Week 6: Perturbations to the FRW metric

March 2, 2017

1 Motivation

Note: These lectures taken largely from Ed Bertschinger's "Cosmological Dynamics," astro-ph/9503125.

Last quarter we introduced the FRW metric to describe an isotropic and homogeneous Universe, determined the equation of motion (the Friedmann equation) for its scale factor $a(t)$, and learned that our Universe is flat (or very close to it) and consists of 70% vacuum energy, 25% nonbaryonic nonrelativistic dark matter, 5% baryons, and roughly 10^{-4} of critical density in a $T = 2.7$ K photon blackbody, and a $T = 1.96$ K neutrino background.

We also saw that the temperature of the CMB is the same to 1 part in 10^5 in every direction on the sky, and we also see fluctuations in the temperature at this level. And since these photons last scattered at a redshift $z = 1100$, we know that the early Universe was very smooth. However, when we look at the Universe today, it is highly clumpy: there are galaxies, clusters of galaxies, and deviations from homogeneity on even larger scales (although the amplitude of these inhomogeneities becomes small at large scales, a statement that the Universe is homogeneous on the largest scales). Heuristically, it is understandable that gravity amplified the tiny primordial perturbations implied by the $\Delta T/T \sim 10^{-5}$ CMB fluctuations into the large-scale structure we see today: Overdense regions will accrete matter from the underdense regions. In this way, the overdensities become increasingly overdense and the underdense regions increasingly underdense. Our purpose this week will be to describe these inhomogeneities quantitatively and to derive equations of motion, from general relativity, for the growth of these perturbations.

2 The perturbed metric

We begin by writing the FRW metric in terms of the conformal time τ, defined by $d\tau = dt/a(t)$, as

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu = a^2(\tau)\left[d\tau^2 - \gamma_{ij}\, dx^i\, dx^j\right],$$

where Latin indices indicate spatial components and Greek indices run over all four spacetime coordinates. We will restrict our attention to a flat Universe, in which case $\gamma_{ij} = \delta_{ij}$. The generalization to a nonflat Universe is straightforward, vastly more complicated, and conceptually not very illuminating.

We now consider small perturbations to this metric. The most general perturbed FRW metric can be written,

$$ds^2 = a^2(\tau)\left\{-(1+2\psi)d\tau^2 + 2w_i d\tau dx^i + [(1-2\phi)\gamma_{ij} + 2h_{ij}]dx^i dx^j\right\},$$

with h_{ij} traceless: $\gamma^{ij}h_{ij} = 0$. Here, $\psi(\vec{x},\tau)$ and $\phi(\vec{x},\tau)$ are scalars, $w_i(\vec{x},\tau)$ is a vector, and $h_{ij}(\vec{x},\tau)$ is a symmetric trace-free tensor. We may choose h_{ij} to be traceless, as any trace can be absorbed into ϕ. The quantities ψ, ϕ, w_i, and h_{ij} are all functions of spacetime, and they are all assumed to be $\ll 1$. Throughout, we will work only to linear order in these perturbation variables. Therefore, we may consistently treat the perturbation variables as 3-tensors with components raised and lowered with the metric γ_{ij}. Note, however, that four-vector indices are still raised and lowered with $g_{\mu\nu}$. We know that this is the most general linear perturbation, as these functions represent $10 = 1+1+3+5$ independent components corresponding to the 10 components of the most general metric $g_{\mu\nu}$. Since we are allowed to choose our four spacetime coordinates $(\tau, \vec{x})$ in any way we want without changing any physical quantities, only six of these fields represent physical degrees of freedom. Since $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ must be invariant under a general coordinate transformation ("gauge transformations"), the fields ϕ, ψ, w_i, and h_{ij} will also change under an infinitesimal coordinate transformation.

Let's first consider the vector $w_i(\vec{x},t)$. This vector can be decomposed into a curl and curl-free part: $\vec{w} = \vec{w}_{\parallel} + \vec{w}_{\perp}$, with $\vec{\nabla}\times\vec{w}_{\parallel} = \vec{\nabla}\cdot\vec{w}_{\perp} = 0$, or in components, $w_i = w_{\parallel,i} + w_{\perp,i}$ with $\epsilon^{ijk}\partial_j w_{\parallel,k} = 0$ and $\partial_i w_{\perp,i} = 0$. This last relation implies that $w_{\parallel,i}$ can be written $w_i = \partial_i w$, for some scalar function w, and so the vector w_i can actually be decomposed into a scalar function $w(\vec{x},t)$ and a vector perturbation $\vec{w}_{\perp}$ that is the transverse component of w_i.

Similarly, the symmetric trace-free tensor field $h_{ij}(\vec{x},t)$ can be decomposed into a longitudinal, solenoidal, and transverse part (or scalar, transverse-vector, and tensor parts): $h_{ij} = h_{\parallel,ij} + h_{\perp,ij} + h_{T,ij}$. The first and second components can be written in terms of a scalar function $h(\vec{x},t)$ and a transverse vector $h_i(\vec{x},t)$, meaning $\partial_i h_i = 0$ (you can distinguish between the scalar, vector, and tensor, by the number of indicies) as follows:

$$h_{\parallel,ij} = D_{ij}h, \qquad h_{\perp,ij} = \nabla_{(i}h_{j)}, \qquad \nabla_i h^i_{T,j} = 0,$$

where the parentheses denote symmetrization, and D_{ij} is a symmetric trace-free second derivative:

$$\nabla_{(i}h_{j)} \equiv \frac{1}{2}(\nabla_i h_j + \nabla_j h_i), \qquad D_{ij} \equiv \nabla_i\nabla_j - \frac{1}{3}\gamma_{ij}\nabla^2.$$

The divergences of $h_{\parallel,ij}$ and $h_{\perp,ij}$ are a longitudinal vector (one that can be written as the gradient of some scalar function) and a transverse vector (an honest-to-goodness vector),

$$\partial_i h_{\parallel,ij} = \frac{2}{3}\nabla_j\nabla^2 h, \qquad \partial_i \cdot h_{\parallel,ij} = \frac{1}{2}\nabla^2 h_j.$$

We have decomposed the tensor h_{ij} into parts that can be obtained from a scalar and a vector, and a part that is pure tensor, but this decomposition is not necessarily unique. First of all, the scalar $h(\vec{x},t)$ and vector $h_i(\vec{x},t)$ are only defined up to a constant. Secondly, the vector h_i can be changed by solutions to Killing's equation, $\partial_i h_j + \partial_j h_i = 0$. Vectors that satisfy this equation are Killing vectors, associated with the symmetries (translations and rotations) of the space (which here is restricted to flat 3-d space). For example, h_i can be changed by the vector field $(h_x, h_y, h_z) = (-y, x, 0)$ (a rotation around the z axis) without changing h_{ij}, although we probably wouldn't want to use this h_i if we impose the more restrictive constraint (although its not necessary in full generality) $h_i \ll 1$, in addition to $h_{ij} \ll 1$.

The tensor part $h_{T,ij}$ is also not unique. The divergence-free condition $\partial_i h_{T,ij} = 0$ and the trace-free condition are still satisfied even if we add to $h_{T,ij}$ a tensor $\zeta_{ij} \equiv D_{ij}\zeta$, where $\zeta(\vec{x},t)$ is a scalar that satisfies $\nabla^2\zeta = 0$. However, this shift is just equivalent to changing the scalar component $h_{\|,ij}$.

We therefore have decomposed the most general metric perturbation into four scalar components (ϕ, ψ, w, and h), each having one degree of freedom, two transverse-vector components ($w_{\perp,i}$, h_i, with $\partial_i w_{\perp,i} = \partial_i h_i = 0$), each with two degrees of freedom, and a symmetric trace-free divergence-free tensor part ($h_{T,ij}$) having two degrees of freedom (five components for a symmetric trace-free tensor minus the thre divergence-free conditions), totalling to 10 degrees of freedom.

3 The stress-energy tensor

According to the Einstein equations, perturbations to the stress-energy tensor will induce perturbations to the FRW metric. We therefore need to understand how we quantify perturbations to the stress-energy tensor.

For a perfect fluid, the stress-energy tensor is $T^{\mu\nu} = (\rho+p)u^\mu u^\nu + pg^{\mu\nu}$, where ρ and p are the proper energy density and pressure in the fluid rest frame, and $u^\mu = dx^\mu/ds$ is the fluid four-velocity. In locally flat coordinates in the perfect-fluid frame, $T^{00} = \rho$ is the energy density, the momentum density $T^{0i} = 0$, and the spatial stress tensor $T_{ij} = p\delta^{ij}$.

For an imperfect fluid, the stress-energy tensor may have additional components that describe shear and bulk viscosity or thermal conduction. The most general stress tensor is thus

$$T^{\mu\nu} = (\rho + p)u^\mu u^\nu + pg^{\mu\nu} + \Sigma^{\mu\nu},$$

where $\Sigma^{\mu\nu}$ can be taken to be traceless Σ^μ_μ and flow orthogonal $\Sigma^\mu_\nu u^\nu = 0$. In locally flat coordinates in the fluid rest frame, only the spatial components Σ^{ij} are nonzero (and $\Sigma^i_i = 0$). With $\Sigma^{\mu\nu}$ so defined, the fluid velocity u^μ is defined so that ρu^μ is the energy-current four-vector, and *not* the mass times particle-number four-vector. Then ρu^μ includes heat conduction, and p includes bulk viscosity, and $\Sigma^{\mu\nu}$ (the *shear stress*) includes shear viscosity.

The components of the stress-energy tensor will need to be evaluated in the comoving coordinate frame of our perturbed metric, and to begin we will need the components of the four velocity in the perturbed spacetime. The comoving frame is defined by that where the fluid is at rest: i.e., $u^i = 0$. The normalization $u^\mu u_\mu = -1$ then requires $u^0 = a^{-1}(1-\psi)$ to lowest order in ψ, and lowering the index with the full metric (again to linear order in perturbation variables) gives $u_0 = -a(1+\psi)$ and $u_i = aw_i$.

The appearance of ψ and w_i in u_0 arises because the proper-time interval $a(\tau)(1+\psi)d\tau$ now depends on position; as we will see, the ψ perturbation reduces in the Newtonian limit to the gravitational potential, and two observers in two different potential wells cannot necessarily synchronize their watches; it is the gravitational redshift associated with the potential well. If $w_i \neq 0$, then an observer at x^i sees a clock at $x^i + dx^i$ run faster by an amount $w_i dx^i$; this is a frame-dragging effect.

So far we have written four-velocity components for a comoving fluid element, defined by $u^i = 0$. Now consider a fluid element that moves with a "peculiar velocity," a coordinate 3-velocity (a Cartesian three-vector), $v^i \equiv dx^i/d\tau = dx^i/dx^0$; i.e., it moves a coordinate distance dx^i during the elapse of a time $d\tau$. If spacetime perturbations are small, then the induced peculiar velocities will be small and so $v^i \ll 1$ is linear in the perturbation variables. Strictly speaking, v^i is not the proper 3-velocity, as with the presence of perturbations, adx^i is not a proper distance and $ad\tau$ is not a proper time. However, to lowest order in the perturbations, v is the proper 3-velocity. To linear order in perturbation variables, the four-velocity components of a fluid element with a 3-velocity v^i are

$$u^0 = a^{-1}(1-\psi), \quad u^i = a^{-1}v^i, \quad u_0 = -a(1+\psi), u_i = 1(v_i + w_i).$$

If $w_i \neq 0$, then the worldline of an observer that comoves in the coordinate system defined by our perturbed metric (i.e., $v_i = 0$) is not normal to the hypersurfaces τ =constant. In other words, $u_\mu \xi^\mu = a w_i \xi^i \neq 0$ for a four-vector that lives in the three-space ($\xi^i \neq 0$ and $\xi^0 = 0$). What this means is that if $w^i \neq 0$, we are not in a locally inertial frame, in which the worldline of a freely-falling observer is normal to the spatial directions. Thus, w_i is interpreted as a frame-dragging effect. By making a local transformation $dx^i \to dx^i + w^i d\tau$, we can eliminate w_i at any given point. The transformation corresponds to choosing a locally inertial frame, the *normal frame*, that moves with 3-velocity $-w_i$ relative to the comoving frame. In the normal frame, the fluid 3-velocity is more generally $v^i + w^i$.

Since there are four coordinates we can choose arbitrarily and 10 metric components, we can always find a coordinate system in which $w_i = 0$. However, if $w_i \neq 0$ and it varies spatially, then this corresponds to shearing and/or rotation of the comoving frame relative to the normal frame...this is dragging of inertial frames. In general, the comoving frame is noninertial: nongravitational forces must be applied to keep a particle at fixed x^i.

With the components of u^μ, we can write the stress-energy tensor. We do so with mixed indices to reduce the presence of metric-perturbation variables. The components are:

$$T^0_0 = -\rho, \qquad T^i_0 = -(\rho + p)v^i,$$

$$T^0_i = (\rho + p)(v_i + w_i), \qquad T^i_j = p\delta^i_j + \Sigma^i_j.$$

Note that the traceless shear stress Σ_{ij} can be decomposed into scalar, vector, and tensor parts, as for h_{ij}, and the energy-flux density $(\rho + p)v_i$ can similarly be decomposed into a scalar and transverse-vector part. The pressure appears in the flux density to account for the pdV work done when the fluid expands. The only approximations that we have made here are that v, and Σ_{ij} are of the same order as the metric-perturbation variables, and we neglect terms of quadratic order in these perturbations.

To linear order, energy-momentum conservation $\nabla_\mu T^\mu_\nu$ becomes

$$\partial_\tau \rho + 3(aH - \dot{\phi})(\rho + p) + \partial_i[(\rho + p)v_i] = 0,$$

and

$$\partial_\tau[(\rho + p)(v_i + w_i)] + 4aH(\rho + p)(v_i + w_i) + \partial_i p + \partial_j \Sigma^i_j + (\rho + p)\partial_i \psi = 0.$$

The first equation generalizes the usual continuity equation. This is clear if we take $p \ll \rho$ and $H = 0$. The presence of p takes into account the modification in the flux density when pdV work is included. The aH term describes the dilution due to expansion, and the $\dot{\phi}$ in that term is included there since the local scale factor is changed from a to $a(1 - \phi)$ in the perturbed spacetime. The second equation is the momentum-conservation (Euler) equation, where the $\partial_i p$, $\partial_j \Sigma^i_j$, and $(\rho + p)\partial_i \psi$ terms are the nongravitational and gravitational acceleration.

4 Gauges and gauge transformations

4.1 Synchronous gauge

Recall that there are ten spacetime-perturbation variables, and four can be eliminated if we choose, with the right choice of the four spacetime coordinates. The *synchronous gauge* corresponds to a choice of coordinates (actually, as we will see, a class of coordinates) in which $w_i = \psi = 0$ everywhere. In these coordinates, there is a set of comoving observers who freely fall without changing their spatial coordinates x_i. These are called "fundamental" comoving observers. These follow from the geodesic equation,

$$\frac{du^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta} u^\alpha u^\beta = 0,$$

for the trajectory $x^\mu(\lambda)$, where $d\lambda = (-ds^2)^{1/2}$ for a timelike geodesic, and $u^\mu = dx^\mu/d\lambda$. Then $\Gamma^i_{00} = 0$, implying that $u^i = 0$ is a geodesic.

In synchronous gauge, the proper time t (or conformal time $\tau = \int dt/a$) measured by a clock carried by each comoving observer, and their fixed spatial coordinates x^i *define* the coordinate system. There is then a residual gauge freedom that comes from the freedom to adjust the initial settings of the clocks and the initial coordinate labels of the observers.

Since each comoving observer carries a fixed x^i, these coordinates are *Lagrangian* coordinates. The coordinate system will break down in the nonlinear regime, when $\delta\rho \sim \bar{\rho}$, during "shell crossing", when the trajectories of different fundamental observers intersect. However, in the linear regime, $\delta\rho \ll \bar{\rho}$, there is no problem.

For consistency with the literature, we absorb ϕ into h_{ij}, and consider an h_{ij} that is no longer traceless: $h \equiv h^i_i \neq 0$. Then, the metric can be written economically as

$$ds^2 = a^2(\tau)[-d\tau^2 + (\gamma_{ij} + h_{ij})dx^i dx^j].$$

Some straightforward but horrendous tensor manipulation then yields the components of the Einstein tensor:

$$-a^2 G^0_0 = 3a^2H^2 + aH\dot{h} - \frac{1}{2}\nabla^2 h + \frac{1}{2}\partial_i\partial_j h^{ij},$$

$$a^2 G^0_i = \frac{1}{2}(\partial_i \dot{h} - \partial_j \dot{h}^j_i), \quad G^i_0 = -\gamma^{ij} G^0_j,$$

$$\begin{aligned}
-a^2 G^i_j &= 2(a\dot{H}) + (aH)^2 \delta^i_j + \left(\frac{1}{2}\partial^2_\tau + aH\partial_\tau - \frac{1}{2}\nabla^2\right)(h\delta^i_j - h^i_j) && (1)\\
&\quad + \frac{1}{2}\gamma^{ik}(\partial_k\partial_j h - \partial_k\partial_l h^l_j - \partial_j\partial_l h^l_k) && (2)\\
&\quad + \frac{1}{2}(\partial_k\partial_l h^{kl})\delta^i_j, && (3)\\
& && (4)
\end{aligned}$$

where, to be clear, $\nabla^2 = \partial^i\partial_i$. Note that the unperturbed parts are what they should be.

The next task will be to separate the perturbed Einstein equations into separate equations for the scalar, vector, and tensor parts. We first write (as before, but now with the trace term)

$$h_{ij} = \frac{1}{3}h\gamma_{ij} + D_{ij}(\nabla^{-2}\xi) + \partial_{(i}h_{j)} + h_{T,ij},$$

requiring $\partial_i h^i = \partial_i h^i_{T\,j} = 0$. Note that there are two vector degrees of freedom, two tensor, and two scalars, for a total of six physical degrees of freedom. As you will show in the homework, the Einstein equations separate into 7 different parts:

$$G^0_0: \qquad \frac{1}{3}\nabla^2(\xi - h) + aH\dot{h} = 8\pi G a^2(\rho - \bar{\rho}),$$

$$G^0_{i,\parallel}: \qquad \frac{1}{3}\partial_i(\dot{h} - \dot{\xi}) - \partial_i(\nabla^{-2}\dot{\xi}) = 8\pi G a^2[(\rho + p)v_i]_parallel,$$

$$G^0_{i,\perp}: \qquad -\frac{1}{4}\nabla^2\dot{h}_i = 8\pi G a^2[(\rho + p)v_i]_\perp,$$

$$G^i_i: \qquad -(\partial^2_\tau + 2aH\partial_\tau)h + \frac{1}{3}\nabla^2(h - \xi) = 24\pi G a^2(p - \bar{p}),$$

$$G^i_{j\neq i,\parallel}: \qquad \left(\frac{1}{2}\partial^2_\tau + aH\partial_\tau\right)D_{ij}(\nabla^{-2}\xi) + \frac{1}{6}D_{ij}(\xi - h) = 8\pi G a^2 \Sigma_{ij,\parallel},$$

$$G^i_{j,\perp}: \qquad \left(\frac{1}{2}\partial^2_\tau + aH\partial_\tau\right)\nabla_{(i}h_{j)} = 8\pi G a^2 \Sigma_{ij,\perp},$$

$$G^i_{j,T}: \qquad \left(\frac{1}{2}\partial^2_\tau + aH\partial_\tau - \frac{1}{2}\nabla^2\right)h_{ij,T} = 8\pi G a^2 \Sigma_{ij,T}.$$

Note that we have too many equations: four scalar equations for h and ξ, two vector equations for h_i, and one tensor equation for $h_{T,ij}$. The G^0_μ equations, the first three equations, involve only one time derivative, while the rest (the G^i_μ equations) involve two time derivatives. A closer look suggests that we can discard either the first three equations, or the second three equations; they are redundant. This follows from the constraint $\nabla_\mu T^\mu_\nu = 0$. Something similar happens in the unperturbed case with the two forms of the Friedmann equations, which both lead to the same dynamics. It can be checked that the two forms of the perturbed equations also lead to the same dynamics.

The final equation is the wave equation for propagation of gravitational waves (symmetric, traceless, and transverse or divergence-free tensor metric perturbations). The aH term simply arises from the expression for the Laplacian in an expanding Universe. The right-hand side is the transverse-traceless stress, which acts as the source for gravitational waves.

4.2 Gauge modes

In the above, we chose a coordinate system in which there were only six perturbation degrees of freedom, corresponding to the six physical modes. However, in the most general perturbed FRW spacetime, there will be 10 perturbation variables, four of which correspond to non-physical, or "gauge" modes. A *gauge transformation* in this context is simply a change of coordinates:

$$\hat{\tau} = \tau + \alpha(\vec{x}, \tau), \qquad \hat{x}^i = x^i + \gamma^{ij}\partial_j\beta(\vec{x}, \tau) + \epsilon^i(\vec{x}, \tau),$$

with $\partial_i \epsilon^i = 0$, and where we have split the transformation into a scalar (for the time) and longitudinal and transverse parts for the spatial transformation.

In an unperturbed FRW metric, there is a natural notion of simultaneity: namely, equidensity hypersurfaces, which are always orthogonal to the worldlines of freely-falling observers. In a perturbed Universe, this is not as clear. In particular, our coordinate choice may seem to indicate perturbations, even when there aren't any. For example, consider an unperturbed FRW Universe, in which the density ρ depends only on τ, and now transform the coordinates with a nonzero α. Then in the transformed $(\hat{x}^i)$ system, $\rho(\hat{\tau}, \vec{x}) = \bar{\rho}(\tau) + (\partial_\tau \bar{\rho})\alpha(\vec{x}, \tau)$. This example shows that we must think a bit harder about what is a physical perturbation and what is a lousy coordinate choice.

When we transform the coordinates, we must also transform the perturbation variables so that ds^2 remains invariant. This results in

$$\hat{\psi} = \psi - \dot{\alpha} - aH\alpha, \qquad \hat{\phi} = \phi + \frac{1}{3}\nabla^2\beta + aH\alpha,$$

$$\hat{w}_i = w_i + \partial_i(\alpha - \dot{\beta}) - \dot{\epsilon}_i, \qquad \hat{h}_{ij} = h_{ij} - D_{ij}\beta - \nabla_{(i}\epsilon_{j)}.$$

The transformed fields are to be evaluated at the same coordinate values as the original fields.

Consider now the synchronous gauge, with $\psi = w_i = 0$ (and defining the trace $h = -6\phi$). You then see that ψ and w_i transform into themselves, while h_{ij} and ϕ transform into themselves. There is thus a whole family of synchronous gauges that are related to each other by

$$\hat{h} = h - 2\nabla^2\beta - 6aH\dot{\beta}, \quad \hat{\xi} = \xi - 2\nabla^2\beta, \quad \hat{h}_i = h_i - 2\epsilon_i,$$

where the variables α and β have been restricted by $\alpha = \dot{\beta}$, and

$$\beta = \beta_0(\vec{x}) \int \frac{d\tau}{a(\tau)}, \qquad \epsilon_i = \epsilon_i(\vec{x}).$$

In other words, there is a family of synchronous gauges related to each other by a scalar function β_0 and a transverse-vector function of the spatial coordinates. In 1980, Bardeen defined scalar perturbations Φ_A and Φ_H (to replace h and ξ),

$$\Phi_A = -\frac{1}{2}\nabla^{-2}(\ddot{\xi} + aH\dot{\xi}), \quad \Phi_H = \frac{1}{6}(h - \xi) - \frac{1}{2}aH\nabla^{-2}\dot{\xi}.$$

that are invariant under synchronous gauge transformations.

4.3 Poisson gauge

There are plenty of other gauge choices that one could make, and some are preferable to others either for numerical reasons, or because they look more like Newtonian equations. Here we will consider the Poisson gauge, by imposing the following gauge conditions:

$$\partial_i w^i = 0, \qquad \partial_i h^i_j = 0,$$

which counts as four constraint equations. In this gauge, there are two scalar potentials ψ and ϕ, one transverse-vector potential w_i, and one transverse-traceless tensor potenial h_{ij}, accounting for 6 degrees of freedom. In the literature, a more restrictive version, known as *longitudinal* or *conformal Newtonian* gauge, imposes $w_i = h_{ij} = 0$, but this can only be used if there are no vector or tensor perturbations. This is therefore not really a gauge choice, but a restriction of the physical degrees of freedom.

The most general perturbed FRW metric can be brought into Poisson gauge if

$$\alpha = w + \dot{h}, \quad \beta = h, \quad , \epsilon_i = h_i,$$

where w is defined from $w_{\|,i=-\nabla w}$, while h and h_i are the fields that describe the longitudinal and solenoidal parts of h_{ij}. Since these transformations involve no derivatives, the transformation is almost unique, except for the addition of arbitrary functions of time alone to α (representing changes in the units of time and length) and ϵ_i (representing a shift in the origin).

In terms of the synchronous-gauge variables, the Poisson-gauge variables are

$$\psi = -\frac{1}{2}\nabla^{-2}(\ddot{\xi} + aH\dot{\xi}), \quad \phi = \frac{1}{6}(\xi - h) + \frac{1}{2}aH\nabla^{-2}\dot{\xi}, \quad w_i = -\frac{1}{2}\partial_\tau h_i,$$

so $\psi = \Phi_A$ and $\phi = -\Phi_H$. The Poisson-gauge vector part is related to the synchronous-gauge solenoidal part.

In this gauge, the Einstein equations are

$$G^0_0: \qquad \nabla^2\phi - 3aH(\dot{\phi} + aH\psi) = 4\pi Ga^2(\rho - \bar{\rho}),$$

$$G^0_{\|,i}: \qquad -\partial_i(\dot{\phi} + aH\psi) = 4\pi Ga^2[(\rho + p)(v_i + w_i)]_{\|},$$

$$G^0_{\perp,i}: \qquad \nabla^2 w_i = 16\pi Ga^2[(\rho + p)(v_i + w_i)]_{\perp},$$

$$G^i_i: \qquad \ddot{\phi} + aH(\dot{\psi} + 2\dot{\phi}) + [2\dot{(aH)} + (aH)^2]\psi - \frac{1}{3}\nabla^2(\phi - \psi) = 4\pi Ga^2(p - \bar{p}),$$

$$G^i_{\|,j\neq i}: \qquad D_{ij}(\phi - \psi) = 8\pi Ga^2\Sigma_{\|,ij},$$

$$G^i_{\perp,j}: \qquad -(\partial_\tau + 2aH)\partial_{(i}w_{j)} = 8\pi Ga^2\Sigma_{\perp,ij},$$

$$G^i_{T,j}: \qquad (\partial^2_\tau + 2aH\partial_\tau - \nabla^2)h_{ij} = 8\pi Ga^2\Sigma_{T,ij}.$$

In this gauge, the perturbation variable ϕ, ψ, and w_i are determined by the instantaneous matter distribution, with no time evolution required. The third equation determines w_i and the fifth determines a combination of ϕ and ψ. Then by combining the first two, we get an equation for ϕ:

$$\nabla^2\phi = 4\pi Ga^2[\delta\rho + 3aH\Phi_f], \quad -\partial_i\Phi_f \equiv [(\rho + p)(v_i + w_i)]_{\|}.$$

Bardeen has defined a matter perturbation variable $\epsilon_m \equiv (\delta\rho + 3aH\Phi_f)/\bar{\rho}$, which is a measure of the energy-density fluctuation in the normal (intertial) frame at rest with matter such that $v_i + w_i = 0$.

Note that with no expansion ($H = 0$), this reduces to the Newtonian Poisson equation. For $H \neq 0$, Φ_f (a longitudinal momentum density) is also a source for ϕ, but for a perturbation of characteristic size (length) $\lambda \ll H^{-1}$, $aH\Phi_f$ is smaller than δ_ρ by a factor λH. We next note that the shear stress can never exceed $\mathcal{O}(\rho c_s^2)$, where c_s is the sound speed, implying that $\phi \simeq \psi$ unless there is relativistic viscous matter. Finally, the third equation implies that $w_i \sim (v_H/c)^2 v_i$, where $v_H = \lambda H$ is the Hubble velocity across a distance λ. Therefore, the scalar potential ϕ generalizes the Newtonian gravitational potential in the Poisson gauge. The Poisson-gauge potential ϕ also has the advantage that it almost always remains small, even when the density perturbations become large, $\delta\rho \sim \rho$. Note also that the tensor-mode equation is identical to that for synchronous gauge. This is simply because the gauge freedom, which consists of scalar and vector degrees of freedom, does not allow for any gauge freedom in the tensor modes.

4.4 Newtonian gauge and conservation of super-horizon curvature

As mentioned above, the restriction of the Poisson gauge to scalar perturbations is the Newtonian (or conformal-Newtonian or longitudinal) gauge, in which the metric is

$$ds^2 = a^2(\tau)\left[-(1+2\psi)d\tau^2 + (1-2\phi)\gamma_{ij}dx^i dx^2\right]. \tag{5}$$

(Note that Weinberg and others switch *phi* and ψ relative to those used here.) After Fourier transforming the metric and density perturbations, the Einstein equations for a Fourier components for wavevector $\vec{k}$ become (suppressing, for notational economy, the $\vec{k}$ subscripts on the perturbation variables),

$$\begin{aligned}
k^2\phi + 3\frac{\dot{a}}{a}\left(\dot{\phi} + \frac{\dot{a}}{a}\psi\right) &= -4\pi G a^2 \delta\rho, &(6)\\
k^2\left(\dot{\phi} + \frac{\dot{a}}{a}\psi\right) &= 4\pi G a^2(\bar{\rho} + \bar{P})\theta, &(7)\\
\ddot{\phi} + \left(2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2}\right)\psi + \frac{k^2}{3}(\phi - \psi) &= -4\pi G a^2 \delta p, &(8)\\
k^2(\phi - \psi) &= 12\pi G a^2(\bar{\rho} + \bar{p})\Theta, &(9)
\end{aligned}$$

where $\bar{\rho}$ and $\bar{p}$ are the mean density and pressure respectively; $\theta = \nabla^i v_i$ is a velocity-perturbation variable; and $(\bar{\rho} + \bar{p}) \equiv -(\hat{k}_i\hat{k}_j - \delta_{ij}/3)\Sigma^i_j$ is the shear. Recall that the dot here denotes derivative with respect to conformal time.

We now show that there is a quantity, the curvature perturbation, that is under certain circumstances, conserved in the superhorizon limit; i.e., when $k/a \ll H$. This is important, as during inflation, the physical wavelength of a comoving mode becomes superhorizon during inflation and then re-enters the horizon during matter/radiation domination. As we will see, curvature perturbations are produced during inflation and then later provide the initial conditions for the growth of density perturbations during matter/radiation domination when the mode re-enters the horizon. The conserved quantity then allows us to relate the amplitude of primordial density perturbations inferred from the CMB and/or galaxy surveys, to the predictions of inflation.

The conserved quantity, $\mathcal{R}_{\vec{k}}$, can be shown to be the curvature perturbation, but for now, think of it simply as a constant that fixes the amplitude of $phi_{\vec{k}}$, $\psi_{\vec{k}}$, $\delta\rho_{\vec{k}}$, and the other perturbation variables inferred from these. Now consider the Newtonian-gauge Einstein equations above in the limit $k/a \ll H$. Next suppose that $\phi = \psi$—the conservation only arises in this case. Fortunately, this is what is expected during inflation and then on superhorizon scales through matter and radiation domination. You can then show that the Einstein equations are solved, in the limit $k/a \ll H$, by

$$\begin{aligned}
\phi(\tau) = \psi(\tau) &= \mathcal{R}\left[-1 + \frac{\dot{a}}{a^3}\int^{\tau} \left[a(\tau')\right]^2\, d\tau'\right], &(10)\\
\theta(\tau) &= k^2\frac{\mathcal{R}}{a}\int^{\tau} \left[a(\tau')\right]^2, &(11)\\
\delta s &= -\frac{\mathcal{R}\dot{\bar{s}}}{\left[a(\tau)\right]^2}\int^{\tau} \left[a(\tau')\right]^2. &(12)
\end{aligned}$$

(Note that the equations look a bit different than those in Weinberg, as we are working with conformal time rather than time.) Here, $s(\vec{x},\tau)$ is any scalar quantity—e.g., the density or the pressure. For example, the first Einstein equation becomes,

$$k^2\phi + 3\frac{\dot{a}}{a}\left(\dot{\phi} + 3\frac{\dot{a}}{a}\phi\right) = -4\pi G a^2 \delta\rho, \tag{13}$$

You then use $H^2 = 8\pi G\rho/3$ to verify that the solutions above satisfy this equation. The second Einstein equation is then satisfied by the solution for θ. The third Einstein equation looks more complicated, but is in fact degenerate (for $\phi = \psi$) with the first, in the same way that the two forms for the Friedmann equation are degenerate.

As the derivation in Weinberg's book shows, the result for δs is obtained in the following way: As $k \to 0$, we are considering perturbations on distance scales far greater than the Hubble distance. Causal processes cannot occur at such large separations, and so the densities, pressures, etc. in one region simply evolve as they would in a homogeneous Universe with a slightly different scale factor as that in another region of the Univerese; i.e., the densities/pressures in one region are related to that in another by a simple rescaling of τ and $\vec{x}$ (i.e., a change in $\phi = \psi$ can be absorbed by a re-scaling of τ and $\vec{x}$. For this reason, the fractional perturbations are said to be *adiabatic*; i.e., $(\delta\rho_\alpha/\bar{\rho}_\alpha) = \delta p_\alpha/\bar{p}_\alpha = -\theta_\alpha/k^2$ are the same for all constituents α (e.g., baryons, dark matter, neutrinos, photons, etc.), regardless of whether there is microscopic energy exchange between these species or not. Another heuristic way to describe adiabatic perturbations is to say that every comoving region of the Universe undergoes the same time evolution, but the starting time for each is slightly different. Adiabatic perturbations arise in single-field slow-roll inflation because during inflation, the energy density at any point is determined by the value of the scalar field; there is nothing else. There is thus only one "quantum number" that we can assign to each point in the Universe. The canonical inflationary scenario is, for this reason, also sometimes referred to as "single-clock" inflation.

Adiabatic perturbations are to be contrasted with models (models other than single-field slow-roll inflation) in which there may also be "isocurvature" or "entropy" perturbations, those in which the fractional perturbations to the energy densities and pressures of different constituents are different. Current measurements, as we will see, are consistent with primordial adiabatic perturbations. The amplitudes of primordial isocurvature perturbations are constrained to be no greater than $\sim 10\%$ of the adiabatic amplitude.