Second order cosmological perturbations: simplified gauge change formulas

Claes Ugglal and John Wainwright

1 Department of Physics, Karlstad University, S-651 88 Karlstad, Sweden
2 Department of Applied Mathematics, University of Waterloo, Waterloo, ON, N2L 3G1, Canada
E-mail: claes.uggla@kau.se and jwainwri@uwaterloo.ca

Received 9 August 2018, revised 8 November 2018
Accepted for publication 27 November 2018
Published 14 January 2019

Abstract
In this paper we present a new formulation of the change of gauge formulas in second order cosmological perturbation theory which unifies and simplifies known results. Our approach is based on defining new second order scalar perturbation variables by adding a multiple of the square of the corresponding first order variables to each second order variable. A bonus is that these new perturbation variables are of broader significance in that they also simplify the analysis of second order scalar perturbations in the super-horizon regime in a number of ways, and lead to new conserved quantities.

Keywords: cosmology, perturbations, gauge

1. Introduction
Cosmological perturbation theory plays a central role in confronting theories of the early universe with observations. The increasing accuracy of the observations, however, has made it desirable to extend the theory from linear to second order (i.e. nonlinear) perturbations, which presents various technical challenges. For example, in applying cosmological perturbation theory at second order it is often desirable to use several gauges since the physical interpretation may require one gauge while the mathematical analysis may be simpler using a different gauge. The change of gauge formulas at second order that are needed in this situation have a complicated structure and are cumbersome to work with due to the presence of source terms that depend quadratically on the first order perturbations. This state of affairs motivated us to revisit the problem of gauge change at second order. Our starting point was two papers that give a fairly comprehensive treatment of this topic, with a description of the general method for deriving the formulas, together with various specific cases, namely, Noh and Hwang [21]

3 See, for example, Bartolo et al [3] and Tram et al [24].
(see sections VI and VII) and Malik and Wands [16] (see sections 6 and 7)⁴. Reading and interpreting the formulas in these papers is not easy due to their complexity, and comparing formulas in the papers is difficult because of the lack of a standard notation. However, while studying the formulas in these papers we noticed that for scalar perturbations they have certain features in common that enables one to write them in a unified and simpler form.

We consider first and second order scalar perturbations of Friedmann–Lemaître (FL) universes subject to the following assumptions:

(i) the spatial background is flat;
(ii) the stress–energy tensor can be written in the form $T^b_a = (\rho + p)u^a u_b + p\delta^b_a$, thereby describing perfect fluids and scalar fields;
(iii) the linear perturbation is purely scalar.

We will use the following five gauges: the Poisson (longitudinal, zero shear) gauge, the uniform curvature (spatially flat) gauge, the total matter gauge, the uniform density gauge, and the uniform scalar field gauge.

The simplification of the gauge change formulas is accomplished by making three choices. First, we use a common fixing of the spatial gauge freedom so that the remaining degrees of freedom in the gauge vector fields $\xi^a_r, r = 1, 2$, are the temporal components at first and second order. Second, we normalize background and perturbation variables so that they are dimensionless. In particular, as the time variable we use the so-called e-fold time $N$ which is defined by $N = \ln(a/a_0)$, where $a$ is the background scale factor⁶. The scalar $N$ represents the number of background e-foldings from some reference time $a_0$. Third, a careful inspection of the source expressions in the known gauge change formulas reveals that a number of quadratic first order terms can be incorporated in a systematic way into the second order perturbation variables and the temporal gauge vector field, leaving much simplified source terms. We will use a hat notation $\hat{f}$ for these source-compensated second order perturbation variables. For the metric and matter variables we can give a unified definition as follows:

$$ (2)\hat{\Box} := (2)\Box + C_\Box (1)\Box^2, $$

where the kernel $\Box$ represents a dimensionless metric or matter perturbation and the coefficient $C_\Box$ depends on the background variables, while for the temporal component of the gauge vector field using e-fold time $N$ as time coordinate we define

$$ (2)\hat{\xi}^N := (2)\xi^N - (1)\xi^N \partial_N (1)\xi^N, $$

where we write $\partial_N \equiv \partial/\partial N$ for brevity. In terms of these quantities, inspection of the known change of gauge formulas leads to the following unified form:

$$ (1)\Box = (1)\Box - (1)e^N, $$

$$ (3a) $$

$$ (2)\hat{\xi}^N = (2)\xi^N - (2)\xi^N + 2(1)\xi^N \partial_N (1)\Box + \Box_{\text{rem}}, $$

$$ (3b) $$

⁴ Other papers that have been influential in developing and applying second order cosmological perturbation theory but do not emphasize change of gauge formulas are Bartolo et al [1] and Nakamura [20]. Examples of recent papers that use change of gauge formulas at second order are Malik [15], Christopherson et al [6], Hidalgo et al [10], Christopherson et al [7], Carrilho and Malik [5], Dias et al [8], Villa and Rampf [31] (see equations (3.11)–(3.14) and (3.22)–(3.25) with the source terms given in equations (C.1)–(C.4)) and Hwang et al [11].

⁵ We refer to Malik and Wands [16], section 7.5, for this terminology. See also Liddle and Lyth [13], page 343. This gauge was apparently introduced by Kodama and Sasaki [12], and called the velocity-orthogonal isotropic gauge (see page 45, case 2b).

⁶ See for example, Martin and Ringeval [17].
where the subscript \( \bullet \) stands for a letter describing a particular gauge choice. We find that the reminder term \( \square_{\text{rem, \bullet}} \) for most metric and matter variables is a simple quadratic function of the first order variables, which in the case of any scalar variable is in fact zero.

The outline of the paper is as follows. In section 2, after introducing the notation that we will use for the metric and matter variables, we present the details concerning the unified definition (1) of the hat variables and the details concerning the unified form of the gauge transformation formula (3). In section 3 for each of the five choices of temporal gauge indicated by \( \cdot \) we give expressions for \( \square_{\text{\bullet}} \) in terms of the metric and matter perturbation variables (see equation (42)). This set of formulas, which provides an efficient unifying algorithm for calculating any gauge invariant in any of the above five gauges, is the main goal of the paper. In section 4 we use our unified scheme to give simple derivations of some important change of gauge formulas previously presented in the cosmological literature. In section 5 we point out that the present paper is the first of four closely connected papers. We also comment on how the new hatted variables result in new conserved quantities, as shown in detail in the sequel papers. In appendix A we make further comparisons of our gauge transformation formulas in section 2 with those in Malik and Wands [16], which served as our main starting point for the present paper.

2. Unified form for gauge transformations to second order

To perturb a flat FL background geometry it is convenient to write the metric as

\[
\text{d}s^2 = a^2 \left(- (1 + 2\phi) \text{d}\eta^2 + f_{ij} \text{d}x^i \text{d}x^j\right),
\]  

(4)

where \( a \) is the background scale factor and \( \eta \) is conformal time in the background. We assume that the metric components can be expanded in powers of a perturbation parameter \( \epsilon \), i.e. as a Taylor series, for example,

\[
\phi = \epsilon^{(1)\phi} + \frac{1}{2} \epsilon^{(2)\phi} + \ldots
\]  

(5)

We furthermore assume that the metric can be decomposed into scalar, vector, and tensor perturbations according to

\[
f_{ij} = D_i B + B_i,
\]  

(6a)

\[
f_{ij} = (1 - 2\psi) \gamma_{ij} + 2 D_i D_j C + 2 D_i (C_j) + 2 C_{ij},
\]  

(6b)

where \( D_i B_j = 0; D_i C_l = 0; C_l = 0; D^l C_{ij} = 0, \) and where \( D_i \) is the spatial covariant derivative corresponding to the flat metric \( \gamma_{ij} \). Use of Cartesian background coordinates yields \( \gamma_{ij} = \delta_{ij} \) and \( D_i = \partial / \partial x^i \). As regards dimensions, we make the choice that the scale factor \( a \) is dimensionless. It then follows that the coordinates \( \eta \) and \( x^i \) have dimensions of length since \( \text{d}s^2 \) has dimension length^2. Hence \( \phi \) and \( \psi \) are dimensionless while \( B \) has dimension length.

We consider a stress–energy tensor of the form:

\[
T_{a b}^\alpha = (\rho + p) u^a u_b + p \delta^\alpha_b,
\]  

(7)

which encompasses perfect fluids and scalar fields. The energy density \( \rho \), the pressure \( p \), and the four-velocity \( u^a \) can be expanded as a Taylor series in \( \epsilon \). Perturbations of the energy density \( \rho \) are therefore given by³

³ We use a subscript zero to denote the background value of some quantity, so that \( \rho_0 \) and \( p_0 \) are the background energy density and pressure.
\[ \rho = \rho_0 + \epsilon^{(1)} \rho + \frac{1}{2} \epsilon^2 (2) \rho + \ldots, \]  
(8)

and similarly for the pressure perturbations. We use the usual background matter variables \( w \) and \( c_s^2 \), and the deceleration parameter \( q \) defined according to

\[ w = \frac{p_0}{\rho_0}, \quad c_s^2 = \frac{p_0'}{\rho_0'}, \quad q = -\frac{H'}{H^2}, \]  
(9)

where \( \cdot' \) denotes the derivative with respect to the conformal background time variable \( \eta \), and \( H = a'/a = aH \) with \( H \) the background Hubble variable. We use units such that \( c = 1 \) and \( 8\pi G = 1 \), where \( c \) is the speed of light and \( G \) the gravitational constant. It follows that \( H \) has dimension of \((\text{length})^{-1}\) and that \( q, w \) and \( c_s^2 \) are dimensionless.

Since we have assumed that the spatial background is flat, the Einstein field equations in the background can be written as

\[ 3H^2 = a^2 \rho_0, \quad 2(-H' + H^2) = a^2 (\rho_0 + p_0), \]  
(10)

which in conjunction with (9) yields the following relation between \( w \) and the deceleration parameter \( q \):

\[ 1 + q = \frac{3}{2} (1 + w). \]  
(11)

a result that we will use frequently.

To define the scalar velocity perturbations we find it convenient to work with the covariant four-velocity \( u_b \), which we normalize with a conformal factor \( a \) according to \( u_b = aV_b \).

\[ u_b = a V_b, \quad \rho = \frac{\rho_0}{\rho_0'}, \quad c_s^2 = \frac{c_s^2}{c_s^2}, \quad q = -\frac{q}{q}. \]  
(12a)

with \( D^r \tilde{V}_r = 0 \), so that \( ^rV \) represents the scalar perturbations. Since the \( V_b \) are dimensionless and the \( x' \) have dimension length it follows from (12) that \( ^rV \) has dimension length.

In the case in which the matter-energy content is provided by a minimally-coupled scalar field, we will use \( \varphi \) to denote the scalar field and define the perturbations according to

\[ \varphi = \varphi_0 + \epsilon^{(1)} \varphi + \frac{1}{2} \epsilon^2 (2) \varphi + \ldots. \]  
(13)

Next we turn to gauge transformations in cosmological perturbation theory. We begin by considering an arbitrary one-parameter family of a tensor field \( A (\epsilon) \), which can be expanded in powers of \( \epsilon \), i.e. as a Taylor series:

\[ A(\epsilon) = A_0 + \epsilon^{(1)} A + \frac{1}{2} \epsilon^2 (2) A + \ldots. \]  
(14)

A gauge transformation induces a change in the first and second order perturbations of \( A(\epsilon) \). Arguably the most geometric and straightforward approach to gauge transformations is the ‘active approach’ using an exponential map described in section 6 of Malik and Wands [16].
and this is the approach we take as our starting point. First and second order gauge transformations are then represented as (equations (6.5) and (6.6), respectively, in [16]):

\[
(1) A[\xi] = (1) A + \mathcal{L}_{\xi} A_0, \quad (15a)
\]

\[
(2) A[\xi] = (2) A + \mathcal{L}_{\xi} A_0 + \mathcal{L}_{2\xi} \left( (1) A + \mathcal{L}_{\xi} A_0 \right), \quad (15b)
\]

where \( (1) \xi^a \) and \( (2) \xi^a \) are independent background gauge vector fields and \( \mathcal{L} \) is the Lie derivative (see also [4], equations (1.1)–(1.3)). Equation (15) describes how the tensor field \( A \) changes under an arbitrary gauge transformation. More importantly from a physical point of view, these equations serve to define gauge invariant quantities in the following way. If we impose a restriction on the perturbation variables that determines the gauge fields uniquely, say \( (1) \xi^a = (1) \xi_a^* \), \( (2) \xi^a = (2) \xi_a^* \), then we say that we have fixed the gauge. If we use these as the gauge fields in (15), then the quantities \((1) A[\xi_a^*]\) and \((2) A[\xi_a^*]\) so defined are gauge invariant quantities. On introducing the shorthand notation

\[
(1) A_a^* = (1) A[\xi_a^*], \quad (2) A_a^* = (2) A[\xi_a^*]. \quad (16)
\]

equations (15) yield

\[
(1) A_a^* = (1) A + \mathcal{L}_{\xi_a^*} A_0, \quad (17a)
\]

\[
(2) A_a^* = (2) A + \mathcal{L}_{\xi_a^*} A_0 + \mathcal{L}_{2\xi_a^*} \left( (1) A + \mathcal{L}_{\xi_a^*} A_0 \right). \quad (17b)
\]

We say that \((1) A_a^*\) and \((2) A_a^*\) are the first and second order gauge invariants associated with the tensor field \( A \) in the gauge specified by the subscript \( a \). We list several gauges and their identifying subscripts at the beginning of section 3.

We fix the spatial gauge freedom completely by setting the metric functions \( C \) and \( C_i \) in (6) to be zero order by order, which up to second order gives

\[
(1) C = 0, \quad (2) C_i = 0, \quad r = 1, 2. \quad (18)
\]

The above spatial gauge fixing is arguably the essence of Hwang and Noh’s so-called ‘gauge ready’ approach, who refer to it as the \( C \)-gauge (see for example Noh and Hwang [21], equation (259)). Note that this is the only way one can algebraically completely fix the spatial gauge by using the metric components and matter variables for the present models (see e.g. the gauge transformations given in [26]), and as a consequence all the gauges listed in the introduction and section 3 are characterized by this condition. The only gauge that is commonly used that does not fulfil this condition is the synchronous gauge, which is useful for treating dust models. However, the synchronous gauge is not a fully fixed gauge and the natural way to completely fix this gauge, and thereby relate quantities to physical observables, is to relate it to the total matter gauge, which does obey the above conditions, see appendix B.7 in [26].

As a consequence of the above spatial gauge fixing, the remaining gauge freedom is described by gauge fields to second order restricted to be of the form

\[\text{Gauge transformations up to second order in cosmological perturbation theory can also be represented in coordinates as follows (see e.g. Malik and Wands [16]):}\]

\[\tilde{x}^a = x^a + \epsilon(1)\xi^a + \frac{1}{2} \xi^a \left( (2) \xi^a + (1)\xi^b \xi^b \right).\]

\[\text{For a recent work using the synchronous gauge for models with dust, see e.g. Gressel and Bruni [9].}\]
where \( (2)\xi \) and \( (2)\bar{\xi} \) are determined by quadratic source terms that arise from the conditions (18), where, in particular, \( (2)\xi \) depends on \( (1)\xi^N \). As in the introduction we are using the e-fold time defined by \( N = \ln(a/a_0) \) as the time coordinate instead of the conformal time \( \eta \). Note that the temporal components of the gauge fields are related according to \( \xi^N = \mathcal{H}\xi^0 \), which follows from

\[
\partial_\eta = \mathcal{H}\partial_N. \tag{20}
\]

In this paper we will primarily use e-fold time \( N \) but we will also use conformal time \( \eta \), depending on the context. Equation (20) enables one to make the transition and we will use it frequently.

We are further restricting our considerations to perturbations that are purely scalar at linear order, i.e. the metric functions that describe vector and tensor perturbations at first order are zero:

\[
(1)B_i = 0, \quad (1)C_{ij} = 0, \quad (1)\bar{\psi}_i = 0. \tag{21}
\]

Bartolo et al [1] (see page 41) argue that this restriction is reasonable on physical grounds, since vector perturbations have decreasing amplitude and are not generated during inflation, while tensor perturbations are expected to be negligible. On the other hand it is well known ([1], see page 41) that even if the vector and tensor perturbations are zero at first order, they will be generated at second order due to the presence of source terms in the vector and tensor governing equations, since these source terms depend on the first order scalar perturbations. Thus even if the first order restriction (21) holds we will have

\[
(2)B_i \neq 0, \quad (2)C_{ij} \neq 0, \quad (2)\bar{\psi}_i \neq 0, \tag{22}
\]

at second order. In this context, however, the second order scalar perturbations are independent of the second order vector and tensor perturbations and hence can be studied separately. In this paper we are choosing to consider only the scalar perturbations at second order, which physically represent density perturbations, leaving the second order vector and tensor perturbations for future work [11]. We are thus studying second order scalar perturbations subject to the first order restriction (21), and they are represented by the functions \( (r)\phi, (r)B_i, (r)\psi_{ij}, (r)V_i, (r)\rho, (r)p, (r)\varphi \), and the remaining gauge freedom which is described by the functions \( (r)\xi_i, r = 1, 2 \).

We now describe how the first order variables \( (1)B, (1)\psi, (1)V_i, (1)\rho, (1)\varphi \) transform under the remaining temporal gauge freedom. From (17a), using e-fold time, one obtains the well-known relations:

\[
(1)B_i = (1)B - \mathcal{H}^{-1}(1)\xi_i, \quad (1)\psi_i = (1)\psi - (1)\xi_i, \tag{23a}
\]

\[
(1)V_i = (1)V - \mathcal{H}^{-1}(1)\xi_i, \quad (1)A_i = (1)A + (\partial_\eta A_0)(1)\xi_i. \tag{23b}
\]

where \( A = \rho \) or \( A = \varphi \). By normalizing the perturbations (apart from \( \psi \)) these transformation rules can be written in the unified form given in equation (3a), which we repeat here:

\[
(1)\Box = (1)\Box - (1)\xi_i. \tag{24}
\]

where the kernel \( \Box \) represents the following variables in the five different cases:

---

10 See equations (B10e) and (B10f) in [26] for the transformation laws for \( C \) and \( C_r \).

11 The tensor mode at second order describes gravitational waves generated by the first order scalar (i.e. matter) perturbations.
\[ \Box = \psi, \quad \Box = \mathcal{H} B, \quad \Box = \mathcal{H} V, \quad \Box = \frac{\rho}{-\partial_N \rho_0}, \quad \Box = \frac{\varphi}{-\partial_N \varphi_0}. \]  

The above normalization ensures that the variables are dimensionless (recall that \( B \) and \( V \) have dimension length and \( \mathcal{H} \) has dimension \( \text{(length)}^{-1} \)).

We now consider the second order perturbation variables. The second order gauge transformation formulas, which follow from (17b), can be written so as to have the same leading order terms as the first order formula (24) but they also include a source term \( S_{\Box} \) that depends quadratically on the first order variables in a sometimes complicated way:

\[ (2) \Box = (2) \Box - (2) \xi^N + S_{\Box}. \]  

As described in the introduction, inspection of the different expressions \( S_{\Box} \) reveals that a number of quadratic first order terms can be incorporated into the second order terms in the formula (26), leaving much simplified source terms. The resulting unified formula is given by equation (3b), which we repeat here for the reader's convenience:

\[ (2) \Box = (2) \Box - (2) \xi^N + 2 (1) \xi^N \partial_N (1) \Box + \Box_{\text{rem}}, \]  

where \( \Box \) stands for one of the variables in (25). The hatted variables are given by equations (1) and (2), which we repeat here:

\[ (2) \hat{\Box} = (2) \Box + C_{\Box} (1) \Box^2, \]  

\[ (2) \hat{\xi}^N = (2) \xi^N - (1) \xi^N \partial_N (1) \xi^N. \]

The details of this unified formulation lie in the coefficients \( C_{\Box} \) in (28a) and in the remaining source terms \( \Box_{\text{rem}} \) in (27). The expressions for these quantities are obtained by comparing the transformation rules obtained from (17b) for each choice of \( \Box \) in (25) with the unified form (27).

First, the coefficients \( C_{\Box} \) in (28a) are given by:

\[ C_{\psi} = 2, \quad C_B = C_V = 1 + q, \quad C_\rho = \frac{\partial_N \rho_0}{\partial_N \rho_0}, \quad C_{\varphi} = \frac{\partial_N \varphi_0}{\partial_N \varphi_0}. \]  

For a non-interacting fluid or a non-interacting scalar field\(^{12}\), so that energy conservation holds in the background:

\[ \partial_N \rho_0 = -3 (\rho_0 + p_0), \]  

it follows that\(^{13}\)

\[ C_\rho = -3 (1 + c_s^2), \]  

\[ C_{\varphi} = (1 + q) - \frac{3}{2} (1 + c_s^2) = \frac{3}{2} (w - c_s^2). \]

\(^{12}\) The stress–energy tensor of a minimally coupled scalar field is equivalent to that of a perfect fluid, thereby defining an energy density and pressure for the scalar field. This equivalence leads to the following relation between \( \varphi_0 \) and \( \rho_0 + p_0 \): \( (\partial_N \varphi_0)^2 = a^2 \mathcal{H}^{-2} (\rho_0 + p_0) \).

\(^{13}\) For the first equation differentiate (30) and with respect to \( N \) and use the definition of \( c_s^2 \) expressed in terms of \( N \). For the second equation differentiate \( (\partial_N \varphi_0)^2 = a^2 \mathcal{H}^{-2} (\rho_0 + p_0) \) using the result of the first equation. One also requires \( \partial_N (a \mathcal{H}^{-1}) = (1 + q) (a \mathcal{H}^{-1}) \), which follows from (9) and (20). The second equality in (31b) depends on (11).
Second, the remaining source terms \( \square_{\text{rem}, \bullet} \) have the following form for the different choices of \( \square \) in (25):

\[
\psi_{\text{rem}, \bullet} = D_2(B) - D_2(B_\bullet), \tag{32a}
\]

\[
\mathcal{H}B_{\text{rem}, \bullet} = (\partial_N + 2q) (D_0(\mathcal{H}B_\bullet) - D_0(\mathcal{H}B))
+ 2S^i[(\phi_p + \phi_\bullet)D_i(\mathcal{H}B_\bullet) - (\phi_p + \phi)D_i(\mathcal{H}B)], \tag{32b}
\]

\[
\mathcal{H}V_{\text{rem}, \bullet} = 2S^i[\phi_v D_i(\mathcal{H}V_\bullet - \mathcal{H}V)], \tag{32c}
\]

\[
\rho_{\text{rem}, \bullet} = 0, \tag{32d}
\]

\[
\varphi_{\text{rem}, \bullet} = 0, \tag{32e}
\]

where the subscripts \( p \) and \( v \) stands for the Poisson and total matter gauge, respectively, which are defined in section 3. The scalar mode extraction operator \( S_i \) and the spatial differential operators\(^{14}\) \( D_0 \) and \( D_2 \) that appear in equation (32) are defined in equations (B.2a) and (B.3a) in appendix B. We note that the temporal gauge on the right side of equation (32) is unspecified and can be chosen to be one of the standard gauges.

At this stage we introduce the normalized density perturbation \( (r)\delta \) according to

\[
(\rho)\delta = \frac{(\rho)}{\rho_0}, \quad r = 1, 2, \tag{33}
\]

which means that the kernel \( \square \) that is associated with the density perturbation in (25) is given by

\[
(\rho)\square = \frac{1}{3}(\rho)\delta. \tag{34}
\]

The factor of \( \frac{1}{3} \) is included so that if conservation of energy holds in the background \((\partial_N\rho_0 = -3(\rho_0 + p_0))\) then (33) becomes

\[
(\rho)\delta = \frac{(\rho)}{\rho_0 + p_0}. \tag{35}
\]

In this case the usual fractional perturbed energy density \( (r)\delta \) is easily obtained via

\[
(\rho)\delta = \frac{(\rho)}{\rho_0} = (1 + w)(\rho)\delta, \tag{36}
\]

provided that the cosmological constant is zero.

For convenience we now explicitly list the normalized source-compensated second order variables given by (28a), where the kernels \( \square \) are given by (25) and (34):

\[
(2)\rho_\bullet = (2)\rho_\bullet + 2(1)\rho_\bullet^2, \tag{37a}
\]

\[
\mathcal{H}(2)B = \mathcal{H}(2)B + (1 + q)(\mathcal{H}(1)B)^2, \tag{37b}
\]

\[
\mathcal{H}(2)V = \mathcal{H}(2)V + (1 + q)(\mathcal{H}(1)V)^2, \tag{37c}
\]

\(^{14}\) We decided to introduce this shorthand notation because these expressions occur frequently in second order cosmological perturbation theory. See appendix B for some notational motivation and historical background concerning these expressions.
\[
\frac{1}{3} \delta^{(2)} = \frac{1}{3} \delta^{(1)} - 3(1 + c_1^2) \left( \frac{1}{3} \delta^{(1)} \right)^2,
\]
\[\lambda^{(2)} \hat{\varphi} = \lambda^{(2)} \varphi + \frac{3}{2} (w - c_2^2) (\lambda^{(1)} \varphi)^2.\]

Here we have introduced the notation
\[\lambda = - (\partial_N \varphi^0)^{-1},\]
for the scale factor associated with the scalar field in equation (25). We note that equations (37d) and (37e) depend on the conservation of energy in the background.

One feature of equation (27) requires comment. In this equation one can replace \((\square)\) on the right side by \((\square)\) using (24), and then modify the definition of \((\hat{\xi})\) by changing the sign in \((28b)\). Although this form of the equation may look more natural, it turns out that the form we have given is more convenient when we actually apply the equation to make a change of gauge in section 3.

Equation (27) with (37) and (32) forms the first main result of this paper and provides the basis for the change of gauge formulas in section 3. For ease of reference we now write out the unified formula (27) with \(\square\) having the values in (25) and (34):

\[\begin{align*}
\delta^{(2) \hat{\psi}_*} &= \delta^{(2) \hat{\psi}} - (\hat{\xi})^{\hat{\psi}_N} + 2 (\hat{\xi})^{\hat{\psi}_N} \partial_N (\hat{\psi})_* + \psi_{\text{rem}, \bullet}, \\
\mathcal{H}^{(2) \hat{B}_*} &= \mathcal{H}^{(2) \hat{B}} - (\hat{\xi})^{\hat{B}_N} + 2 (\hat{\xi})^{\hat{B}_N} \partial_N (\mathcal{H}^{(1) \hat{B}_*}) + \mathcal{H}_{\text{rem}, \bullet}, \\
\mathcal{H}^{(2) \hat{V}_*} &= \mathcal{H}^{(2) \hat{V}} - (\hat{\xi})^{\hat{V}_N} + 2 (\hat{\xi})^{\hat{V}_N} \partial_N (\mathcal{H}^{(1) \hat{V}_*}) + \mathcal{H}_{\text{rem}, \bullet}, \\
\frac{1}{3} \delta^{(2) \hat{\psi}_*} &= \frac{1}{3} \delta^{(2) \hat{\psi}} - (\hat{\xi})^{\hat{\psi}_N} + 2 (\hat{\xi})^{\hat{\psi}_N} \partial_N \left( \frac{1}{3} \delta^{(1) \hat{\psi}_*} \right), \\
\lambda^{(2) \hat{\varphi}_*} &= \lambda^{(2) \hat{\varphi}} - (\hat{\xi})^{\hat{\varphi}_N} + 2 (\hat{\xi})^{\hat{\varphi}_N} \partial_N (\lambda^{(1) \hat{\varphi}_*}),
\end{align*}\]

where \(\lambda\) is given by (38) and the remainder terms by (32). However, we need to augment this set of equations with transformation equations for the metric variable \(\phi\), which has to be treated separately since its transformation law involves the time derivative of the gauge field. At first order we have:

\[\delta^{(1) \psi}_* = \delta^{(1) \psi} + (\partial_N + 1 + q) (\hat{\xi})^{(1) \hat{\psi}_N},\]

and at second order,

\[\delta^{(2) \psi}_* = \delta^{(2) \psi} + (\partial_N + 1 + q) (\hat{\xi})^{(2) \hat{\psi}_N} + 2 (\hat{\xi})^{\hat{\psi}_N} \partial_N (\hat{\psi})_* + \psi_{\text{rem}, \bullet},\]

where

\[\delta^{(2) \psi} = \delta^{(2) \psi} - 2 (\hat{\psi})_*^2,\]

\[\psi_{\text{rem}, \bullet} = (\partial_N (\hat{\xi})^{(1) \hat{\psi}_N})^2 - (\partial_N \hat{\psi}(\hat{\xi})^{(1) \hat{\psi}_N})^2.\]

We end this section by noting that there are other ways defining the curvature perturbation \(\psi\). In this paper we write the scalar part of the perturbed spatial metric as \(1 - 2 \psi \hat{\psi}\_\bullet\) (which we refer to as the Malik–Wands form, see for example Malik and Wands [16]), while another
choice is an exponential form \( e^{-2\psi_{SB}} \delta_j \) first introduced by Salopek and Bond [22] (see for example, Lyth and Rodriguez [14], section IIB). Equating the two forms, Taylor expanding the exponential and performing a perturbation expansion for \( \psi_{SB} \) yields \((1)\psi_{SB} = (1)\psi\) and \((2)\psi_{SB} = (2)\psi + 2(1)\psi^2\) \(\), showing that \((2)\psi_{SB} = (2)\psi\). We refer to section 2.1 in Carrilho and Malik [5] for two other possibilities.

### 3. Performing a change of gauge

Having fixed the spatial gauge (see equation (18)), we can now choose a temporal gauge to second order by setting to zero the first and second perturbations of one of the variables \(B, \psi, V, \delta, \varphi\), thereby specifying the gauge uniquely. We use the following terminology and subscripts to label the gauges:

(i) Poisson gauge, subscript \(p\), defined by \(B_p = 0\),
(ii) uniform curvature gauge, subscript \(c\), defined by \(\psi_c = 0\),
(iii) total matter gauge, subscript \(\varphi\), defined by \(V_\varphi = 0\),
(iv) uniform density gauge, subscript \(\rho\), defined by \(\delta_\rho = 0\),
(v) uniform scalar field gauge\(^{15}\), subscript \(sc\), defined by \(\varphi_{sc} = 0\).

In order to introduce a specific gauge labelled by \(\bullet\) we must determine the transition function \((2)\xi_N\) using equation (39). (We will not list the expressions for \((1)\xi_N\) below since they can easily be read off from the second order equations: replace \((2)\) by \((1)\), omit the hats and drop the \(rem\) terms.) Referring to the above definition of the gauges we choose \(\bullet = p\) in (39b), \(\bullet = \varphi\) in (39c), \(\bullet = c\) in (39a), \(\bullet = \rho\) in (39d) and \(\bullet = sc\) in (39e), to obtain the following results:

\[
\begin{align*}
(\rho)B_p &= 0 \implies (2)\xi_N^p = \mathcal{H}(2)\hat{B} + \mathcal{H}B_{rem,p}, \\
(\rho)V_\varphi &= 0 \implies (2)\xi_N^\varphi = \mathcal{H}(2)\hat{V} + \mathcal{H}V_{rem,\varphi}, \\
(\rho)\psi_c &= 0 \implies (2)\xi_N^c = (2)\hat{\psi} + \psi_{rem,c}, \\
(\rho)\delta_\rho &= 0 \implies (2)\xi_N^\rho = \frac{1}{3}(2)\hat{\delta}, \\
(\rho)\varphi_{sc} &= 0 \implies (2)\xi_N^{sc} = \lambda(2)\hat{\varphi}.
\end{align*}
\]

These expressions for the gauge fields at second order represent the second main result of this paper. Their concise form is a consequence of using the hatted variables.

The final step is to successively substitute the expressions (41) into (27). This immediately gives the following change of gauge formulas at second order:

\[
\begin{align*}
(2)\xi_p &= (2)\hat{\xi} - \mathcal{H}(2)\hat{B} + 2\mathcal{H}(1)B\partial_N(1)\xi_p + \square_{rem,p} - \mathcal{H}B_{rem,p}, \\
(2)\xi_\varphi &= (2)\hat{\xi} - \mathcal{H}(2)\hat{V} + 2\mathcal{H}(1)V\partial_N(1)\xi_\varphi + \square_{rem,\varphi} - \mathcal{H}V_{rem,\varphi}, \\
(2)\xi_c &= (2)\hat{\xi} - (2)\hat{\psi} + 2(1)\psi\partial_N(1)\xi_c + \square_{rem,c} - \psi_{rem,c}.
\end{align*}
\]

\(^{15}\) This gauge is naturally only available in a perturbed universe with a scalar field. In this context it is in fact equivalent to the total matter gauge, but it is helpful to give it a separate name. This equivalence is established in a subsequent paper [28].
\( (2) \varphi^\rho = (3) \varphi - \frac{1}{3} (2) \delta + 2 \left( \frac{1}{3} (1) \delta \right) \partial_N (1) \varphi^\rho + \square_{\text{rem},\rho}, \)  
\( (42d) \)

\( (2) \varphi_{\text{sc}} = (2) \varphi - \lambda (2) \varphi + 2 (\lambda (1) \varphi) \partial_N (1) \varphi_{\text{sc}} + \square_{\text{rem},\text{sc}}, \)  
\( (42e) \)

where the \( \square_{\text{rem},\bullet} \) terms are given by (32), and \( \square \) represents any of the symbols in equation (25) and (34). The gauge on the right side is unspecified and can be chosen to be one of the standard gauges. For example if one wishes to transform from the total matter gauge to the uniform curvature gauge, one would use the third equation with subscripts \( v \) added on the right side:

\( (2) \varphi^c = (2) \varphi - (2) \varphi^v + 2 \left( \frac{1}{3} \varphi^v \right) \partial_N (1) \varphi^c + \square_{\text{rem},c,v} - \varphi_{\text{rem},c,v}, \)  
\( (43) \)

with

\( (1) \varphi^c = (1) \varphi - (1) \varphi^v, \)  
\( (44) \)

at first order\(^\text{16} \). The remainder terms are obtained from equation (32) by choosing the total matter gauge on the right side. In the present example we obtain

\[ \varphi_{\text{rem},c,v} = D_2^{(1)} (B_v) - D_2^{(1)} (B_c). \]  
\( (45) \)

In summary equation (42), in conjunction with the definition (37) of the hatted variables, represent the main goal of this paper. They provide an efficient algorithm for calculating any of the gauge invariants in any of the five gauges, as illustrated in the next section. Although our primary motivation was to simplify and unify the change of gauge formulas at second order, we note that equation (42) also give a useful overview of the situation at linear order. By replacing \( (2) \) with \( (1) \) and by omitting the hats and dropping the \( \text{rem} \) terms one can read off familiar relations such as

\[ \psi^v = \psi^p - \mathcal{H} V_p, \quad \psi_p = - \mathcal{H} B_c, \quad \delta^v = \delta^p - 3 \mathcal{H} V_p, \quad \psi^\rho = - \frac{1}{3} \delta^\rho, \quad \psi_{\text{sc}} = - \lambda \varphi_c. \]  
\( (46) \)

Finally, we recall that the change of gauge formulas for the metric perturbation \( \phi \) have to be treated separately and are given by equation (40), with the specific gauge field to be obtained from equation (41) once the two gauges have been chosen. An example at linear order is

\[ (1) \varphi^v = (1) \varphi^c + (\partial_N + 1 + q) (1) \xi^N_{\text{v,c}}, \quad \text{with} \quad (1) \xi^N_{\text{v,c}} = \mathcal{H} V_c. \]  
\( (47) \)

4. Examples

In this section we give examples of using the general equation (42) to calculate second order gauge invariants of interest in current research in cosmology. The expressions we obtain are more concise than those in the literature because of our use of the hatted variables and the differential operators \( D_0 \) and \( D_2 \). The latter feature, in particular, simplifies the representation of the terms involving spatial derivatives. In order to make comparisons with the literature it is necessary to expand our expressions by using the definition (37) of the hatted variables and the definition \( (B.3b,c) \) of \( D_0 \) and \( D_2 \). The latter definitions lead to the following identities that will be useful when making comparisons:

\[ D_0 (A) - D_0 (B) = S^{ij} [ D_i (A + B) D_j (A - B) ], \]  
\( (48a) \)

\(^{16} \) As with equation (41), the first order formulas can be read off from the second order formulas by inspection, since they correspond to the leading order terms.
\[ \square_2(A) - \square_2(B) = \frac{1}{3} (\square^2 S^d - \delta^d) [\square_1(A + B) \square_1(A - B)], \]

(48b)

where the scalar mode extraction operator \(S^d\) is defined by (B.2a). We will frequently change from \(e\)-fold time \(N\) to conformal time \(\eta\) using equation (20) in order to make comparisons with the literature. We note that the process of expanding our expressions to make comparisons with the literature, as illustrated in section 4 and in appendix A, can be tedious. We regard this as a measure of how concise our expressions are, and we emphasize that this is not something one has to do when using our formalism in practice, as discussed in section 5. A bonus of the conciseness is that it is easier to avoid errors and to find simpler unifying expressions.

Our first example concerns \((\hat{\psi})_\rho\), the second order curvature perturbation in the uniform density gauge, which is important as a conserved quantity on super-horizon scales\(^{17}\). Choosing \(\square = \psi\) in (42d) and using (32a) for the remainder term immediately gives

\[ (\hat{\psi})_\rho = (\hat{\psi})_\rho - \frac{1}{3} \delta^\rho + (1) \delta \partial^\rho \psi - \square_2(B_\rho) + \square_2(B_\rho), \]

(49)

Equation (49) is a concise version of equation (7.71) in Malik and Wands [16]\(^{18}\). If we choose the arbitrary temporal gauge to be the uniform curvature gauge \((\psi_\epsilon = 0)\) equation (49) becomes

\[ (\hat{\psi})_\rho = -\frac{1}{3} \delta^\rho + (1) \delta \partial^\rho \psi - \square_2(B_\rho) + \square_2(B_\rho), \]

(50)

which is a concise version of equation (3.3) in Christofferson et al [7], which relates \((\hat{\psi})_\rho\) to \((\hat{\delta})_\rho\), the second order density perturbation in the uniform curvature gauge. To compare with earlier literature we change from \(N\) to \(\eta\) and make the replacement \((\hat{\delta}) = -3(\mathcal{H}/\rho_0)\eta\rho\), which leads to

\[ (\hat{\psi})_\rho = \frac{\mathcal{H}}{\rho_0} (\hat{\psi})_\rho - \frac{\mathcal{H}}{(\rho_0)^2} (1) \rho_\epsilon \left( 2(1) \rho_\epsilon + (5 + 3c_s^2) \mathcal{H}(1) \rho_\epsilon \right) - \square_2(B_\rho) + \square_2(B_\rho), \]

(51a)

where

\[ \square_2(B_\rho) - \square_2(B_\rho) = \frac{1}{27} \mathcal{H}^{-2} (\square^2 S^d - \delta^d) [\square_1(1) \delta_\epsilon - 6 \mathcal{H}(1) \delta_\epsilon] [\square_1(1) \delta_\epsilon], \]

(51b)

the latter relation following from (48b) and \(\mathcal{H}(1) B_\rho = \mathcal{H}(1) B_\rho - \frac{1}{3} (1) \delta_\epsilon\). We find that equation (3.3) in [7] agrees with (51)\(^{19}\). Equation (51) has also been given by Carrilho and Malik [5] (see equation (3.3)).

Our second example concerns the density perturbation at second order in the total matter gauge \(\delta_\epsilon\), which is used when deriving the generalized Poisson equation in second order perturbation theory in relativistic cosmology (see Hidalgo et al [10]). Choosing \(\square = \frac{1}{4} \delta\) in (42b) and using (32d) for the remainder term we express \(\delta_\epsilon\) in terms of an arbitrary temporal gauge:

\[ (\hat{\delta})_\epsilon = (\hat{\delta}) - 3 \mathcal{H}(1) V + 2 \mathcal{H}(1) V \partial_\epsilon (1) \delta_\epsilon + 6 S^d(1) \partial_\epsilon \mathcal{H}(1) V, \]

(52)

where the mode extraction operator \(S^d\) is defined in equation (B.2a).

Expanding our equation, changing from \(N\) to \(\eta\) and using (33) leads to

\(^{17}\) See, for example, Bartolo et al [3], equation (36).

\(^{18}\) Set \(E_i = 0\) in their equation to fix the spatial gauge. There are a number of typos.

\(^{19}\) In rearranging the \(O(\square^2)\) terms one has to use (33), and the definition of \(S_\rho\).
\( (2)\rho_s = (2)\rho + \rho_0^f(2)V + (1)V[2(1)\rho_c' + 3\rho_0(1 + c_s^2 + \frac{1}{2}(1 + w))H(1)V] - 2\rho_0s'[(1)\phi, D_j(1)V]. \) (53)

Equation (3.10) in [10] can be simplified to have the form (53) (subject to a few differing coefficients) when we choose the arbitrary temporal gauge on the right side of (53) to be the Poisson gauge. To make the comparison it is easier to perform further manipulations on (53) and transform it into the Hidalgo et al form\(^\text{20}\).

As our third example, by choosing \( \Box = \frac{1}{4}\delta \) in (42c) and using (32d) for the remainder term, we express \( \delta_c \) in terms of an arbitrary temporal gauge:

\[ (2)\delta_c = (2)\delta - 3(2)\tilde{\psi} + 2(1)\tilde{\psi}\partial_N(1)\delta_c + 3([\Box_2(1)\mathcal{B}_c] - \Box_2(1)\mathcal{B}). \] (54)

Expanding our equation, changing from \( N \) to \( \eta \) and using (33) leads to

\[ (2)\rho_c = (2)\rho + \rho_0^f(2)\tilde{\eta} + \frac{(1)\tilde{\psi}}{H} \left[ 2(1)^r + \rho_0^f \left( 2(1)^r \right) \psi - \partial_N(1)^r \right] \]
\[ - \rho_0^f \left( [\Box_2(1)\mathcal{B}_c] - \Box_2(1)\mathcal{B} \right). \] (55a)

where

\[ \Box_2(1)\mathcal{B}_c - \Box_2(1)\mathcal{B} = \frac{1}{3}H^{-2}(D^\gamma S^\gamma - \delta^\gamma)[D_j(1)\psi - 2H(1)\mathcal{B}D_j(1)\psi], \] (55b)

the latter relation following from (48b) and \( H(1)\mathcal{B}_c = H(1)\mathcal{B} - (1)\psi \). Equation (55) agrees with equation (7.35) in Malik and Wands [16] with the gauge fixed so that \( E_1 = 0 \).\(^\text{21}\)

Our final section in this section concerns the curvature perturbation in the uniform scalar field gauge \( (2)\tilde{\psi}_w \) which is a conserved quantity on super-horizon scales in a scalar field dominated universe (Vernizzi [30]). Choosing \( \Box = \psi \) in (42e) and using (32a) for the remainder term immediately gives

\[ (2)\tilde{\psi}_w = (2)\tilde{\psi} - \lambda(2)^r - 2(1)^r \partial_N(1)^r \tilde{\psi}_w - [\Box_2(1)\mathcal{B}_w] + \Box_2(1)\mathcal{B}. \] (56)

Expanding our equation leads to\(^\text{22}\)

\[ (2)\tilde{\psi}_w = (2)\tilde{\psi} - \lambda(2)^r + (1)^r \left[ -2\lambda\partial_N(1)^r \tilde{\psi} - (\lambda\partial_N^2(1)\tilde{\psi} + 2(1)^r \partial_N + 2(1)\tilde{\psi}) \right] 
\[ - \Box_2(1)\mathcal{B}_w + \Box_2(1)\mathcal{B}, \] (57)

with \( \lambda = -(\partial_N\varphi_0)^{-1} \). Converting to \( \eta \) as time variable yields equation (30) in [30], when specialized to the long wavelength limit\(^\text{23}\). We note in passing that this transformation formula plays a central role in finding a conserved quantity and explicit solutions at second order, as discussed in the next section and in detail in the follow up papers [27] and [29], called UW3 and UW4, respectively, below.

\(^{20}\) Use \( (1)\rho_0 = (1)\rho + (1)V' + (1)H', \) introduce \( \rho = \rho/\rho_0 \), use \( 2S[V D_j, V] = V^2 \), replace \( 1 + c_s^2 \) using \( w' = 3H(1 + w)/(1 + w) - (1 + c_s^2) \), and assume conservation of energy \( \rho_0^f = -3H(1 + w)\rho_0 \).

\(^{21}\) To make the comparison note that \( H(1)^r \tilde{\psi}_w = H(1)^r - 3\rho_0 H^2(1 + c_s^2) \), and write the spatial derivative terms using our notation \( D_j, D^\gamma \) and \( S^\gamma \).

\(^{22}\) Here we have used equation (28a) for \( (2)^r \psi \) and equation (29) for \( C_\phi \), as well as the first order relation \( (1)\tilde{\psi}_w = (1)\psi - (1)^r \).

\(^{23}\) Note that \( H^2(\lambda c_s^2 \varphi_0 + 2) = \lambda c_s^2 + H' + 2\lambda c_s^2 \), where \( \lambda = -(\partial_N\varphi_0)^{-1} \) in our notation. However, Vernizzi uses the convention \( f = \partial_N \). In addition, since Vernizzi restricts consideration to long wavelength perturbations the terms \( \Box_2(\mathcal{B}_w) - \Box_2(\mathcal{B}) \), which are \( O(D^2) \), do not appear.
5. Discussion

The present paper is the first of four closely connected papers dealing with scalar perturbations up to second order. In the present paper, which we will refer to as UW1, we have introduced new second order variables and a new second order gauge vector field, which simplifies the change of gauge formulas at second order, and provides an efficient unifying algorithm for calculating any gauge invariant in the commonly used gauges. This is important since it is often desirable to use several gauges when addressing a given problem. In the second paper, called UW2 [28], we present five ready-to-use systems of governing equations for second order perturbations. These two papers constitute the foundation for subsequent physical applications, illustrated by UW3 [27] and UW4 [29].

In UW3 we use the new variables and gauge transformation formulas, and apply them to the equations given in UW2 to produce new dimensionless gauge-invariant conserved quantities and explicit general solutions, containing both the so-called growing and decaying modes, for second order perturbations in the super-horizon regime. This is made possible due to that the dimensionless source-compensated \( \hat{h} \) second order perturbation variables simplify the analysis of perturbations in the super-horizon regime in a number of ways. For example, in this regime the perturbed energy conservation equations can be written in the following form in terms of hatted variables:

\[
\partial_N \left( (1) \delta - 3(1)\psi \right) \approx -3(1)\Gamma, \tag{58a}
\]

\[
\partial_N \left( (2) \delta - 3(2)\psi \right) \approx -3(2)\Gamma - 2(1)\Gamma^2 + 2\partial_N (1)\Gamma (1)\delta, \tag{58b}
\]

with a simple quadratic source term in the second order equation (see UW3 [27]). Here \( (r)\Gamma, \ r = 1, 2 \), are the non-adiabatic pressure perturbations (see UW2 [28]). It follows from (58), specialized to the uniform curvature gauge \( (r)\psi = 0, \ r = 1, 2 \) that \( (1)\delta \) and the hatted second order perturbation \( (2)\delta \) are conserved for adiabatic perturbations \( (r)\Gamma = 0, \ r = 1, 2 \) in the super-horizon regime. Note, however, that the unhatted second order density perturbation \( (2)\delta \) is not conserved unless \( c_s^2 \) is constant, as follows from (37d). Another example is that when using the uniform curvature gauge in the super-horizon regime, the perturbed Einstein equations assume a particularly simple form when expressed in terms of hatted variables, which leads to further conserved quantities. In particular if the source is a minimally coupled scalar field with an arbitrary scalar field potential we obtain a new second order conserved quantity for the scalar field, which is used in UW4 [29] to obtain new physical results for scalar fields for second order perturbations in the long wavelength limit, without imposing the slow-roll approximation.

Appendix A. Relation with Malik and Wands [16]

We begin by listing the main differences between Malik and Wands [16] and the present paper. First, they give a more general framework for the gauge transformation formulas than we do in that they do not require the perturbations at first order to be purely scalar. Second, they do not use the gauge freedom to set \( C = 0, \ C_i = 0 \) as we have done (\( C \) and \( C_i \) are labeled \( E \) and \( F_i \) in their paper). Their generating vector field \( \xi \) at first order, given by their equation (6.17),

\[
\xi^\mu = (\alpha, \delta' \beta_1 + \gamma_1), \tag{A.1}
\]

is thus more general than ours, which is given by equation (19) (also note that they use subscripts to denote the order of the perturbation). As a result of these differences when comparing
equations it is necessary to set $E = 0$, $F = 0$, $S_i = 0$, $h_{ij} = 0$ at first order in the metric perturbations (their equations (2.8)–(2.12)), and to set $\beta_1 = 0$, $\gamma_1^i = 0$ in their generating gauge vector field. Then identify $\alpha_1 \equiv (1)\xi^N = H^{-1}(1)\xi^N$. Their generating gauge vector field at second order is general, $\xi_i^N = (\alpha_2, \partial^i \beta_2 + \gamma_2^i)$. In accordance with (19), we identify

$$\alpha_2 \equiv (2)\xi^N = H^{-1}(2)\xi^N, \quad \beta_2 \equiv (2)\xi, \quad \gamma_2^i = (2)\xi^i.$$  \hspace{1cm} (A.2)

The third difference is in the treatment of the density perturbation. Malik and Wands begin with the standard unscaled perturbation expansion (their equation (6.16)):

$$\rho = \rho_0 + \rho_1 + \frac{1}{2} \rho_2,$$  \hspace{1cm} (A.3)

and do not assume local energy conservation. We, on the other hand, in several equations assume local energy conservation and then define the scaled perturbations

$$\langle r \rangle \delta = \frac{\rho_r}{\rho_0 + \rho_0}, \quad r = 1, 2.$$  \hspace{1cm} (A.4)

However, normalizing with $\rho_0 + \rho_0$ is equivalent to normalizing with $-\rho_0^2/(3H)$, which suggests a natural generalization in the case local energy conservation does not hold.

The fourth difference is in the treatment of the velocity perturbation. Malik and Wands begin with the perturbation expansion (their equation (4.4)):

$$u^i = a^{-1}(v^i_1 + \frac{1}{2} v^i), \quad v^i_r = \partial^i v_r + \tilde{v}_r, \quad r = 1, 2.$$  \hspace{1cm} (A.5)

of the contravariant spatial components of the four-velocity, whereas we expand the covariant components as in (12), using $V$ instead of $v$ as the scalar perturbation of the velocity. It follows that $v$ and $V$ are related as follows\(^{24}\):

$$\begin{align*}
V_1 &= v_1 + B_1, \\
V_2 &= v_2 + B_2 - 2S[\phi D B_1 + 2\psi D V_1].
\end{align*}  \hspace{1cm} (A.6a)\hspace{1cm} (A.6b)$$

We are using the same letters for the scalar metric perturbations, namely $(\phi, B, \psi)$ as Malik and Wands. Their transformation laws for these variables are given by equations (6.37)–(6.39) at first order and (6.47), (6.51) and (6.58) at second order. In order to facilitate a comparison we write our formulas using the Malik and Wands variables and notation. For $(\phi)_i$, given by equation (40b), we use (A.2) to obtain

$$\tilde{\phi}_2 = \phi_2 + \phi_2' + H\alpha_2 + \alpha_1(\phi_i'' + 5H\phi_i' + (H' + 2H^2)\alpha_1 + 2\phi_i' + 4H\phi_i) + 2\alpha_1(\phi_i' + 2\phi_i).$$  \hspace{1cm} (A.7)

We have consistency with Malik and Wands equation (6.47). For $(\psi)_i$, given by equation (39a), we use (A.2) to obtain

$$\begin{align*}
\tilde{\psi}_2 &= \psi_2 - H\alpha_2 - \alpha_1 \left( H\alpha_i' + (H' + 2H^2)\alpha_1 - 2\psi_i' + 4H\psi_i \right) \\
&\quad + \frac{1}{3}(D^2 S^{ij} - \delta^{ij})[D_i(2B_1 - \alpha_1)D_j\alpha_1].
\end{align*}  \hspace{1cm} (A.8)$$

To obtain agreement with Malik and Wands we write their (6.58) in the form

$$\tilde{\psi}_2 = \psi_2 - H\alpha_2 + \frac{1}{6}(D^2 S^{ij} - \delta^{ij})\chi_{ij},$$  \hspace{1cm} (A.9)

\(^{24}\) Expand the relation $u_a = g_{ab} u^b$. Here we are using the Malik and Wands subscript convention to indicate the order of the perturbation.
where $\chi$ is given by (6.54) specialized to scalar perturbations at first order by setting $C_{ij} = -\psi \delta_{ij}$, $B_{ij} = \mathbf{D}_i \mathbf{B}$ and $\xi^i = 0$, which yields

$$
\chi = 2\alpha (H \alpha' + (H' + 2H^2) \alpha - 2\psi' - 4H\psi) \delta_{ij}
+ 2(\mathbf{D}_i B_j + \mathbf{D}_j B_i - \mathbf{D}_i \mathbf{D}_j \alpha_1).
$$

Substituting (A.10) in (A.9) yields (A.8).

For $\mathbf{B}_2$, given by equation (39b), we use (A.2) to obtain

$$
\tilde{B}_2 = B_2 - \alpha_2 + \beta_2^i = \alpha_1 (\alpha_1 + H \alpha_1) - 2S^i [-2\phi \mathbf{D}_i \alpha_1 + \alpha_1 \mathbf{D}_i (B_i' + H B_i) + (\alpha_1' + H \alpha_1) \mathbf{D}_i (B_1 - \alpha_1)],
$$

where we have introduced the Malik–Wands notation $\beta_2^i = (2 \xi)$, which is given by

$$
\beta_2 = -S^i \mathbf{D}_i (2B_1 - \alpha_1) \mathbf{D}_j \alpha_1.
$$

To obtain agreement with Malik and Wands we write their (6.27) and (6.28) as follows:

$$
\bar{B}_2 = B_2 - \alpha_2 + \beta_2 + S^i \chi_{Bi},
$$

where $\beta_2$ is given by (6.59) with the spatial gauge fixed so that $\bar{E}_2 = 0 = E_2$,

$$
\beta_2 = -\frac{1}{2} S^i \chi_{ij},
$$

and where $\chi_{Bi}$ is given by (6.49) specialized to scalar perturbations at first order by setting $B_{ij} = \mathbf{D}_i \mathbf{B}$ and $\xi^i = 0$, which yields

$$
\chi_{Bi} = -\mathbf{D}_i [\alpha_1 (\alpha_1' + H \alpha_1)] + 2[-2\phi \mathbf{D}_i \alpha_1 + \alpha_1 \mathbf{D}_i (B_i' + H B_i) + (\alpha_1' + H \alpha_1) \mathbf{D}_i (B_1 - \alpha_1)].
$$

Note that (A.14) with (A.10) yields (A.12). Substituting (A.15) in (A.13) yields (A.11).

We now consider the matter perturbations. For $\mathbf{V}$ the ‘Malik and Wands form’ of our equation (39c) is as follows:

$$
\mathbf{V}_2 = \mathbf{V}_2 - \alpha_2 - \alpha_1 (\alpha_1' + H \alpha_1) + 2S^i [-2\phi \mathbf{D}_i \alpha_1 + \alpha_1 \mathbf{D}_i (V_i' + H V_i)].
$$

To compare with Malik and Wands we need to use equations (A.6b) and (A.11) to obtain the transformation law for $v_2$:

$$
\tilde{v}_2 = v_2 - \beta_2^i + 2S^i [\alpha \mathbf{D}_i (v_i' - H v_i)],
$$

where $\beta_2$ is given by (A.12). This agrees with Malik and Wands equations (6.27) and (6.28) apart from a differing sign in (6.27)\(^25\).

Finally, for $\mathbf{\delta}$ the un-scaled form of our equation (39d) in Malik and Wands notation is as follows:

$$
\bar{\mathbf{\delta}}_2 = \rho_2 + \rho_0 \alpha_2 + \alpha_1 (\rho_1'' \alpha_1 + \rho_0 \alpha_1' + 2 \rho_1'),
$$

which agrees with Malik and Wands equation (6.20) after setting $\xi^i = 0$.

**Appendix B. Spatial differential operators**

In this appendix we introduce the spatial differential operators that are used in this paper. First, we require the spatial Laplacian and the trace-free symmetric second order derivative:

\(^{25}\) We find that after setting $\xi^i = 0$ (6.27) should read $\chi_{ij} = 2\alpha_1 (\alpha_1' + H v_i)$.
\[ D^2 := \gamma^i(D_i D)_i, \quad D_{ij} := D_i(D_j) - \frac{1}{3} \gamma_0 D^2. \]  \hfill (B.1)

We use these to define the scalar mode extraction operators (see [25], equation (85)) \(^{26}\).

\[ S^i = D^{-2}D^i, \quad S^{ij} = \frac{3}{2}(D^{-2})^2D^{ij}, \]  \hfill (B.2a)

where \(D^{-2}\) is the inverse spatial Laplacian. Note that \(S^i\) is the inverse operator of \(D_i\) and \(S^{ij}\) is the inverse operator of \(D_{ij}\):

\[ S^i D_i C = C, \quad S^{ij} D_{ij} C = C. \]  \hfill (B.2b)

We can now define the following spatial differential operators:

\[ (DC)^2 := \gamma^i(D_i(C)(D_iC)), \]  \hfill (B.3a)

\[ D_0(C) := S^{ij}(D_i(C)(D_jC)), \]  \hfill (B.3b)

\[ D_2(C) := \frac{1}{3} (D^2 D_0(C) - (DC)^2), \]  \hfill (B.3c)

which act on an arbitrary function \(C\).

The operator \((DC)^2\) is familiar, being the square of the magnitude of the gradient \(D_i C\), whereas the other two are less so. The expression \(D_0(C)\) can be viewed as the scalar mode of the rank two tensor \((D_i(C)(D_jC))\) while \(D_2(C)\) is defined by taking the Laplacian of \(D_0(C)\). We mention one property of these operators that suggests their physical role. When taking limits such as the long wavelength limit and the Newtonian limit it is essential to count how quantities change under a scaling of the spatial coordinates. More precisely, if some expression \(L(D_i)\) involving \(D_i\) scales as \(L(\lambda D_i) = \lambda^p L(D_i)\) under a rescaling of spatial coordinates \(x' \rightarrow \lambda^{-1} x'\), we say that \(L(D_i)\) has weight \(p\) in \(D_i\) \(^{27}\). It follows from equations (B.2a) and (B.3) that \(D_0(C)\) has weight 0 while \(D_2(C)\) has weight 2 in \(D_i\). We thus expect that \(D_0(C)\) will be dominant and \(D_2(C)\) will be negligible in the long wavelength limit, while the reverse will be true in the Newtonian limit.

In this paper the operators \(D_0\) and \(D_2\) serve to simplify the quadratic source terms in the gauge change formulas at second order. They play a similar role in other source terms, for example the source terms of the perturbed Einstein tensor, and hence in the solutions of the perturbed Einstein equations at second order. The fact that these operators occur frequently motivates our choice of notation: the symbol \(D\) suggests a spatial differential operator acting on an arbitrary function \(C\), while the subscript 0 or 2 indicates the weight in \(D_i\). The use of \(D_{ij}\) in general, rather than \(i\), also helps to clarify the structure of the spatial derivative terms.

It turns out that the operators \(D_0\) and \(D_2\), as defined in (B.3), are related to two quantities, denoted \(\Psi_0\) and \(\Theta_0\), that have been used in the literature on second order perturbations since 1997. These quantities were defined in the case of a flat background by Mollerach and Matarrese \([19]\) as follows \(^{28}\):

\[ \Psi_0 := \frac{1}{2} D^{-2} \left( D^i D^j C D_i D_j C - (D^2 C)^2 \right), \]  \hfill (B.4a)

---

\(^{26}\) If the background is not flat, \(S^{ij} = \frac{3}{2} D^{-2}(D^2 + K)^{-1}D^{ij}\).

\(^{27}\) Note that \(D^2\) and \(D_{ij}\) both have weight 2 while in contrast \(S^i\) and \(S^{ij}\) have weights \(-1\) and \(-2\), respectively.

\(^{28}\) See equation (3.7) for \(\Psi_0\) and equation (3.11) for \(\Theta_0\) in \([19]\). See also Materrese et al \([18]\), equations (4.36) and (6.6).
\[
\Theta_0 := D^{-2} \left( \Psi_0 - \frac{1}{3} (DC)^2 \right),
\]
where \( C \) is a function that determines the spatial dependence of the perturbations. The explicit relations are simple, namely
\[
\Theta_0 = -\frac{1}{3} D_0(C), \quad \Psi_0 = -D_2(C),
\]  
but their derivation requires the use of some complicated identities satisfied by \( D_i \), as follows.

Expand the second derivatives on the left sides to get
\[
D_i D_j (D_iC D_jC) = 2 (D_iC) D_i (D_2C) + (D_iD_jC)(D_iD_jC) + (D_2^2 C)^2,
\]

which is the key identity. The desired relations (B.5) follow immediately from (B.7) on using the definitions (B.3) and (B.4).

Although \( \Theta_0 \) and \( \Psi_0 \) were first introduced to describe second order perturbations of the Einstein-de Sitter universe, since 2005 they have also been used to describe perturbed \( \Lambda \)CDM universes. See, for example, Bartolo et al [2] following equation (9), Tomita [23], equation (2.11), Villa and Rampf [31], equations (5.21) and (5.38), and Tram et al [24], equations (D.11) and (D.12). In these references one sees that \( \Theta_0 \) (i.e. \( D_0(C) \)) contributes to perturbations on super-horizon scales, while \( D_2^2 \Psi_0 \) (i.e. \( D^2 D_2(C) \)) contributes to the Newtonian part of the second order density perturbation. As mentioned earlier, this physical interpretation is a consequence of the fact that \( D_0(C) \) is of weight zero in \( D_i \) while \( D_2(C) \) is of weight two and \( D^2 D_2(C) \) is of weight four. The important point, however, is that \( D_0(C) \) and \( D_2(C) \) are not restricted in use to the \( \Lambda \)CDM universes: they arise in the general change of gauge formulas in this paper, in the source terms of the second order perturbations of the Einstein tensor, and \( D_0(C) \) contributes to perturbations on super-horizon scales in general. For example, \( D_0(\psi_p) \) contributes to the CMB anisotropy at second order on large scales.

ORCID iDs

John Wainwright @ https://orcid.org/0000-0002-3452-0907

References

[1] Bartolo N, Komatsu E, Matarrese S and Riotto A 2004 Non-Gaussianity from inflation: theory and observations Phys. Rep. 402 103–266
[2] Bartolo N, Matarrese S and Riotto A 2005 Signatures of primordial non-Gaussianity in the large-scale structure of the universe J. Cosmol. Astropart. Phys. JCAP10(2005)010

29 In a suggestion of physical interpretation Villa and Rampf [31] (see page 15 following equation (5.39)) refer to \( \Theta_0 \) as the ‘GR kernel’ and to \( \Psi_0 \) as the ‘Newtonian kernel’.
30 See for example Bartolo et al [3], equation (53).
[3] Bartolo N, Matarrese S and Riotto A 2010 Non-Gaussianity and the cosmic microwave background anisotropies Adv. Astron. 2010 157079
[4] Bruni M, Matarrese S, Mollerach S and Sonego S 1997 Perturbations of spacetime: Gauge transformations and gauge invariance at second order and beyond Class. Quantum Grav. 14 2585
[5] Carrilho P and Malik K A 2015 Vector and tensor contributions to the curvature perturbation at second order J. Cosmol. Astropart. Phys. JCAP02(2016)021
[6] Christopherson A J, Malik K A, Matravers D R and Nakamura K 2011 Comparing different formulations of nonlinear perturbation theory Class. Quantum Grav. 28 225024
[7] Christopherson A J, Naslon E and Malik K A 2015 A short note on the curvature perturbation at second order Class. Quantum Grav. 32 075005
[8] Dias M, Elliston J, Frazer J, Mulryne D and Seery D 2015 The curvature perturbation at second order J. Cosmol. Astropart. Phys. JCAP02(2015)040
[9] Gressel H A and Bruni M 2018 fNL – gNL mixing in the matter density field at higher orders J. Cosmol. Astropart. Phys. JCAP06(2018)016
[10] Hidalgo J C, Christopherson A J and Malik K A 2013 The Poisson equation at second order in relativistic cosmology J. Cosmol. Astropart. Phys. JCAP11(2005)005
[11] Kodama H and Sasaki M 1984 Cosmological perturbation theory Prog. Theor. Phys. Suppl. 78 1–166
[12] Liddle A R and Lyth D H 2000 Cosmological Inflation and Large-Scale Structure (Cambridge: Cambridge University Press)
[13] Malik K A and Wands D 2009 Cosmological perturbations Phys. Rep. 475 1–51
[14] Martin J and Ringeval C 2006 Inflation after WMAP3: confronting the slow-roll and exact power spectra with CMB data J. Cosmol. Astropart. Phys. JCAP08(2006)009
[15] Matarrese S, Mollerach S and Bruni M 1998 Relativistic second-order perturbations of the Einstein–de Sitter universe Phys. Rev. D 58 043504
[16] Mollerach S and Matarrese S 1997 Cosmic microwave background anisotropies from second order gravitational perturbations Phys. Rev. D 56 4494
[17] Nakamura K 2007 Second order gauge invariant cosmological perturbation theory Prog. Theor. Phys. 117 17
[18] Noh H and Hwang J C 2004 Second order perturbations of the Friedmann world model Phys. Rev. D 69 104011
[19] Salopek D S and Bond J R 1990 Nonlinear evolution of long-wavelength metric fluctuations in inflationary models Phys. Rev. D 42 3936
[20] Tomita K 2005 Relativistic second-order perturbations of nonzero-λ flat cosmological models and CMB anisotropies Phys. Rev. D 71 083504
[21] Tram T, Fidler C, Crittenden R, Koyama K, Pettinari G W and Wands D 2016 The intrinsic matter bispectrum in ΛCDM J. Cosmol. Astropart. Phys. JCAP08(2016)058
[22] Uggla C and Wainwright J 2013 Asymptotic analysis of perturbed dust cosmologies to second order Gen. Relativ. Gravit. 45 1467
[23] Uggla C and Wainwright J 2014 Second order density perturbations for dust cosmologies Phys. Rev. D 90 043511
[24] Uggla C and Wainwright J 2019 Second order cosmological perturbations: conserved quantities and explicit solutions at large scales (unpublished)
[25] Uggla C and Wainwright J 2018 Second order cosmological perturbations: dynamics Phys. Rev. D 98 103534
[26] Vernizzi F 2005 On the conservation of second-order cosmological perturbations in a scalar field dominated universe Phys. Rev. D 71 06130
[27] Villa E and Rampf C 2016 Relativistic perturbations in ΛCDM: Eulerian and Lagrangian approaches J. Cosmol. Astropart. Phys. JCAP01(2016)030