R. Ellis, J. Hwang, and M. Bruni. Covariant and Gauge-independent Perfect-Fluid Robertson-Walker Perturbations. *Phys. Rev. D*, 40(6):1819–1826, 1989. doi: 10.1103/PhysRevD.40.1819.

[8] George F. R. Ellis and Henk van Elst. Cosmological Models (Cargèse lectures 1998), 1998. [http://arxiv.org/abs/gr-qc/9812046](http://arxiv.org/abs/gr-qc/9812046).

[9] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger. Theory of Cosmological Perturbations. *Phys. Rep.*, 215(5 & 6):203–333, 1992. doi: 10.1016/0370-1573(92)90044-Z.

[10] V. Mukhanov. *Physical Foundations of Cosmology*. Cambridge University Press, 2005. ISBN 978-0521563987.

[11] Marco Bruni, Peter K.S. Dunsby, and George F.R. Ellis. Cosmological perturbations and the physical meaning of gauge invariant variables. *Astrophys.J.*, 395:34–53, 1992. doi: 10.1086/171629.

[12] J. W. York, jr. Covariant Decomposition of symmetric Tensors in the Theory of Gravitation. *Annales de l’ Institut Henry Poincaré - Section A: Physique théorique*, XXI(4):319–332, 1974.

[13] J. M. Stewart and M. Walker. Perturbations of Space-times in General Relativity. *Proc. R. Soc. London*, A 341:49–74, 1974. doi: 10.1098/rspa.1974.0172.
[14] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields*, volume 2 of *Course of Theoretical Physics*. Oxford: Butterworth-Heinemann (Elsevier), 4 edition, 2010. ISBN 978-0-7506-2768-9.

[15] P. G. Miedema and W. A. van Leeuwen. Density Perturbations in the Early Universe. *eprint* 2003. [http://arXiv.org/abs/gr-qc/0303004](http://arXiv.org/abs/gr-qc/0303004).

[16] Steven Weinberg. *Cosmology*. Oxford University Press Inc., New York, first edition, 2008. ISBN 978-0198526827.

[17] S. Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley & Sons, Inc. New York, 1972. ISBN 978-0471925675.

[18] Adam J. Christopherson. Applications of Cosmological Perturbation Theory. *eprint*, 2011. [http://arXiv.org/abs/1106.0446](http://arXiv.org/abs/1106.0446).

[19] Karim A. Malik and David Wands. Cosmological Perturbations. *Phys.Rept.*, 475:1–51, 2009. doi: 10.1016/j.physrep.2009.03.001. [http://arXiv.org/abs/0809.4944](http://arXiv.org/abs/0809.4944).

[20] Christian Knobel. An Introduction into the Theory of Cosmological Structure Formation. *eprint*, 2012. [http://arXiv.org/abs/1208.5931](http://arXiv.org/abs/1208.5931).

[21] Patrick Peter. Cosmological Perturbation Theory. *eprint*, 2013. [http://arXiv.org/abs/1303.2509](http://arXiv.org/abs/1303.2509).

[22] K. A. Malik and D. R. Matravers. Comments on Gauge-invariance in Cosmology. *eprint*, 2012. [http://arXiv.org/abs/1206.1478](http://arXiv.org/abs/1206.1478) doi: 10.1007/s10714-013-1573-2.

[23] Maxima. Maxima, a Computer Algebra System. Version 5.32.1, 2014. [http://maxima.sourceforge.net](http://maxima.sourceforge.net).

[24] S. M. Carroll. *Space Time and Geometry*. Benjamin Cummings, first edition, 2003. ISBN 978-0805387322.