1 Intro to perturbation theory

We're now ready to start considering perturbations to the FRW universe. This is needed in order to understand many of the cosmological tests, including:

- CMB anisotropies.
- Galaxy clustering.
- Gravitational lensing.

In order to keep life simple, we will only do flat universes for now. We will consider each major constituent of the Universe and write down the variables that can be perturbed:

- *Metric tensor*: 10 components $\delta g_{\mu\nu}(x^k, \eta)$.
- *Dark matter*: treat pressure as exactly zero ("cold"), so is described by 4 numbers, $\delta\rho_{dm}(x^k, \eta)$ and $v^i_{dm}(x^k, \eta)$.
- *Baryonic matter*: for most of the course we will ignore the pressure, which is good on sufficiently large scales. (We'll see how large later on.) Then also described by 4 numbers, $\delta\rho_b(x^k, \eta)$ and $v^i_b(x^k, \eta)$.
- *Neutrinos*: Described by a phase space density as a function of phase space location: $\delta f(x^i, P_i, \eta)$.
- *Photons*: Also described by a phase space density as a function of phase space location: $\delta f(x^i, P_i, \eta)$. Later on we'll add polarization.

We will do linear perturbation theory, which means that we will throw out terms in the equations that are higher-order in perturbations, such as $\delta g_{01}\delta\rho_b$. We will also often make the transformation to Fourier-space variables, defined by e.g.:

$$\delta\rho_{dm}(k_i, \eta) = \int_{\mathbf{R}^3} \delta\rho_{dm}(x^i, \eta) e^{-ik_i x^i}\, d^3x^i. \tag{1}$$

The big advantage is that for a homogeneous background, each Fourier mode evolves independently. For this reason we will use the Cartesian coordinates for the spatial metric. (Spherical coordinates are possible but much messier.)

Often people put tildes on Fourier transformed quantities but we won't do that here to keep things simple – just look to see if there's a $\mathbf{k}$ or an $\mathbf{x}$ as the argument.

Warnings – VERY IMPORTANT:

- *Time variable* – during our study of perturbation theory, we will use η rather than t as our basic time variable. This is much more convenient from the point of view of the equations. In particular, the overdot $\dot{}$ will now represent $\partial/\partial\eta$, NOT $\partial/\partial t$.

- *Gauge dependence* – this is a GR-based programme, most variables are gauge-dependent and their numerical values have NO MEANING unless a choice of gauge is specified. This has been an endless source of confusion!

Relevant material from Dodelson – we're going to start approximately following the book (and use the same notation, although the presentation will be a bit different):

- Photon and neutrino equations: §§4.1–4.4.
- Dark matter: §4.5.
- Baryons: §4.6.
- Metric tensor: §§5.1–5.5.

In this set of lectures we'll treat the metric tensor as fixed, and the matter species will simply move as it dictates. Later we'll promote $g_{\mu\nu}$ to dynamical variables and complete the treatment of perturbation theory.

2 The collisionless Boltzmann equation for massless particles

This section will be devoted to deriving Equation (4.33) of Dodelson. The derivation will be more general and formal than what's in the book, because I don't want to have to re-derive everything later when we do gravitational waves.

Metric perturbations and observers. Let's write the general perturbation to the FRW metric:

$$ds^2 = a^2(\eta)\left\{-(1+2A)d\eta^2 - 2B_i d\eta\, dx^i + [(1+2D)\delta_{ij} + 2E_{ij}]dx^i\, dx^j\right\}, \quad (2)$$

where the perturbation variables A, B_i, D, and E_{ij} depend on x^k and η. (Dodelson considers the case where $A = \Psi$ and $D = \Phi$ are nonzero, but we're more general.)

We're going to follow the trajectory of a particle as it navigates through the Universe on a geodesic. In order to do so we'll have to specify its position x^k and momentum p_k as a function of time. The position is easy enough, but for the momentum it will be convenient to describe not "p_k" but rather the physical momentum components measured by some observer $\mathcal{O}$. In the homogeneous universe there was a unique preferred observer, the comoving observer, who carried a tetrad:

$$\begin{aligned} u^\mu = (\mathbf{e}_{\hat{0}})^\mu &= a^{-1}(1,0,0,0) \\ (\mathbf{e}_{\hat{1}})^\mu &= a^{-1}(0,1,0,0) \\ (\mathbf{e}_{\hat{2}})^\mu &= a^{-1}(0,0,1,0) \\ (\mathbf{e}_{\hat{3}})^\mu &= a^{-1}(0,0,0,1). \end{aligned} \quad (3)$$

(Note that there's an a^{-1} for u^μ now as well because we switched variables to η.) The comoving observer saw an isotropic CMB and hence was preferred.

In the perturbed universe, there is no longer any symmetry that picks out one observer. We are free to pick any tetrad as long as it is orthonormal:

$$u^\mu u_\mu = -1; \quad u^\mu (\mathbf{e}_{\hat{i}})_\mu = 0; \quad (\mathbf{e}_{\hat{i}})^\mu (\mathbf{e}_{\hat{j}})_\mu = 1. \tag{4}$$

Let's try to construct the most general such tetrad. Since there are 16 components of the tetrad $(\mathbf{e}_{\hat{\alpha}})^\mu$, and the orthonormality provides 10 constraints, we expect that there are 6 degrees of freedom in the choice of tetrad.

First construct u^μ by perturbing Eq. (3). In general it can be written as:

$$u^\mu = a^{-1}(1 + V^0, V^1, V^2, V^3), \tag{5}$$

where $V^0 ... V^3$ are perturbation variables. The orthonormality equation gives

$$u^\mu u_\mu = g_{\mu\nu} u^\mu u^\nu = -(1+2A)(1+V^0)^2 - 2B_i V^i + (1+2D) V^i V^i + 2E_{ij} V^i V^j. \tag{6}$$

Since A, B, D, E, V are perturbation variables, we only keep the first-order terms in them,

$$u^\mu u_\mu = -1 - 2A - 2V^0. \tag{7}$$

This has to be -1, so $V^0 = -A$ and

$$u^\mu = a^{-1}(1 - A, V^1, V^2, V^3). \tag{8}$$

Next we need the spatial metric components. We can write them by perturbing Eq. (3):

$$(\mathbf{e}_{\hat{i}})^0 = a^{-1}\xi_i, \quad (\mathbf{e}_{\hat{i}})^j = a^{-1}(\delta_i^j + S_i^j + \vartheta_i^j), \tag{9}$$

where ξ_i is a 3-vector, S_i^j is a symmetric 3×3 matrix, and ϑ_i^j is an antisymmetric 3×3 matrix. We can use the spacelike condition for these vectors:

$$\begin{aligned} u^\mu (\mathbf{e}_{\hat{i}})_\mu \quad = \quad & -(1+2A)(1-A)\xi_i - B_j(1-A)(\delta_i^j + S_i^j + \vartheta_i^j) - B_j V^j \xi_i \\ & +(1+2D)V^j(\delta_i^j + S_i^j + \vartheta_i^j) + 2E_{jk}(\delta_i^j + S_i^j + \vartheta_i^j)V^k. \end{aligned} \tag{10}$$

Working to first order in the perturbation variables:

$$u^\mu (\mathbf{e}_{\hat{i}})_\mu = -\xi_i - B_i + V^i. \tag{11}$$

(See the simplification?) This must be zero so

$$\xi_i = V^i - B_i. \tag{12}$$

The last orthonormality condition between spatial vectors gives (to first order)

$$(\mathbf{e}_{\hat{i}})^\mu (\mathbf{e}_{\hat{j}})_\mu = (1 + 2D)\delta_{ij} + 2E_{ij} + 2S_i^j. \tag{13}$$

Since S is symmetric, and this must evaluate to 1, we have

$$S_i^j = -D\delta_i^j - E_{ij}. \tag{14}$$

There is no constraint on ϑ_i^j. We thus have completely defined the observer's reference frame:

$$\begin{aligned} u^\mu &= a^{-1}(1-A, V^1, V^2, V^3) \\ (\mathbf{e}_{\hat{i}})^\mu &= a^{-1}(V^i - B_i, (1-D)\delta_{ij} - E_{ij} + \vartheta_{ij}), \end{aligned} \tag{15}$$

which depends on the 3-vector V^i and the antisymmetric tensor (3 components) ϑ_{ij}. So there are indeed 6 degrees of freedom in the tetrad. They represent:

- V^i: the velocity of the observer relative to the coordinate system (i.e. trajectory of constant x^1, x^2, x^3).
- ϑ_{ij}: orientation of observer's basis vectors – ϑ_{23}, ϑ_{31}, and ϑ_{12} correspond to infinitesimal rotations around 1, 2, and 3 axes.

All of these degrees of freedom are associated with the description of the Universe, they are not real propagating modes.

It's often useful to have the covariant components of the basis vectors:

$$\begin{aligned} u_\mu &= a(-(1+A), V^1 - B_1, V^2 - B_2, V^3 - B_3) \\ (\mathbf{e}_{\hat{i}})_\mu &= a(-V^i, (1+D)\delta_{ij} + E_{ij} + \vartheta_{ij}), \end{aligned} \tag{16}$$

Specific choices of tetrad. Of the many possible choices of tetrad, we want to choose the one that will be most convenient. It turns out that in linear perturbation theory, ϑ_{ij} will completely decouple from the problem. The reason is that in the unperturbed isotropic universe, all quantities (e.g. phase space densities) are invariant under 3-dimensional rotations, so ϑ_{ij} can appear only multiplying a perturbaton. So it doesn't matter how we choose ϑ_{ij}; for definitiveness we'll just set it to zero.

For the velocity V^i there are three interesting choices:

- *The coordinate observer*: $V^i = 0$. Conceptually the simplest choice; this corresponds to an observer who sits at constant spatial coordinates x^1, x^2, x^3.
- *The comoving observer*: $V^i = T^i_0/(\rho + p)$. This is the observer for whom the momentum density $T_{\mu\nu}u^\mu(\mathbf{e}_{\hat{i}})^\nu$ vanishes. (Prove on homework.)
- *The normal observer*: $V^i = B_i$. This is the observer who is moving orthogonal to the surface of constant η, i.e. who sees such a surface as their local surface of simultaneity. To this observer vectors with spacelike contravariant components $w^\mu = (0, w^1, w^2, w^3)$ are physically spacelike.

The normal observer has the most convenient properties; we (and Dodelson) will use it for all calculations.

Additional properties of the normal observer:

- Spatial covariant components of 4-velocity $u_i = 0$.
- Coincides with coordinate observer when $B_i = 0$, which will be true in most of our calculations (but not always; depends on gauge choice).
- Coincides with comoving observer when $T^0_i = 0$. (Homework!)
- Mapping from covariant momentum components P_i to observer-frame components $P_{\hat{i}}$ is linear, because $(\mathbf{e}_{\hat{i}})^0 = 0$:

$$P_{\hat{i}} = \frac{1-D}{a} P_i - \frac{E_{ij}}{a} P_j; \qquad P_i = a(1+D)P_{\hat{i}} + aE_{ij}P_{\hat{j}}. \tag{17}$$

Phase space coordinates. We may now write the phase space density f of particles as a function of position and momentum. Phase space is 6-dimensional (7 if we include time), and one can write down many coordinate systems for it:

- *Canonical coordinates*, (x^i, P_i): these have the advantage of being canonically conjugate $\{x^i, P_j\} = \delta^i_j$ (homework exercise), so the total number of particles in a given region is given by simple integration:

$$N_{\text{particles}} = \frac{g}{(2\pi)^3} \int f(x^i, P_i)\, d^3x^i\, d^3P_i. \tag{18}$$

- *Observer-frame momenta*, $(x^i, P_{\hat{i}})$: these are the most intuitive since $P_{\hat{i}}$ is a measured momentum. But these are not canonically conjugate, and don't even conserve volume: there is a phase space volume Jacobian:

$$J = \begin{vmatrix} dx^i/dx^j & dx^i/dP_{\hat{j}} \\ dP_i/dx^j & dP_i/dP_{\hat{j}} \end{vmatrix} = \left| \frac{dP_i}{dP_{\hat{j}}} \right| = a^3(1+3D). \tag{19}$$

  Therefore in these coordinates the number of particles in a given region is:

$$N_{\text{particles}} = \frac{g}{(2\pi)^3} \int a^3(1+3D) f(x^i, P_{\hat{i}})\, d^3x^i\, d^3P_{\hat{i}}. \tag{20}$$

- *Spherical coordinates*, $(x^i, p, \hat{p}^i)$: Dodelson introduces the spherical coordinates in momentum space,

$$p = \sqrt{P^{\hat{i}} P_{\hat{i}}}; \quad \hat{p}^i = \frac{P^{\hat{i}}}{p}, \tag{21}$$

  so that $\hat{p}^i$ is a unit vector that lives on the unit sphere, and p is the magnitude of the 3-momentum seen by an observer. This is useful for observations because we're an observer and we measure magnitudes of momenta (photon frequency; $p = 2\pi\nu$) and directions in our reference frame. Our observations are at $a = 1$ so we directly observe p. The Jacobian is now $(1+3D)p^2$, and the number of particles in a region of phase space is:

$$N_{\text{particles}} = \frac{g}{(2\pi)^3} \int a^3(1+3D)p^2\, f(x^i, p, \hat{p}^i)\, d^3x^i\, dp\, d^2\hat{p}^i. \tag{22}$$

Dodelson (and us) work in the spherical coordinates in momentum space (but Cartesian coordinates in position space).

The Boltzmann equation. In the absence of collisions, the Boltzmann equation says that the phase space density of particles is conserved along a trajectory:

$$\frac{df}{d\eta} \equiv \frac{\partial f}{\partial \eta} + \frac{dx^i}{d\eta}\frac{\partial f}{\partial x^i} + \frac{dp}{d\eta}\frac{\partial f}{\partial p} + \frac{d\hat{p}^i}{d\eta}\frac{\partial f}{\partial \hat{p}^i} = 0. \tag{23}$$

Here f is a function $f(x^i, p, \hat{p}^i; \eta)$, $df/d\eta$ means that we take the derivative of f with respect to η but changing the coordinates and momentum to follow a specific trajectory; and $\partial f/\partial \eta$ means that we take the derivative with $x^i, p, \hat{p}^i$ fixed. The factors $dx^i/d\eta$, etc. come from the chain rule and are given by the equations of motion. Since $\hat{p}^i$ is a unit vector, $\partial f/\partial \hat{p}^i$ should be thought of as a tangent vector to the unit sphere.

We can write the Boltzmann equation with dots:

$$\dot{f} + \dot{x}^i \frac{\partial f}{\partial x^i} + \dot{p}\frac{\partial f}{\partial p} + \dot{\hat{p}}^i \frac{\partial f}{\partial \hat{p}^i} = 0. \tag{24}$$

In order to go further we'll need to start using equations of motion. One of these is trivial: in the unperturbed universe, $\dot{\hat{p}}^i = 0$ (no change in 3-direction of propagation), so $\dot{\hat{p}}^i$ is first order in perturbation theory. But $\partial f/\partial \hat{p}^i$ is also first order, so the product is second order. Therefore the last term drops out. Rearranging, and solving for $\dot{f}$:

$$\dot{f} = -\dot{x}^i \frac{\partial f}{\partial x^i} - \dot{p}\frac{\partial f}{\partial p}. \tag{25}$$

We can also see that $\partial f/\partial x^i$ is first order in perturbation theory, since it is zero in the unperturbed universe. Therefore we only need $\dot{x}^i$ to zeroeth order. In the unperturbed universe, the requirement of a null trajectory gives

$$g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = 0 \quad \rightarrow \quad \dot{x}^i \dot{x}^i = 1, \tag{26}$$

and the direction is $\hat{p}^i$, so $\dot{x}^i = \hat{p}^i$:

$$\dot{f} = -\hat{p}^i \frac{\partial f}{\partial x^i} - \dot{p}\frac{\partial f}{\partial p}. \tag{27}$$

However $\partial f/\partial p$ is nonzero so we need to get the first-order perturbation theory result for $\dot{p}$.

The momentum evolution. Our last step here will be to consider the momentum evolution of the particle. Recall that the physical momentum of the particle is p, and that for a massless particle this is the energy $-P^\mu u_\mu$. Then

$$\dot{p} = \frac{d}{d\eta}(-u_\mu P^\mu). \tag{28}$$

In terms of the affine parameter λ, $P^0 = d\eta/d\lambda$, so

$$\dot{p} = \frac{1}{P^0}\frac{d}{d\lambda}(-u_\mu P^\mu). \tag{29}$$

This is a rather messy calculation, which you will do on the homework; the answer is

$$\dot{p} = -\left(aH + \hat{p}^i\frac{\partial A}{\partial x^i} + \hat{p}^i\hat{p}^j\frac{\partial B_i}{\partial x^j} + \dot{D} + \hat{p}^i\hat{p}^j\dot{E}_{ij}\right)p. \tag{30}$$

The Boltzmann equation thus simplifies to:

$$\dot{f}(x^i,p,\hat{p}^i;\eta) = -\hat{p}^i\frac{\partial f}{\partial x^i} + p\frac{\partial f}{\partial p}\left(aH + \hat{p}^i\frac{\partial A}{\partial x^i} + \hat{p}^i\hat{p}^j\frac{\partial B_i}{\partial x^j} + \dot{D} + \hat{p}^i\hat{p}^j\dot{E}_{ij}\right). \tag{31}$$

Dodelson does this for a special coordinate system known as the Newtonian gauge where $A = \Psi$, $D = \Phi$, and $B = E = 0$. This gauge exists for density perturbations, but does not allow for gravitational waves or vorticity. We'll maintain generality in class, but for Dodelson's choice of gauge, the above equation reduces to

$$\dot{f}(x^i,p,\hat{p}^i;\eta) = -\hat{p}^i\frac{\partial f}{\partial x^i} + p\frac{\partial f}{\partial p}\left(aH + \hat{p}^i\frac{\partial \Psi}{\partial x^i} + \dot{\Phi}\right), \tag{32}$$

which is Eq. (4.33) in Dodelson.

3 Perturbed Boltzmann equation

In order to do perturbation theory, we will want to write f as a homogeneous solution plus a perturbation. We've already done the hard work of writing the evolution equation; but it will simplify matters later on if we choose our perturbation variables carefully.

Perturbation variables. The homogeneous Universe solution for f is the blackbody:

$$f^{(0)}(x^i,p,\hat{p}^i;\eta) = \frac{1}{e^{p/T}-1}, \tag{33}$$

where $T \propto 1/a$ is the background photon temperature. The superscript $^{(0)}$ denotes the background value. Rather than using the obvious choice of perturbation variable δf, we will write:

$$f(x^i,p,\hat{p}^i;\eta) = \left\{\exp\frac{p}{T(\eta)[1+\Theta(x^i,p,\hat{p}^i;\eta)]} - 1\right\}^{-1}, \tag{34}$$

where the perturbation variable is Θ. This is called the *blackbody temperature perturbation* (also called the thermodynamic temperature perturbation). It is related to δf by differentiation:

$$\delta f = -\frac{\partial f^{(0)}}{\partial p}p\Theta = \frac{p}{T}\frac{e^{p/T}}{(e^{p/T}-1)^2}\Theta, \tag{35}$$

in first-order perturbation theory.

Evolution equation. We can determine how Θ evolves by plugging our formula for f in terms of η and Θ into the Boltzmann equation and keeping terms through first order. The left-hand side is

$$\dot{f} = \dot{f}^{(0)} + \delta\dot{f} = \dot{f}^{(0)} - \Theta\dot{T}\frac{\partial}{\partial T}\left(\frac{\partial f^{(0)}}{\partial p}p\right) - \frac{\partial f^{(0)}}{\partial p}p\dot{\Theta}. \tag{36}$$

Then we have:

$$\begin{aligned}
&\dot{f}^{(0)} - \Theta\dot{T}\frac{\partial}{\partial T}\left(\frac{\partial f^{(0)}}{\partial p}p\right) - \frac{\partial f^{(0)}}{\partial p}p\dot{\Theta} = -\hat{p}^i\frac{p}{T}\frac{e^{p/T}}{(e^{p/T}-1)^2}\frac{\partial\Theta}{\partial x^i} \\
&+p\left[\frac{\partial f^{(0)}}{\partial p} + \Theta\frac{\partial}{\partial p}\left(\frac{p}{T}\frac{e^{p/T}}{(e^{p/T}-1)^2}\right) + \frac{\partial\Theta}{\partial p}\frac{p}{T}\frac{e^{p/T}}{(e^{p/T}-1)^2}\right] \\
&\times\left(aH + \hat{p}^i\frac{\partial A}{\partial x^i} + \hat{p}^i\hat{p}^j\frac{\partial B_i}{\partial x^j} + \dot{D} + \hat{p}^i\hat{p}^j\dot{E}_{ij}\right).
\end{aligned} \tag{37}$$

This looks like a mess but in fact there are some cancellations:

- $\dot{f}^{(0)}$ (LHS) cancels $Hp\partial f^{(0)}/\partial p$ (RHS). This is in fact just the cancellation of zeroeth-order terms; it reduces our problem to:

$$\begin{aligned}
&-\Theta\dot{T}\frac{\partial}{\partial T}\left(\frac{\partial f^{(0)}}{\partial p}p\right) - \frac{\partial f^{(0)}}{\partial p}p\dot{\Theta} = -\hat{p}^i\frac{p}{T}\frac{e^{p/T}}{(e^{p/T}-1)^2}\frac{\partial\Theta}{\partial x^i} \\
&+aHp\left[\Theta\frac{\partial}{\partial p}\left(\frac{p}{T}\frac{e^{p/T}}{(e^{p/T}-1)^2}\right) + \frac{\partial\Theta}{\partial p}\frac{p}{T}\frac{e^{p/T}}{(e^{p/T}-1)^2}\right] \\
&+p\frac{\partial f^{(0)}}{\partial p}\left(\hat{p}^i\frac{\partial A}{\partial x^i} + \hat{p}^i\hat{p}^j\frac{\partial B_i}{\partial x^j} + \dot{D} + \hat{p}^i\hat{p}^j\dot{E}_{ij}\right).
\end{aligned} \tag{38}$$

- Divide everything by $p\partial f^{(0)}/\partial p$. Recall that $-pe^{p/T}/T(e^{p/T}-1)^2 = p\partial f^{(0)}/\partial p$:

$$\begin{aligned}
-\Theta\dot{T}\frac{\partial}{\partial T}\ln\left(\frac{\partial f^{(0)}}{\partial p}p\right) - \dot{\Theta} &= \hat{p}^i\frac{\partial\Theta}{\partial x^i} + aHp\left[-\Theta\frac{\partial}{\partial p}\ln\left(\frac{\partial f^{(0)}}{\partial p}p\right) - \frac{\partial\Theta}{\partial p}\right] \\
&\quad +\hat{p}^i\frac{\partial A}{\partial x^i} + \hat{p}^i\hat{p}^j\frac{\partial B_i}{\partial x^j} + \dot{D} + \hat{p}^i\hat{p}^j\dot{E}_{ij}.
\end{aligned} \tag{39}$$

- The operator $\dot{T}\partial/\partial T$ on the LHS is equal to $-aHT\partial/\partial T$, and because $f^{(0)}$ depends only on the ratio p/T this cancels the partial derivative term on the RHS, leaving us with:

$$-\dot{\Theta} = \hat{p}^i\frac{\partial\Theta}{\partial x^i} - aHp\frac{\partial\Theta}{\partial p} + \hat{p}^i\frac{\partial A}{\partial x^i} + \hat{p}^i\hat{p}^j\frac{\partial B_i}{\partial x^j} + \dot{D} + \hat{p}^i\hat{p}^j\dot{E}_{ij}. \tag{40}$$

We can push through a minus sign and get the full evolution equation for the temperature perturbation:

$$\dot{\Theta} = -\hat{p}^i\frac{\partial\Theta}{\partial x^i} + aHp\frac{\partial\Theta}{\partial p} - \hat{p}^i\frac{\partial A}{\partial x^i} - \hat{p}^i\hat{p}^j\frac{\partial B_i}{\partial x^j} - \dot{D} - \hat{p}^i\hat{p}^j\dot{E}_{ij}. \tag{41}$$

What's great about this equation is that the source terms for Θ depend only on the metric perturbations, they don't depend on p. We will show later that the collision terms also don't depend on p, so that $\partial\Theta/\partial p$ is forever zero. This will greatly simplify our calculations.

4 Thomson scattering and the collision term

Up until now we've assumed the photons are collisionless. But that's wrong in the early Universe because they will scatter off the free electrons. So that means that now we need to go back to our Boltzmann equation and put in a collision term.

Form of the collision term. The collision term is:

$$\frac{df}{d\eta} = C[f], \tag{42}$$

where $C[f]$ is the number of photons scattered into a given mode per unit conformal time, minus the number of photons scattered out. We can re-write it as:

$$\begin{aligned} C[f](x^i, p, \hat{p}^i; \eta) \quad &= \quad -a\Gamma_{\rm scat}(x^i, p, \hat{p}^i; \eta) f(x^i, p, \hat{p}^i; \eta) \\ &\quad +a \int \Pi(p', \hat{p}'^i \to p, \hat{p}^i)\Gamma_{\rm scat}(x^i, p', \hat{p}'^i; \eta) f(x^i, p', \hat{p}'^i; \eta) \\ &\quad \times p'^2\, dp'\, d^2\hat{p}'^i, \end{aligned} \tag{43}$$

where $\Gamma_{\rm scat}$ is the number of Thomson scatterings per unit time t (the a is due to the conformal part: $dt/d\eta = a$), and $\Pi(p', \hat{p}'^i \to p, \hat{p}^i)$ is the 3D probability distribution of the final momentum p of a scattered photon with initial momentum p'.

We can simplify this by using the principle of reciprocity: ignoring electron recoil the scattering rate from the primed to the unprimed momentum state must equal the scattering rate from unprimed to primed (this is trivial to see in the baryon rest frame where these two rates must be the same by isotropy, but the principle is more general):

$$\Gamma_{\rm scat}(x^i, p, \hat{p}^i; \eta)\Pi(p, \hat{p}^i \to p', \hat{p}'^i) = \Gamma_{\rm scat}(x^i, p', \hat{p}'^i; \eta)\Pi(p', \hat{p}'^i \to p, \hat{p}^i). \tag{44}$$

Since probability is normalized,

$$\int \Pi(p, \hat{p}^i \to p', \hat{p}'^i)\, p'^2\, dp'\, d^2\hat{p}'^i = 1, \tag{45}$$

we can simplify the collision term:

$$\begin{aligned} C[f](x^i, p, \hat{p}^i; \eta) \quad &= \quad -a\Gamma_{\rm scat}(x^i, p, \hat{p}^i; \eta) \int \Pi(p, \hat{p}^i \to p', \hat{p}'^i) \\ &\quad \times [f(x^i, p, \hat{p}^i; \eta) - f(x^i, p', \hat{p}'^i; \eta)] p'^2\, dp'\, d^2\hat{p}'^i. \end{aligned} \tag{46}$$

In the unperturbed universe, the photon cannot change its energy during scattering, so $p' = p$ and $f(x^i, p, \hat{p}^i; \eta) = f(x^i, p', \hat{p}'^i; \eta)$. This leads to the important conclusion the integral in Eq. (46) is (at most) first order in the perturbations, so the prefactor can be computed in the unperturbed universe. That is, we can use the unperturbed scattering rate $n_e \sigma_T$:

$$\begin{aligned} C[f](x^i, p, \hat{p}^i; \eta) &= -a n_e \sigma_T \int \Pi(p, \hat{p}^i \to p', \hat{p}'^i) \\ &\quad \times [f(x^i, p, \hat{p}^i; \eta) - f(x^i, p', \hat{p}'^i; \eta)] p'^2 \, dp' \, d^2\hat{p}'^i. \end{aligned} \tag{47}$$

The scattering probability distribution. In order to proceed we need the function $\Pi(p, \hat{p}^i \to p', \hat{p}'^i)$. This is most easily accomplished first in the rest frame of the baryons, and then in the more general case.

In the rest frame of the baryons, we learned in undergraduate physics that Thomson scattering has an angular distribution,

$$\frac{dP}{d^2\hat{\mathbf{n}}'} = \frac{3}{16\pi}[1 + (\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}')^2], \tag{48}$$

where $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}'$ are the incoming and outgoing photon directions. However, the baryons are moving at some velocity $\mathbf{v}_b$. We can relate $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}'$ to the lab-frame (actually normal-frame) directions $\hat{\mathbf{p}}$ and $\hat{\mathbf{p}}'$ by a Lorentz transformation. The photon has energy ω in the baryon frame, so in the baryon frame its 4-momentum is:

$$(\omega, \omega\hat{\mathbf{n}}). \tag{49}$$

The lab-frame 4-momentum is obtained by multiplying by the Lorentz transformation matrix, which to order v_b is

$$\begin{pmatrix} 1 & \mathbf{v}_b \\ \mathbf{v}_b & 1 \end{pmatrix} \begin{pmatrix} \omega \\ \omega\hat{\mathbf{n}} \end{pmatrix} = \begin{pmatrix} \omega(1 + \mathbf{v}_b \cdot \hat{\mathbf{n}}) \\ \omega(\hat{\mathbf{n}} + \mathbf{v}_b) \end{pmatrix}. \tag{50}$$

The lab-frame direction of propagation is the photon momentum divided by its energy,

$$\hat{\mathbf{p}} = \frac{\hat{\mathbf{n}} + \mathbf{v}_b}{1 + \mathbf{v}_b \cdot \hat{\mathbf{n}}}, \tag{51}$$

and similarly

$$\hat{\mathbf{p}}' = \frac{\hat{\mathbf{n}}' + \mathbf{v}_b}{1 + \mathbf{v}_b \cdot \hat{\mathbf{n}}'}. \tag{52}$$

To get the final momentum probability distribution we need to convert $dP/d^2\hat{\mathbf{n}}'$ into $dP/d^2\hat{\mathbf{p}}'$, which involves taking a Jacobian:

$$\frac{dP}{d^2\hat{\mathbf{p}}'} = \frac{3}{16\pi}[1 + (\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}')^2] \left| \frac{d^2\hat{\mathbf{n}}'}{d^2\hat{\mathbf{p}}'} \right|. \tag{53}$$

We also need the probability distribution for the magnitude of the momentum, p'. In fact since the photon frequency is conserved in the baryon rest frame,

p' is determined by p, $\hat{\mathbf{p}}$, and $\hat{\mathbf{p}}'$. This is because the Lorentz transformation showed:

$$p = \omega(1 + \mathbf{v}_b \cdot \hat{\mathbf{n}}) \quad \text{and} \quad p' = \omega(1 + \mathbf{v}_b \cdot \hat{\mathbf{n}}'). \tag{54}$$

From these we can solve for p':

$$p' = p[1 + \mathbf{v}_b \cdot (\hat{\mathbf{n}}' - \hat{\mathbf{n}})] = p[1 + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}})], \tag{55}$$

where we have worked to first order. Then the full 3D probability distribution is

$$\Pi(p, \hat{p}^i \to p', \hat{p}'^i) = \frac{1}{p'^2} \frac{dP}{d^2\hat{\mathbf{p}}'} \delta\left(p' - p[1 + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}})]\right). \tag{56}$$

Collision term: results. We can plug this formula into Eq. (47), and get:

$$C[f](x^i, p, \hat{p}^i; \eta) = -an_e\sigma_T \int \frac{dP}{d^2\hat{\mathbf{p}}'} \left\{ f(x^i, p, \hat{p}^i; \eta) - f\left(x^i, p[1 + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}})], \hat{p}'^i; \eta\right) \right\} d^2\hat{p}'^i. \tag{57}$$

The difference of phase space densities can be simplified by taking the first-order perturbation results:

$$f(p, \hat{p}^i) = f^{(0)} - p\frac{\partial f^{(0)}}{\partial p} \Theta(p, \hat{p}^i) \tag{58}$$

and

$$f\left(p[1 + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}})], \hat{p}'^i\right) = f^{(0)} + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}}) p \frac{\partial f^{(0)}}{\partial p} - p\frac{\partial f^{(0)}}{\partial p} \Theta(p, \hat{p}'^i); \tag{59}$$

so

$$f(p, \hat{p}^i) - f\left(p[1 + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}})], \hat{p}'^i\right) = -p\frac{\partial f^{(0)}}{\partial p} \left[\Theta(p, \hat{p}^i) - \Theta(p, \hat{p}'^i) + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}})\right]. \tag{60}$$

The collision term now simplifies to

$$C[f](x^i, p, \hat{p}^i; \eta) = an_e\sigma_T p \frac{\partial f^{(0)}}{\partial p} \int \frac{dP}{d^2\hat{\mathbf{p}}'} \left[\Theta(p, \hat{p}^i) - \Theta(p, \hat{p}'^i) + \mathbf{v}_b \cdot (\hat{\mathbf{p}}' - \hat{\mathbf{p}})\right] d^2\hat{p}'^i. \tag{61}$$

Since the quantity in square brackets is a perturbation, we can use the unperturbed formula for $dP/d^2\hat{\mathbf{p}}'$

$$\frac{dP}{d^2\hat{\mathbf{p}}'} = \frac{3}{16\pi}[1 + (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')^2]. \tag{62}$$

When we substitute this in, the $\mathbf{v}_b \cdot \mathbf{p}'$ term goes away (it is odd under $\mathbf{p}' \to -\mathbf{p}'$).

$$C[f](x^i, p, \hat{p}^i; \eta) = \frac{3}{16\pi} an_e\sigma_T p \frac{\partial f^{(0)}}{\partial p} \int [1 + (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')^2] \left[\Theta(p, \hat{p}^i) - \Theta(p, \hat{p}'^i) - \mathbf{v}_b \cdot \hat{\mathbf{p}}\right] d^2\hat{p}'^i. \tag{63}$$

When we write the evolution equation in terms of Θ instead of f, we need to multiply this by

$$\frac{d\Theta}{df} = \left(\frac{df}{d\Theta}\right)^{-1} = \left(-p\frac{\partial f^{(0)}}{\partial p}\right)^{-1}, \tag{64}$$

so

$$C[\Theta](x^i, p, \hat{p}^i; \eta) = -\frac{3}{16\pi} a n_e \sigma_T \int [1+(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}')^2] \left[\Theta(p, \hat{p}^i) - \Theta(p, \hat{p}'^i) - \mathbf{v}_b \cdot \hat{\mathbf{p}}\right] d^2\hat{p}'^i. \tag{65}$$

Putting this together with the collisionless terms, we get:

$$\begin{aligned} \dot{\Theta} &= -\hat{p}^i \frac{\partial \Theta}{\partial x^i} + aHp\frac{\partial \Theta}{\partial p} - \hat{p}^i \frac{\partial A}{\partial x^i} - \hat{p}^i \hat{p}^j \frac{\partial B_i}{\partial x^j} - \dot{D} - \hat{p}^i \hat{p}^j \dot{E}_{ij} \\ &\quad -\frac{3}{16\pi} a n_e \sigma_T \int [1 + (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')^2] \left[\Theta(p, \hat{p}^i) - \Theta(p, \hat{p}'^i) - \mathbf{v}_b \cdot \hat{\mathbf{p}}\right] d^2\hat{p}'^i. \end{aligned} \tag{66}$$

If the CMB is initially a blackbody everywhere (but possibly not at the same temperature in every location), which should happen if it is thermalized in the early universe, then initially Θ is independent of p, i.e. $\partial\Theta/\partial p = 0$. As one can see from this equation this situation is maintained during the subsequent evolution, so we may drop the explicit dependence of Θ on p. This removes some additional terms from the equation.

$$\begin{aligned} \dot{\Theta} &= -\hat{p}^i \frac{\partial \Theta}{\partial x^i} - \hat{p}^i \frac{\partial A}{\partial x^i} - \hat{p}^i \hat{p}^j \frac{\partial B_i}{\partial x^j} - \dot{D} - \hat{p}^i \hat{p}^j \dot{E}_{ij} \\ &\quad -\frac{3}{16\pi} a n_e \sigma_T \int [1 + (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')^2] \left[\Theta(\hat{p}^i) - \Theta(\hat{p}'^i) - \mathbf{v}_b \cdot \hat{\mathbf{p}}\right] d^2\hat{p}'^i. \end{aligned} \tag{67}$$

A further simplification is achieved by performing the $\hat{\mathbf{p}}'$ integration on the $\Theta(\hat{p}^i)$ and $\mathbf{v}_b \cdot \hat{\mathbf{p}}$ terms.

$$\begin{aligned} \dot{\Theta} &= -\hat{p}^i \frac{\partial \Theta}{\partial x^i} - \hat{p}^i \frac{\partial A}{\partial x^i} - \hat{p}^i \hat{p}^j \frac{\partial B_i}{\partial x^j} - \dot{D} - \hat{p}^i \hat{p}^j \dot{E}_{ij} \\ &\quad - a n_e \sigma_T (\Theta - \mathbf{v}_b \cdot \hat{\mathbf{p}}) + \frac{3}{16\pi} a n_e \sigma_T \int [1 + (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')^2] \Theta(\hat{p}'^i) d^2\hat{p}'^i. \end{aligned} \tag{68}$$

Compare to Dodelson (4.56).