CHAPTER 5

CLASSICAL INFLATIONARY PERTURBATIONS

> That's all I can tell you. Once you get into cosmological shit like this, you got to throw away the instruction manual.
>
> Stephen King, *It*

Although the inflationary paradigm was originally formulated as a solution to the flatness and horizon problems, these are not its most important consequences. As we will discover in the following chapters, the background dynamics of the inflaton field sources quantum fluctuations and generates macroscopic cosmological perturbations seeding the formation of *all* the structures in the Universe. Not bad, isn't it?

The fluctuations of the inflaton field ϕ inevitably source the metric perturbations. This means that in order to compute the inflationary perturbations, either classical or quantum, we need to consider the full inflaton-gravity system

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2\kappa^2}R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi - V(\phi)\right] . \tag{5.1}$$

Fortunately for us, the measured CMB fluctuations are small enough as to justify a linearized analysis. Even in this case, the detailed computation of the primordial density perturbation spectrum can be rather involved. The subtleties and complications are associated to the fact that we are dealing with a diffeomorphism invariant theory. In particular, the inflaton perturbations $\delta\phi(\vec{x},t)$ following from a naive splitting $\phi = \phi(t) + \delta\phi(\vec{x},t)$ are not gauge/diffeomorphism invariant. This can be easily seen by performing a temporal shift $t \to t + \delta t$ with infinitesimal δt. Under this coordinate transformation the perturbation $\delta\phi(\vec{x},t)$ does not transform as a scalar but rather shifts to a new value $\delta\phi \to \delta\phi - \dot{\phi}(t)\delta t$. If not taken into account this gauge-dependence can completely jeopardize the computation of the inflationary spectrum. The standard formulation of inflation in Eq. (5.1) is certainly aesthetical but not particularly convenient for the problem at hand. In particular, one cannot easily see how the ten metric components talk to each other or distinguish proper degrees of freedom from simple redundancies. The identification of physical degrees of freedom and the computation of inflationary perturbations becomes particularly simple in the so-called Arnowitt-Deser-Misner (ADM) formalism. In the next section, we describe this framework in detail.

5.1 The ADM formalism

The ADM formalism was introduced by Arnowitt, Deser and Misner an an attempt to formulate General Relativity in terms of Hamiltonian mechanics. In this formalism, a general 4-dimensional (globally hyperbolic)[1] manifold M with metric $g_{\mu\nu}$ is foliated into a family of non-intersecting spacelike hypersurfaces Σ labelled by a "time" coordinate $t \in \mathbb{R}$

$$M \to \mathbb{R} \times \Sigma \,. \tag{5.2}$$

From the mathematical point of view, this decomposition breaks the symmetry between space and time and allows to formulate the resolution of Einstein equations as an initial value problem with constraints. Although the foliation (5.2) may appear to break diffeomorphism invariance, this is not the case due to the arbitrariness in the choice of the coordinate time t.

Each spatial slice is equipped with its own Riemannian structure. The induced metric $\gamma_{\mu\nu}$ on Σ can be uniquely determined by the conditions

$$\gamma_{\mu\nu} n^\mu = 0\,, \qquad\qquad \gamma_{\mu\nu} s^\mu = g_{\mu\nu} s^\mu\,, \tag{5.3}$$

with n^μ *a* normal vector to the hypersurface and s^μ *any* tangent vector to it. The normal and tangent vectors are normalized by the conditions

$$g_{\mu\nu} n^\mu n^\nu = -1\,, \qquad\qquad g_{\mu\nu} s^\mu n^\nu = 0\,. \tag{5.4}$$

Taking into account these expressions we can rewrite $\gamma_{\mu\nu}$ and its inverse as

$$\gamma_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu\,, \qquad\qquad \gamma^{\mu\nu} = g^{\mu\nu} + n^\mu n^\nu\,. \tag{5.5}$$

Note that even though $\gamma_{\mu\nu}$ is a metric on 3-dimensional space, the components γ_{00} and γ_{0i} are generically non-zero since $g_{0i} \neq 0$.

In order two connect the spatial coordinates in two different slices, we introduce a set of curves intersecting them and use t as the affine parameter along the curves. Note that we do not require the curves to be geodesics or orthogonal to the spatial hypersurfaces since this would be over-restrictive. The (normalized) vector field $t^\mu = \partial x^\mu/\partial t$ defining the direction of time derivatives,

$$t = t^\mu \nabla_\mu\,, \qquad\qquad t^\mu \nabla_\mu t = 1\,, \tag{5.6}$$

can be decomposed into a spatial and a normal part by introducing a *shift vector* $N^\mu \equiv \gamma^{\mu\nu} t_\nu$ and a *lapse function* $N \equiv -n_\mu t^\mu$,[2]

$$t^\mu = N n^\mu + N^\mu\,. \tag{5.7}$$

Taking into account that in a coordinate basis x^μ with $t^\mu \nabla_\mu = \partial/\partial t$ the infinitesimal displacement dx^μ takes the form

$$dx^\mu = t^\mu dt + dx^i = (N dt) n^\mu + (dx^i + N^i dt)\,, \tag{5.8}$$

[1] The name globally hyperbolic stems from the fact that the scalar wave equation is well posed. Apart from extreme regions such as the centers of black holes, spacetimes are generally globally hyperbolic.

[2] In numerical relativity, the lapse and shift functions are usually denoted by α and β^i. We follow here the notation of the orginal ADM paper.

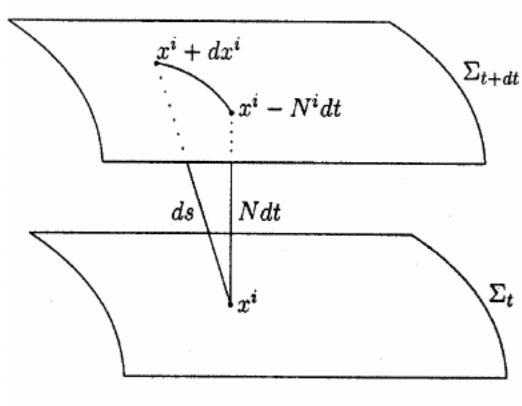

Figure 5.1: The ADM decomposition.

we can rewrite the spacetime line element $\mathrm{d}s^2 \equiv g_{\mu\nu}dx^\mu dx^\nu$ as (see Fig. 5.1)

$$\mathrm{d}s^2 = -N^2\mathrm{d}t^2 + \gamma_{ij}(\mathrm{d}x^i + N^i\mathrm{d}t)(\mathrm{d}x^j + N^j\mathrm{d}t)\,, \tag{5.9}$$

with $\gamma_{ij} = g_{ij}$ and the latin indices denoting spatial components. The space-time geometry is therefore described by the spatial geometry of slices, encoded in γ_{ij}, together with the deformations of neighboring slices with respect to each other, encoded in N and N^i. In particular, the lapse function measures proper time between two adjacent hypersurfaces while the shift function relates spatial coordinates between them. Given Eq. (5.9) we can easily compute the covariant and contravariant components of the metric in terms of γ^{ij}, N and N^i as well as the metric determinant. We obtain

$$g_{00} = \gamma_{ij}N^iN^j - N^2\,, \qquad g_{0i} = \gamma_{ij}N^j\,, \qquad g_{ij} = \gamma_{ij}\,, \tag{5.10}$$

$$g^{00} = -N^{-2}\,, \qquad g^{0i} = N^{-2}N^i\,, \qquad g^{ij} = \gamma^{ij} - N^{-2}N^iN^j\,, \tag{5.11}$$

$$\sqrt{-g} = N\sqrt{\gamma}\,. \tag{5.12}$$

5.1.1 Intrinsic and extrinsic curvatures

As General Relativity is based on the concept of curvature, it is important to analyze curvature in the language of the 3+1 ADM decomposition.

The (intrinsic) curvature of the 3-dimensional hypersurface Σ can be defined and computed using the standard methods. In particular, the spatial metric $\gamma_{\mu\nu}$ allows to construct a covariant derivative D_σ on Σ such that

$$D_\sigma\gamma_{\mu\nu} = 0\,. \tag{5.13}$$

The covariant derivative D_ρ can be written in terms of the covariant derivative ∇_σ constructed out of $g_{\mu\nu}$. For a general tensor $T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}$ we have

$$D_\rho T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = (\gamma^{\mu_1}{}_{\rho_1}\cdots\gamma^{\mu_k}{}_{\rho_k}\gamma_{\nu_1}{}^{\kappa_1}\cdots\gamma_{\nu_l}{}^{\kappa_k})\gamma_\rho{}^\lambda\,\nabla_\lambda T^{\rho_1\cdots\rho_k}{}_{\kappa_1\cdots\kappa_k}\,, \tag{5.14}$$

with

$$\gamma^\mu{}_\rho = \delta^\mu{}_\rho + n^\mu n_\rho \tag{5.15}$$

the natural projection tensor onto Σ. The commutator of two covariant derivatives $[D_\mu, D_\nu]$ acting on any spatial form V_σ ($V_\sigma n^\sigma = 0$) defines the 3-dimensional Riemann tensor

$$[D_\mu, D_\nu]V_\rho \equiv {}^{(3)}R_{\mu\nu\rho}{}^\sigma V_\sigma \,. \tag{5.16}$$

This tensor measure measures the change of a vector on Σ when it is transported around a close loop. The indices μ and ν in the commutator define the "direction" of the loop. The Ricci tensor ${}^{(3)}R_{\mu\nu}$ and the Ricci scalar ${}^{(3)}R$ are obtained by performing the standard non-vanishing contractions,

$${}^{(3)}R_{\mu\nu} = {}^{(3)}R^\rho{}_{\mu\rho\nu}\,, \qquad {}^{(3)}R = \gamma^{\mu\nu}\,{}^{(3)}R_{\mu\nu}\,. \tag{5.17}$$

All the objects defined till now are *intrinsic* quantities describing the properties of the space-like hypersurfaces Σ. To recover the full spacetime information, we need to describe how the Σ hypersurfaces are embedded into the 4-dimensional geometry, i.e. how they bend inside M. As we need to "move off" the Σ hypersurface to detect the embedding, the bending of Σ cannot be captured by intrinsic objects defined on it. We are led therefore to define the *extrinsic curvature* tensor as

$$K_{\mu\nu} \equiv D_\mu n_\nu = \gamma^\rho{}_\mu \gamma^\sigma{}_\nu \nabla_\rho n_\sigma = \gamma^\rho{}_\mu \nabla_\rho n_\nu \,. \tag{5.18}$$

with n^μ the normal vector to Σ. The last equality in (5.18) follows from writing $\gamma^\sigma{}_\nu = \delta^\sigma{}_\nu + n^\sigma n_\nu$ and taking into account that

$$n_\sigma n^\sigma = -1 \qquad \longrightarrow \qquad n^\sigma \nabla_\rho n_\sigma = 0\,. \tag{5.19}$$

Physically, the extrinsic curvature $K_{\mu\nu}$ measures the change of this vector along Σ or if you prefer the curvature of the slice relative to the enveloping 4-geometry.

Some terminology

The induced metric and the extrinsic curvature tensor are also known as the 1st and 2nd fundamental forms of Σ.

The extrinsic curvature has some interesting properties:

1. It is symmetric

$$K_{\mu\nu} = K_{\nu\mu}\,. \tag{5.20}$$

   This can be easily proved by rewriting Eq. (5.18) as

$$K_{\mu\nu} = \gamma^\rho{}_\mu \nabla_\rho n_\nu = \nabla_\mu n_\nu + n^\rho n_\mu \nabla_\rho n_\nu \,, \tag{5.21}$$

   and expanding the second term using the Frobenius theorem

$$n_{[\mu} \nabla_\nu n_{\rho]} = 0\,, \tag{5.22}$$

to get

$$K_{\mu\nu} = \nabla_\nu n_\mu + n^\rho n_\nu \nabla_\rho n_\mu = \gamma^\rho{}_\nu \nabla_\rho n_\mu = K_{\nu\mu} \,. \tag{5.23}$$

This also implies that all spatial projections of $\nabla_\mu n_\nu$ are symmetric and therefore that $\nabla_\mu n_\nu = \nabla_\nu n_\mu$.

2. Since $K_{\mu\nu} n^\mu = K_{\mu\nu} n^\nu = 0$ and $n_i = 0$, we can conclude that $K^{0\mu} = 0$ and that the contravariant component $K^{\mu\nu}$ are also purely spatial.

3. Using the symmetry of $K_{\mu\nu}$ we can write $K_{\mu\nu} = \frac{1}{2}(K_{\mu\nu} + K_{\nu\mu})$. Combining this with Eqs. (5.15), (5.19) and taking into account metric compatibility, $\nabla_\rho g_{\mu\nu} = 0$, we can express the extrinsic curvature as

$$\begin{aligned} K_{\mu\nu} &= \frac{1}{2}\left(\gamma^\rho{}_\mu \nabla_\rho n_\nu + \gamma^\rho{}_\nu \nabla_\rho n_\mu\right) \\ &= \frac{1}{2}\left(n^\rho n_\mu \nabla_\rho n_\nu + n^\rho n_\nu \nabla_\rho n_\mu + \nabla_\mu n_\nu + \nabla_\nu n_\mu\right) \\ &= \frac{1}{2}\left(n^\rho \nabla_\rho (n_\mu n_\nu) + g_{\rho\nu} \nabla_\mu n^\rho + g_{\mu\rho} \nabla_\nu n^\rho\right) \\ &= \frac{1}{2}\left(n^\rho \nabla_\rho \gamma_{\mu\nu} + \gamma_{\rho\nu} \nabla_\mu n^\rho + \gamma_{\mu\rho} \nabla_\nu n^\rho\right) \,. \end{aligned} \tag{5.24}$$

The last term in this equation is proportional to the Lie derivative [3] $\mathcal{L}$ of the intrinsic metric $\gamma_{\mu\nu}$ along the unit normal

$$K_{\mu\nu} = \frac{1}{2}\left(n^\rho \nabla_\rho \gamma_{\mu\nu} + \gamma_{\rho\nu} \nabla_\mu n^\rho + \gamma_{\mu\rho} \nabla_\nu n^\rho\right) \equiv \frac{1}{2} \mathcal{L}_n \gamma_{\mu\nu} \,. \tag{5.25}$$

4. Using Eq. (5.25) we can write

$$\begin{aligned} K_{\mu\nu} &= \frac{1}{2}\left(n^\rho \nabla_\rho \gamma_{\mu\nu} + \gamma_{\rho\nu} \nabla_\mu n^\rho + \gamma_{\mu\rho} \nabla_\nu n^\rho\right) \\ &= \frac{1}{2N}\left[N n^\rho \nabla_\rho \gamma_{\mu\nu} + \gamma_{\rho\nu} \nabla_\mu (N n^\rho) + \gamma_{\mu\rho} \nabla_\nu (N n^\rho)\right] \\ &= \frac{1}{2N} \gamma_\mu{}^\rho \gamma_\nu{}^\sigma \mathcal{L}_{t-N} \gamma_{\rho\sigma} \\ &= \frac{1}{2N} \gamma_\mu{}^\rho \gamma_\nu{}^\sigma \left(\mathcal{L}_t \gamma_{\rho\sigma} - \mathcal{L}_N \gamma_{\rho\sigma}\right) \,, \end{aligned} \tag{5.26}$$

(5.27)

[3] The Lie derivative is a geometrical generalization of the directional derivative. For a scalar function ϕ in a manifold with connection ∇_μ, it is given by

$$\mathcal{L}_X \phi = X^\mu \nabla_\mu \phi = x^\mu \partial_\mu \phi \,.$$

For a contravariant vector field V^ν the Lie derivative is given by the commutator

$$\mathcal{L}_X V^\nu = X^\mu \nabla_\mu V^\nu - V^\mu \nabla_\mu X^\nu = [X, V]^\nu \,.$$

For a covariant vector field V_ν we rather have

$$\mathcal{L}_X V_\nu = X^\mu \nabla_\mu V_\nu + V_\nu \nabla_\mu X^\nu \,.$$

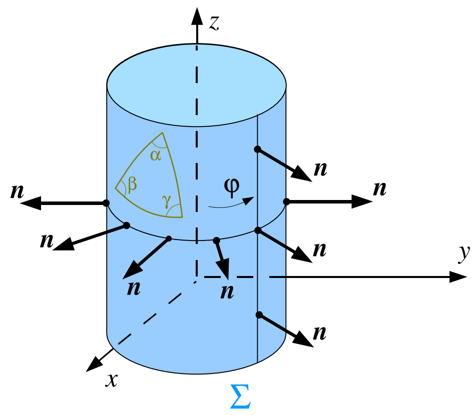

Figure 5.2: A cylinder Σ as a hypersurface in the Euclidean space $\mathbf{R}^3$. The unit normal vector n^i stays constant when z varies at fixed x and y, whereas its direction changes as x and y vary at fixed z. Consequently the extrinsic curvature of Σ vanishes in the z direction, but it is non zero in the other directions.

where we have substituted $Nn^\mu = t^\mu - N^\mu$ a and smuggled in inoffensive projections $\gamma_\mu{}^\rho \gamma_\nu{}^\sigma$ (remember that $K_{\mu\nu}$ is only spatial). Taking into account that $\mathcal{L}_N \gamma_{\mu\nu} = D_\mu N_\nu + D_\nu N_\mu$ together with the fact that the time derivative of a tensor field is defined as the Lie derivative along the time-evolution vector field t, $\dot{\gamma}_{\mu\nu} = \gamma_\mu{}^\rho \gamma_\nu{}^\sigma \mathcal{L}_t \gamma_{\rho\sigma}$, we can recast Eq. (5.26) in the form

$$K_{\mu\nu} = \frac{1}{2N} \left(\dot{\gamma}_{\mu\nu} - D_\mu N_\nu - D_\nu N_\mu \right) . \tag{5.28}$$

5. The trace of the extrinsic curvature is equal to the space-time divergence of the normal vector,

$$K \equiv g^{\mu\nu} K_{\mu\nu} = \gamma^{\mu\nu} K_{\mu\nu} = \nabla_\sigma n^\sigma . \tag{5.29}$$

A workout example: a cylinder in $\mathbf{R}^3$

Consider a cylinder Σ as a hypersurface in the Euclidean space $\mathbf{R}^3$. In cylindrical coordinates $(x^i) = (\rho, \varphi, z)$, the associated line element reads

$$\gamma_{ij}\, dx^i\, dx^j = R^2 d\varphi^2 + dz^2 \,, \tag{5.30}$$

with $R =$ constant the radius of the cylinder. As shown in Fig. 5.2, the unit normal vector n^i stays constant when z varies at fixed x and y, whereas its direction changes as x and y vary at fixed z. Consequently the extrinsic curvature of Σ should be expected to vanish in the z direction, but to be non-zero in the other directions. To evaluate the extrinsic curvature of Σ, we consider the unit normal n^i to Σ,

$$n^i = \left(\frac{x}{\sqrt{x^2+y^2}},\ \frac{y}{\sqrt{x^2+y^2}},\ 0 \right) , \tag{5.31}$$

and compute its divergence,

$$\nabla_j n^i = (x^2+y^2)^{-3/2} \begin{pmatrix} y^2 & -xy & 0 \\ -xy & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{5.32}$$

The trace of this quantity is different from zero

$$K = \frac{1}{R}. \tag{5.33}$$

We conclude that, although Σ is an intrinsically flat plane, its immersive image, the cylinder, has an extrinsic curvature, as intuitively expected.

5.1.2 The Gauss-Codazzi equations

The curvature invariants intrinsic to the hypersurfaces Σ together with the extrinsic curvature encode the whole information in M. By performing pure spatial projections of the indices in the 4-dim Riemann tensor we obtain the *Gauss equation*

$$\gamma_\mu{}^\alpha \gamma_\nu{}^\beta \gamma_\rho{}^\delta R^\sigma_{\alpha\beta\delta} = {}^{(3)}R_{\mu\nu\rho}{}^\sigma + K_{\mu\rho} K_\nu{}^\sigma - K_{\nu\rho} K_\mu{}^\sigma \,. \tag{5.34}$$

Making 3 spatial projections and one timelike projection we obtain the *Codazzi equation* (also called *Codazzi-Mainardi relation* in the mathematical literature)

$$\gamma^\mu{}_\alpha \gamma^\nu{}_\beta \gamma^\rho{}_\delta R_{\mu\nu\rho\sigma} n^\sigma = D_\alpha K_{\beta\delta} - D_\beta K_{\alpha\delta} \,. \tag{5.35}$$

Finally, taking 2 timelike projections we obtain the *Ricci equation*[4]

$$R_{\mu\rho\nu\sigma} n^\rho n^\sigma = n^\sigma (\nabla_\mu \nabla_\rho - \nabla_\rho \nabla_\mu) n_\nu = -\mathcal{L}_n K_{\mu\nu} + K_{\mu\rho} K^\rho_\nu + D_{(\mu} a_{\nu)} + a_\mu a_\nu \,, \tag{5.36}$$

with $a^\mu \equiv n^\rho \nabla_\rho n^\mu$ the normal acceleration[5] and the parenthesis in $D_{(\mu} a_{\nu)}$ denoting symmetrization. The above equations constitute all the possible projections of the 4-dimensional Riemann tensor. Indeed a projection involving three times n^μ vanishes identically due to the partial antisymmetry of $R_{\mu\rho\nu\sigma}$.

[4] Not to be confused with the *Ricci identity* (5.16).

[5] Since n^μ is a timelike unit vector, it can be regarded as the 4-velocity of some observer.

The Gauss, Codazzi and Ricci equations allow us to express the 4-dimensional Ricci scalar,

$$R = g^{\mu\nu}g^{\rho\sigma}R_{\mu\nu\rho\sigma} = (\gamma^{\mu\nu} - n^\mu n^\nu)(\gamma^{\rho\sigma} - n^\rho n^\sigma)R_{\mu\nu\rho\sigma} = \gamma^{\mu\nu}\gamma^{\rho\sigma}R_{\mu\nu\rho\sigma} - 2R_{\mu\nu}n^\mu n^\nu\,, \quad (5.37)$$

as

$$R = {}^{(3)}R + K_{\mu\nu}K^{\mu\nu} - K^2 - 2\nabla_\mu v^\mu\,, \quad (5.38)$$

where we have made use of the relation

$$R_{\mu\nu}n^\mu n^\nu = K^2 - K_\mu{}^\nu K_\nu{}^\mu + \nabla_\mu v^\mu\,, \quad (5.39)$$

and defined $v^\mu \equiv -n^\mu \nabla_\rho n^\rho + a^\mu$. Equation (5.38) contains a quadratic piece and a total divergence that can be explicitly cancelled at the level of the action by introducing the famous Gibbons-Hawking-York boundary term.

5.2 Inflaton-gravity action in the ADM formalism

After all these preliminaries, let us now come back to the inflaton-gravity action (5.1) and express it in terms of γ_{ij}, N and N^i. Combining the quadratic piece in Eq. (5.38) with the straightforward expansions of the metric determinant and the inflaton kinetic term,

$$\sqrt{-g} = N\sqrt{\gamma}\,, \qquad -g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi = \frac{1}{N^2}(\dot\phi - N^i\partial_i\phi)^2 - \partial^i\phi\partial_i\phi \equiv (\Pi_\phi)^2 - \partial^i\phi\partial_i\phi\,, \quad (5.40)$$

we can rewrite the action (5.1) as

$$S = \int d^4x N\sqrt{\gamma}\Big[\frac{1}{2\kappa^2}\left({}^{(3)}R + K_{ij}K^{ij} - K^2\right) + \frac{1}{2}[(\Pi^\phi)^2 - \partial^i\phi\partial_i\phi] - V(\phi)\Big]\,, \quad (5.41)$$

where we have explicitly taken into account that the extrinsic curvature tensor $K_{\mu\nu}$ is a purely spatial tensor,

$$K_{\mu\nu}K^{\mu\nu} = K_{ij}K^{ij}\,, \qquad K = g^{\mu\nu}K_{\mu\nu} = g^{ij}K_{ij}\,. \quad (5.42)$$

Equation (5.41) is a well-defined action from the point of view of the variational principle since, contrary to the original action (5.1), it only contains first time derivatives. The structure of K_{ij} allows us to interpret $K_{ij}K^{ij} - K^2$ as a sort of "kinetic term" governing the dynamics of γ_{ij} in a "potential" ${}^{(3)}R$. Note also that the action (5.41) does not contain any (time) derivatives for N and N^i, meaning that the lapse and shift functions are just Lagrange multipliers. The absence of derivatives for N and N^i is not accidental, but rather a consequence of diffeomorphism invariance. The lapse and shift are just *labels* defining the spatial hypersurface at the next instant of time.

The variation of the action with respect to N^i and N gives rise to the *momentum and energy (or Hamiltonian) constraints*

$$\nabla_j[K_i^j - \delta_i^j K] = \kappa^2 T^0{}_i\,, \qquad R^{(3)} - K_{ij}K^{ij} + K^2 = 2\kappa^2 T_{00}\,, \quad (5.43)$$

with

$$T_{00} = \frac{1}{2}(\Pi_\phi)^2 + \frac{1}{2}\partial^i\phi\partial_i\phi + V\,, \qquad T^0{}_i = \Pi^\phi\partial_i\phi\,. \quad (5.44)$$

These equations are nonlinear elliptic partial differential equations (hence not containing time derivatives) to be satisfied everywhere on the spatial hypersurface Σ. They are the necessary and sufficient *integrability conditions* for the embedding of the spacelike hypersurfaces in the 4-dimensional spacetime.

Gravity vs. electromagnetism

Due to the parallelism between gravity and electrodynamics people usually talk about *geometrodynamics* when referring to this section's content. In both cases the field equations can be separated into constraint equations and dynamic equations. In the Maxwell's theory the constraints on the electric and magnetic fields are embodied in the $\nabla \cdot \vec{E} = \rho$ and $\nabla \cdot \vec{B} = 0$ equations. In gravity they are given by the momentum and Hamiltonian constraints in (5.43). In both cases, any field configuration satisfying the constraint equations alone represents a valid solution. In order to find a solution of the equations we must first find an initial solution of the constraint equations and then evolve it using the dynamic equations. That is precisely what we will do in the Section 5.3.

In an unperturbed Universe with

$$N = N^{(0)} = 1\,, \qquad N_i^{(0)} = 0\,, \qquad \gamma_{ij} = a^2\delta_{ij}\,, \qquad \phi = \phi(t)\,, \tag{5.45}$$

the momentum constraint vanishes trivially due to the underlying isotropy of the zeroth order background. On the other hand, the energy constraint reduces to the Friedman equation (4.5),

$$H^2 = \frac{\kappa^2}{3}\left(\frac{1}{2}\phi^2 + V\right)\,. \tag{5.46}$$

5.2.1 Scalar-vector-tensor decomposition

Since we are interested in perturbations around homogeneous and isotropic FLRW cosmologies, it is convenient to decompose the quantities N_i and γ_{ij} in Eq. (5.9) into irreducible representations of the 3-dimensional rotation group. The relevant representations are scalars, vectors and tensors, so the following procedure is called scalar-vector-tensor decomposition or SVT decomposition. As it is well known, one can always decompose a vector into its longitudinal and transverse parts. This allows us to rewrite the shift function N^i as the gradient of a scalar plus a vector that is divergence free (div-free), namely

$$N_i = \partial_i\psi + \bar{N}_i\,, \tag{5.47}$$

with $\nabla_i\bar{N}^i = 0$. Repeating this procedure for each of the indices in the rank-2 tensor γ_{ij} we get

$$\begin{aligned} \gamma_{ij} &= e^{2\zeta}a^2\left[\delta_{ij} + h_{ij} + \left(\frac{1}{2}\left(\nabla_i\nabla_j + \nabla_j\nabla_i\right) - \frac{1}{3}\delta_{ij}\nabla^2\right)E + \nabla_iF_j + \nabla_jF_i\right] \\ &= e^{2\zeta}a^2\left[\delta_{ij} + h_{ij} + \left(\nabla_i\nabla_j - \frac{1}{3}\delta_{ij}\nabla^2\right)E + \nabla_iF_j + \nabla_jF_i\right]\,, \end{aligned} \tag{5.48}$$

with ζ and E two scalar functions, F_i a divergence-free vector obeying $\nabla_i F^i = 0$ and h_{ij} a transverse-traceless (TT) symmetric tensor satisfying

$$\nabla_i h^{ij} = \partial_i h^{ij} = 0\,, \qquad\qquad h^i_i = 0\,. \tag{5.49}$$

5.2.2 Physical degrees of freedom

Before proceeding with the proper computation of cosmological perturbations, let us perform a quick counting of the physical degrees of freedom (d.o.f.). To highlight the procedure, we will work in arbitrary spacetime dimensions.

In $D+1$ dimensions, the symmetric metric tensor $g_{\mu\nu}$ has $(D+1)\,(D+2)/2$ degrees of freedom, so we should expect to recover this number from the ADM decomposition. Indeed, in the ADM splitting we have:

Type		# metric d.o.f.
4 scalars	N, ψ, ζ and E	4
2 div-free (spatial) vectors	$\bar{N}_i$ and F_i	$2\,(D-1)$
1 TT symmetric (spatial) tensor	h_{ij}	$D\,(D+1)/2-(D+1)$

which makes a total of

$$\frac{1}{2}(D+1)\,(D+2) \tag{5.50}$$

degrees of freedom, as expected. Note however that not all these degrees of freedom are physical. In particular, we can always perform $D+1$ coordinate transformations to eliminate $D+1$ of them. These coordinate transformations can be written as

$$t \to t+\delta t \qquad\qquad x^i \to x^i + \delta x^i + \nabla^i \delta x, \tag{5.51}$$

with δx a scalar function and δx^i a divergence-free vector. Taking into account this *gauge freedom*, we are left with:

Type	# metric d.o.f	-	# coordinate d.o.f		
Scalars	N, ψ, ζ E	-	δt and δx	=	2 for $\forall D$
Div-free vectors	$\bar{N}_i$ and F_i	-	δx^i	=	$D-1$
TT tensors	h_{ij}	-	—	=	$D\,(D+1)/2-(D+1)$

In $3+1$ dimensions, we have 2 scalar, 2 vector and 2 tensor modes in the metric sector. These degrees of freedom must be complemented with an additional degree of freedom from the matter sector (the inflaton) and two constraints (5.43) for the lapse and shift functions. Up to vector degrees of freedom, this leaves us with 1 (physical) scalar and 2 (physical) tensor degrees of freedom.

5.2.3 Gauge fixing

As anticipated in the introduction of this chapter, not only the metric but also the scalar field changes under coordinate transformations. For instance, a temporal shift $t \to t+\delta t$

translates into a change on the metric perturbation ζ and on the field perturbation $\delta\phi$,

$$\widetilde{\zeta} = \zeta + H\delta t\,, \qquad \widetilde{\delta\phi} = \delta\phi - \dot{\phi}\delta t\,, \tag{5.52}$$

meaning that neither ζ nor $\delta\phi$ are invariant under general coordinate transformations. If not taken into account this *gauge-dependence* can give rise to strong contradictions. Imagine for instance a featureless Universe with no perturbations at all. If we perform a temporal shift $t \to t + \delta t$ we will end up in a Universe which seems to contain perturbations! ,

$$\widetilde{\zeta} = H\delta t\ , \qquad \widetilde{\delta\phi} = -\dot{\phi}\delta t\ . \tag{5.53}$$

Of course, this is just a *gauge artifact* indicating that the phase space we are working with is too big and must be restricted. One possibility is to work with gauge invariant objects such as

$$\zeta_{\rm GI} \equiv \zeta + H\frac{\delta\phi}{\dot{\phi}}\,. \tag{5.54}$$

Another possibility is to *fix the gauge*. One can find many gauge choices in the literature, each of them with its own pros and cons. We will work in the so-called *comoving gauge* in which $E = 0$ and the inflaton field ϕ has no perturbations (i.e. the observers move together with the cosmic fluid without measuring any flux)[6]

$$\delta\phi = 0\,, \qquad \gamma_{ij} = e^{2\zeta}a^2\left[\delta_{ij} + h_{ij}\right]\ . \tag{5.55}$$

The comoving gauge has some interesting properties:

1. Surfaces of constant ϕ coincide with surfaces of constant time.
2. The only fluctuating degrees of freedom are in the metric. The inflaton degree of freedom has been "eaten" by the metric, which acquires a longitudinal polarization ζ. This is analogous to what happens in spontaneously broken gauge theories. For this reason, we will sometimes refer to the comoving gauge as the *unitary gauge*.
3. The scalar perturbation ζ is directly associated to ${}^{(3)}R$,

$$ {}^{(3)}R = -2a^{-2}e^{-2\zeta}\left[2\partial^2\zeta + (\partial\zeta)^2\right]\ . \tag{5.56}$$

   For this reason, we will often refer to ζ as the *(comoving) curvature perturbation*.
4. (For adiabatic matter fluctuations) the scalar perturbation ζ is conserved on super-horizon scales

$$\lim_{k\ll aH}\dot{\zeta}_k = 0\,. \tag{5.57}$$

   The constancy of ζ outside the horizon allows to relate the primordial perturbations to CMB observations, while ignoring any postinflationary physics.

[6] The remaining (non-dynamical) scalar degrees of freedom N and ψ will be expressed in terms of ζ via the constraint equations (5.43). See Section (5.3).

5.3 Second-order action for perturbations

Solving the constraint equations (5.43) order by order in perturbation theory we can obtain algebraic solutions for N and N^i. The substitution of these expressions into the ADM action (5.41) will leave us with an action for the dynamical variables h_{ij} and ζ. On general grounds, these degrees of freedom will couple to each other. However, it is well known that scalar and tensor perturbations do not mix at first order in perturbation theory. Since we are interested only on the leading perturbations, we will make use of this *decomposition theorem*[7] and study scalar and tensor perturbations separately. In practice, this reduces to setting to zero one type of perturbation while studying the other but keeping always in mind that this has nothing to do with fixing the gauge!

5.3.1 Scalar perturbations

Setting all perturbations to zero except for the scalar mode and using the comoving gauge (5.55) we obtain

$$\gamma_{ij} = a^2 e^{2\zeta}\delta_{ij}\,, \qquad \Longrightarrow \qquad \dot\gamma_{ij} = 2a^2(H+\dot\zeta)e^{2\zeta}\,, \tag{5.58}$$
$$\gamma^{ij} = a^{-2} e^{-2\zeta}\delta^{ij}\,, \qquad \Longrightarrow \qquad \dot\gamma^{ij} = -2a^{-2}(H+\dot\zeta)e^{-2\zeta}\,. \tag{5.59}$$

Be careful

Note the change of sign in $\dot\gamma^{ij}$ with respect to $\dot\gamma_{ij}$: $\gamma^{ik}\gamma_{kj} = \delta^i{}_k \Longrightarrow \dot\gamma^{ik}\gamma_{kj} = -\gamma^{ik}\dot\gamma_{kj}$.

Using Eqs. (5.58) and (5.59) we can compute the Levi-Civitá connection

$$\Gamma^k_{ij} = \frac{1}{2}\gamma^{kl}\left(\partial_j\gamma_{ik} + \partial_i\gamma_{kj} - \partial_k\gamma_{ij}\right) = \delta^{kl}\left(\partial_j\zeta\delta_{ik} + \delta_i\zeta\delta_{jk} - \partial_k\zeta\delta_{ij}\right)\,, \tag{5.60}$$

the extrinsic curvature invariants

$$K_{ij} = \frac{1}{N}\Big[a^2e^{2\zeta}\delta_{ij} - \partial_{(i}N_{j)} + \big(2N_{(i}\partial_{j)}\zeta - N_k\partial_k\zeta\delta_{ij}\big)\Big]\,, \tag{5.61}$$
$$K^{ij} = \gamma^{ik}\gamma^{jl}K_{kl} = a^{-4}e^{-4\zeta}\delta^{ik}\delta^{jl}K_{kl}\,, \tag{5.62}$$
$$K = \gamma^{ij}K_{ij} = \frac{1}{N}\Big[3(H+\dot\zeta) - a^{-2}e^{-2\zeta}(\partial_kN_k + N_k\partial_k\zeta)\Big]\,, \tag{5.63}$$

and the action terms

$${}^{(3)}R = -2a^{-2}e^{-2\zeta}\left[2\partial^2\zeta + (\partial\zeta)^2\right]\,, \tag{5.64}$$
$$N^2(K_{ij}K^{ij} - K^2) = -6(H+\dot\zeta)^2 + 4a^{-2}e^{-2\zeta}(H+\dot\zeta)(\partial_iN_i + N_i\partial_i\zeta) - a^{-4}e^{-4\zeta}\left[(\partial_iN_i)^2 + 2(\partial_iN_i\zeta)^2 - (\partial_{(i}N_{j)} - (\partial_iN_i + N_i\partial_i\zeta))^2\right]\,, \tag{5.65}$$

where the parenthesis around the indices denote symmetrization, namely $2\partial_{(i}N_{j)} = \partial_iN_j + \partial_jN_i$.

[7] The proof of this theorem is straightforward but rather tedious so we will not perform it here. The curious reader is referred to Weinberg's book on Cosmology.

Be careful

Note that the indices in some expressions are still summed over but not contracted!, i.e. $\delta^{ij}\partial_j V_i = \partial_i V_i$. In particular, the quantity $\partial^2 \equiv \delta^{ij}\partial_i\partial_j$ is not the same as $\partial^i\partial_i$. These two quantities differ by a factor of the space metric.

Exercise

Derive Eqs. (5.60)-(5.65).

To obtain the quadratic action for ζ we need to solve the constraints equations (5.43) at first order in perturbation theory. Expanding the lapse and shift functions around the unperturbed FLRW values $N^{(0)} = 1$ and $N_i^{(0)} = 0$,

$$N = \sum_{n=0}^{\infty} N^{(n)} = 1 + N^{(1)} + \ldots\,, \qquad N_i = \sum_{n=0}^{\infty} N_i^{(n)} = \sum_{n=1}^{\infty}\left(\partial_i\psi^{(n)} + \bar{N}_i^{(n)}\right)\,, \tag{5.66}$$

and taking into account Eqs. (5.61)-(5.65) we get

$$a^{-2}\partial^2\psi^{(1)} = -a^{-2}\frac{\partial^2\zeta}{H} + \frac{\dot{\phi}^2}{2H^2}\dot{\zeta}\,, \qquad \nabla_i(HN^{(1)} - \dot{\zeta}) = 0\,, \qquad \partial^2\bar{N}_i^{(1)} = 0\,. \tag{5.67}$$

The solution of these equations reads

$$\psi^{(1)} = -\frac{\zeta}{H} + a^2\frac{\dot{\phi}^2}{2H^2}\,\partial^{-2}\dot{\zeta}\,, \qquad N^{(1)} = \frac{\dot{\zeta}}{H}\,, \qquad \bar{N}_i^{(1)} = 0\,, \tag{5.68}$$

with $\partial^{-2}(\partial^2\dot{\zeta}) = \dot{\zeta}$.

Exercise

Show that the equations in (5.68) do indeed verify the differential equations in (5.67).

Substituting the first-order solutions for N and N_i back into the ADM action and expanding it to second order we get

$$S_s^{(2)} = \frac{1}{2}\int d^4x\, a^3\frac{\dot{\phi}^2}{H^2}\left[\dot{\zeta}^2 - a^{-2}(\partial_i\zeta)^2\right] = \kappa^{-2}\int d^4x\, a^3\epsilon\left[\dot{\zeta}^2 - a^{-2}(\partial_i\zeta)^2\right]\,, \tag{5.69}$$

where we have made use of the background equations of motion and performed "a lot of integration by parts" (J. Maldacena). Note that the final expression is suppressed by the Hubble flow parameter (4.7). This reflects the fact that ζ is a pure gauge mode in (exact) de Sitter.

Exercise

1. Derive Eq. (5.69).
2. Show that on super-Hubble scales ($k \ll aH$)

$$\zeta(t) = c_1 + c_2 \int^t dt_1 \ \exp\left\{-\int^{t_1} [3+\eta(t_2)]H(t_2)dt_2\right\} . \tag{5.70}$$

with η the Hubble flow parameter in Eq. (3.50). What happens at long times?

For the purposes of the next chapter, it is convenient to introduce the *Mukhanov-Sasaki variable*

$$v \equiv z\,\zeta\,, \qquad \text{with} \qquad z^2 \equiv a^2 \frac{\dot{\phi}^2}{H^2} = 2a^2\epsilon\,. \tag{5.71}$$

In terms of v the action for the curvature perturbation (5.69) becomes

$$S_s^{(2)} = \int d\tau\, d^3\mathbf{x}\, \mathcal{L} = \frac{1}{2}\int d\tau\, d^3x \left[v'^2 - (\partial_i v)^2 - m^2(\tau)v^2\right] , \tag{5.72}$$

with τ ithe conformal time, $' = \partial/\partial\tau$ and

$$m^2(\tau) = -\frac{z''}{z}\,. \tag{5.73}$$

Note that the action (5.72) coincides formally with that of a massive scalar field in Minkowski spacetime (with the coordinate time t replaced by the conformal time τ). All the interactions among the scalar field and gravity are effectively encoded in the time-dependent mass $m^2(\tau)$. f

5.3.2 Tensor perturbations

The computation of the second order action for tensor perturbations is considerably simpler. Expanding the Einstein-Hilbert action at second order we get

$$S_t^{(2)} = -\frac{1}{2}\int \mathrm{d}\tau \mathrm{d}^3x \frac{a^2}{4\kappa^2}\eta^{\mu\nu}\partial_\mu h_{ij}\partial_\nu h_{ij}\,. \tag{5.74}$$

Defining a new field variable

$$v_{ij} \equiv \frac{a}{2\kappa}h_{ij}\,, \tag{5.75}$$

we can recast the action (5.74) as

$$S_t^{(2)} = \int \mathrm{d}\tau \mathrm{d}x^3 \left[-\frac{1}{2}\eta^{\mu\nu}\partial_\mu v_{ij}\partial_\nu v_{ij} - \frac{1}{2}\bar{m}^2(\tau)v_{ij}^2\right] . \tag{5.76}$$

where we have defined

$$\bar{m}^2(\tau) = \frac{a''}{a}\,. \tag{5.77}$$

Eq. (5.74) should be recognized as essentially two copies of the curvature perturbation equation (5.72), one for each polarization mode of the gravitational waves. As for curvature perturbations, all the interactions are encoded in the time-dependent mass $\bar{m}^2(\tau)$.