Chapter 2

The Boltzmann Equation

Wonderful, the process which fashions and transforms us! What is it going to turn you into next, in what direction will it use you to go?

–Chuang-tzu, 6

The study of the formation and evolution of CMB fluctuations in both real and frequency space begins with the radiative transport, or Boltzmann equation. In this pedagogically motivated chapter, we will examine its derivation. The Boltzmann equation written in abstract form as

$$\frac{df}{dt} = C[f] \tag{2.1}$$

contains a collisionless part df/dt, which deals with the effects of gravity on the photon distribution function f, and collision terms $C[f]$, which account for its interactions with other species in the universe. The collision terms in the Boltzmann equation have several important effects. Most importantly, Compton scattering couples the photons and baryons, keeping the two in kinetic equilibrium. This process along with interactions that create and destroy photons determines the extent to which the CMB can be thermalized. We will examine these issues more fully in §3 where we consider spectral distortions. Compton scattering also governs the evolution of inhomogeneities in the CMB temperature which lead to anisotropies on the sky. This will be the topic of §6 and §7.

In this chapter, we will first examine gravitational interactions and show that the photon energy is affected by gradients in the gravitational potential, *i.e.* the gravitational redshift, and changes in the spatial metric, *i.e.* the cosmological redshift from the scale factor and dilation effects due to the space curvature perturbation. Compton scattering in its non-relativistic limit can be broken down in a perturbative expansion based on the energy transfer between the photons and electrons. We will examine the importance of each term in turn and derive its effects on spectral distortions and temperature inhomogeneities in the CMB.

2.1 Gravitational Interactions

Gravity is the ultimate source of spatial fluctuations in the photon distribution and the cause of the adiabatic cooling of the photon temperature from the expansion. Its effects are described by the collisionless Boltzmann, or Liouville, equation which controls the evolution of the photon distribution $f(x,p)$ as the photons stream along their geodesics. Here x and p are the 4-position and 4-momentum of the photons respectively. It is given by

$$\frac{df}{dt} = \frac{\partial f}{\partial x^\mu}\frac{dx^\mu}{dt} + \frac{\partial f}{\partial p^\mu}\frac{dp^\mu}{dt} = 0. \tag{2.2}$$

In other words, the phase space density of photons is conserved along its trajectory. The gravitational effects are hidden in the time dependence of the photon momentum. The solution to equation (2.2) is non-trivial since the photons propagate in a metric distorted by the lumpy distribution of matter. To evaluate its effect explicitly, we need to examine the geodesic equation in the presence of arbitrary perturbations.

2.1.1 Metric Fluctuations

The big bang model assumes that the universe is homogeneous and isotropic on the large scale. All such cases can be described by the Friedman-Robertson-Walker metric, where the line element takes the form

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + (a/a_0)^2\gamma_{ij}dx^i dx^j, \tag{2.3}$$

with γ_{ij} as the background three-metric on a space of constant curvature $K = -H_0^2(1-\Omega_0-\Omega_\Lambda)$ and the scale factor is related to the redshift by $a/a_0 = (1+z)^{-1}$. We will be mainly interested in the flat $K=0$ and negatively curved (open) $K<0$ cases. For these cases, a convenient representation of the three-metric which we will have occasion to use is the radial representation

$$\gamma_{ij}dx^i dx^j = -K^{-1}[d\chi^2 + \sinh^2\chi(d\theta^2 + \sin^2\theta d\phi^2)], \tag{2.4}$$

where the radial coordinate is scaled to the curvature length $(-K)^{-1/2}$.

Small scalar perturbations to the background metric can in general be expressed by two spatially varying functions. The exact form of the metric fluctuations varies with the choice of hypersurface on which these perturbations are defined, *i.e.* the gauge. We will discuss the subtleties involving the choice of gauge in §4.3. For now, let us derive the evolution equations for the photons using the conformal Newtonian gauge where the metric takes the form

$$\begin{aligned} g_{00} &= -[1+2\Psi(\mathbf{x},t)], \\ g_{ij} &= (a/a_0)^2[1+2\Phi(\mathbf{x},t)]\gamma_{ij}. \end{aligned} \tag{2.5}$$

Note that Ψ can be interpreted as a Newtonian potential. Φ is the fractional perturbation to the spatial curvature as the form of equation (2.4) shows. As we shall see in §4.2.6, they are related by the Einstein equations as $\Phi = -\Psi$ when pressure may be neglected. We will therefore often loosely refer to both as "gravitational potentials."

The geodesic equation for the photons is

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0, \tag{2.6}$$

where Γ is the Christoffel symbol. The affine parameter λ is chosen such that the photon energy satisfies $p^0 = dx^0/d\lambda$. Since the photon momentum is given by

$$\frac{p^i}{p^0} = \frac{dx^i}{dt}, \tag{2.7}$$

the geodesic equation then becomes

$$\frac{dp^i}{dt} = g^{i\nu}\left(\frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\nu} - \frac{\partial g_{\nu\alpha}}{\partial x^\beta}\right)\frac{p^\alpha p^\beta}{p^0}. \tag{2.8}$$

This equation determines the gravitational effects on the photons in the presence of perturbations as we shall now show.

2.1.2 Gravitational Redshift and Dilation

Let us rewrite the Boltzmann equation in terms of the energy p and direction of propagation of the photons γ^i in a frame that is locally orthonormal on constant time hypersurfaces,

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}\frac{dx^i}{dt} + \frac{\partial f}{\partial p}\frac{dp}{dt} + \frac{\partial f}{\partial \gamma^i}\frac{d\gamma^i}{dt} = 0. \tag{2.9}$$

Notice that $d\gamma^i/dt \neq 0$ only in the presence of curvature from K or Φ because otherwise photon geodesics are straight lines. Since the anisotropy $\partial f/\partial\gamma^i$ is already first order in the perturbation, it may be dropped if the background curvature $K=0$. In the presence of negative curvature, it makes photon geodesics deviate from each other exponentially with distance. Two photons which are observed to have a given angular separation were in the past separated by a larger (comoving) physical distance than euclidean analysis would imply. We shall see that this property allows the curvature of the universe to be essentially read off of anisotropies in the CMB. Formal elements of this effect are discussed in §4.2.4.

On the other hand, the redshift term dp/dt is important in all cases – even in the absence of perturbations. Since static curvature effects are unimportant in determining the redshift contributions, we will assume in the following that the background three-metric is flat, *i.e.* $\gamma_{ij} = \delta_{ij}$ without loss of generality. The energy and direction of propagation are explicitly given by

$$p^2 = p^i p_i, \qquad \gamma^i = \frac{a}{a_0}\frac{p^i}{p}(1+\Phi), \tag{2.10}$$

which implies $p^0 = (1+\Psi)p$. The geodesic equation (2.8) then yields to first order in the fluctuations

$$\frac{1}{p}\frac{dp^0}{dt} = -\left(\frac{\partial\Psi}{\partial t} + \frac{da}{dt}\frac{1}{a}(1-\Psi) + \frac{\partial\Phi}{\partial t} + 2\frac{\partial\Psi}{\partial x^i}\frac{a_0}{a}\gamma^i\right). \tag{2.11}$$

From this relation, we obtain

$$\begin{aligned}\frac{1}{p}\frac{dp}{dt} &= \frac{1}{p}\frac{dp^0}{dt}(1+\Psi) + \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi}{\partial x^i}\frac{dx^i}{dt} \\ &= -\left(\frac{da}{dt}\frac{1}{a} + \frac{\partial \Phi}{\partial t} + \frac{\partial \Psi}{\partial x^i}\frac{a_0}{a}\gamma^i\right),\end{aligned} \tag{2.12}$$

which governs the gravitational and cosmological redshift effects on the photons.

Now let us discuss the physical interpretation of the energy equation (2.12). Consider first a small region where we can neglect the spatial variation of Ψ and Φ. In the presence of a gravitational potential, clocks naturally ticking at intervals Δt run slow by the dilation factor (see *e.g.* [173]),

$$\delta t = (-g_{00})^{-1/2}\Delta t \simeq (1-\Psi)\Delta t. \tag{2.13}$$

For light emitted from the point 1, crests leave spaced by $\delta t_1 = [1-\Psi(t_1)]\Delta t$. If they arrive at the origin spaced by δt_0, they should be compared with a local oscillator with crests spaced as $[1-\Psi(t_0)]\Delta t$, *i.e.* the shift in frequency (energy) is

$$\frac{p_1}{p_0} = [1+\Psi(t_1) - \Psi(t_0)]\frac{\delta t_1}{\delta t_0}. \tag{2.14}$$

Now we have to calculate the in-transit delay factor $\delta t_1/\delta t_0$. Since null geodesics from the origin are radial in the FRW metric, choose angular coordinates such that along the $\chi(t)$ geodesic

$$-(1+2\Psi)dt^2 + (a/a_0)^2(-K)^{-1}(1+2\Phi)d\chi = 0. \tag{2.15}$$

A wave crest emitted at (t_1, χ_1) is received at $(t_0, 0)$ where the two are related by

$$\int_{t_1}^{t_0}(1+\Psi-\Phi)\frac{a_0}{a}dt = \int_0^{\chi_1}(-K)^{1/2}d\chi. \tag{2.16}$$

At χ_1, the source emits a second crest after δt_1 which is received at the origin at $t_0 + \delta t_0$ where

$$\int_{t_1}^{t_0}(1+\Psi-\Phi)\frac{a_0}{a}dt = \int_{t_1+\delta t_0}^{t_0+\delta t_0}(1+\Psi-\Phi)\frac{a_0}{a}dt. \tag{2.17}$$

This can be manipulated to give

$$\int_{t_1}^{t_1+\delta t_1}(1+\Psi-\Phi)\frac{a_0}{a}dt = \int_{t_0}^{t_0+\delta t_0}(1+\Psi-\Phi)\frac{a_0}{a}dt, \tag{2.18}$$

or

$$\frac{\delta t_1}{\delta t_0} = \frac{a(t_1)}{a(t_0)}\frac{1-\Psi(t_1)+\Phi(t_1)}{1-\Psi(t_0)+\Phi(t_0)}. \tag{2.19}$$

Inserting this into equation (2.14), the ratio of energies becomes

$$\frac{p_1}{p_0} = \frac{a(t_1)[1+\Phi(t_1)]}{a(t_0)[1+\Phi(t_0)]}. \tag{2.20}$$

Notice that the space curvature Φ but *not* the Newtonian potential Ψ enters this expression. This is easy to interpret. Heuristically, the wavelength of the photon itself scales with the space-space component of the metric, *i.e.* $a(1+\Phi)$. In the background, this leads to the universal redshift of photons with the expansion. The presence of a space curvature perturbation Φ also stretches space. We shall see that it arises from density fluctuations through the Einstein equations (see §4.2.6). Overdense regions create positive curvature and underdense regions negative curvature. From equation (2.20), the rate of change of the energy is therefore given by

$$\frac{1}{p}\frac{\partial p}{\partial t} = -\frac{da}{dt}\frac{1}{a} - \frac{\partial \Phi}{\partial t}, \tag{2.21}$$

which explains two of three of the terms in equation (2.12).

Now let us consider the effects of spatial variations. Equation (2.14) becomes

$$\frac{p_1}{p_0} = [1+\Psi(t_1,\chi_1) - \Psi(t_0,0)]\frac{\delta t_1}{\delta t_0}. \tag{2.22}$$

The additional factor here is the potential difference in space. Photons suffer gravitational redshifts climbing in and out of potentials. Thus the gradient of the potential along the direction of propagation leads to a redshift of the photons, *i.e.*

$$\begin{aligned}\frac{1}{p}\frac{\partial p}{\partial x^i}\frac{dx^i}{dt} &= -\frac{\partial \Psi}{\partial x^i}\frac{dx^i}{dt} \\ &= -\frac{\partial \Psi}{\partial x^i}\frac{a_0}{a}\gamma^i,\end{aligned} \tag{2.23}$$

as required. This explains why a uniform Ψ does not lead to an effect on the photon energy and completes the physical interpretation of equation (2.12).

2.1.3 Collisionless Brightness Equation

The fractional shift in frequency from gravitational effects is independent of frequency $p' = p(1+\delta p/p)$. Thus, a blackbody distribution will remain a blackbody,

$$\begin{aligned}f'(p') = f(p) &= \{\exp[p'/T(1+\delta p/p)] - 1\}^{-1} \\ &= \{\exp[p'/T'] - 1\}^{-1},\end{aligned} \tag{2.24}$$

with a temperature shift $\delta T/T = \delta p/p$. Let us therefore integrate the collisionless Boltzmann equation over energy, *i.e.* define

$$4\Theta \equiv \frac{1}{\pi^2 \rho_\gamma}\int p^3 dp f - 1 = \frac{\delta\rho_\gamma}{\rho_\gamma}, \tag{2.25}$$

where ρ_γ is the spatially and directionally averaged energy density of the photons. Since $\rho_\gamma \propto T^4$, $\Theta(\eta, \mathbf{x}, \boldsymbol{\gamma})$ is the fractional temperature fluctuation for a blackbody.

Employing equation (2.12) in (2.9) and integrating over frequencies, we obtain the collisionless Boltzmann (or brightness) equation,

$$\dot\Theta + \gamma^i\frac{\partial}{\partial x^i}(\Theta+\Psi) + \dot\gamma^i\frac{\partial}{\partial \gamma^i}\Theta + \dot\Phi = 0, \tag{2.26}$$

where the overdots represent derivatives with respect to conformal time $d\eta = dt/a$. Notice that since the potential $\Psi(\eta, \mathbf{x})$ is not an explicit function of angle γ and $\gamma^i = \dot{x}^i$, we can write this in a more compact and suggestive form,

$$\frac{d}{d\eta}[\Theta + \Psi](\eta, \mathbf{x}, \boldsymbol{\gamma}) = \dot{\Psi} - \dot{\Phi}, \tag{2.27}$$

which also shows that in a static potential $\Theta + \Psi$ is conserved. Thus the temperature fluctuation is just given by the potential difference:

$$\Theta(\eta_0, \mathbf{x}_0, \boldsymbol{\gamma}_0) = \Theta(\eta_1, \mathbf{x}_1, \boldsymbol{\gamma}_1) + [\Psi(\eta_1, \mathbf{x}_1) - \Psi(\eta_0, \mathbf{x}_0)]. \tag{2.28}$$

This is the Sachs-Wolfe effect [138] in its simplest form.

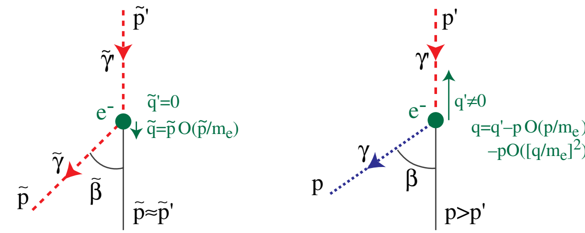
2.2 Compton Scattering

Compton scattering $\gamma(p) + e(q) \leftrightarrow \gamma(p') + e(q')$ dominates the interaction of CMB photons with electrons. By allowing energy exchange between the photons and electrons, it is the primary mechanism for the thermalization of the CMB. It also governs the mutual evolution of photon and baryon inhomogeneities before last scattering. The goal of this section is to derive its collision term in the Boltzmann equation to second order in the small energy transfer due to scattering. The approach taken here provides a coherent framework for all Compton scattering effects. In the proper limits, the equation derived below reduces to more familiar forms, *e.g.* the Kompaneets equation in the homogeneous and isotropic limit and the temperature Boltzmann equation for blackbody spectra. Furthermore, new truly second order effects such as the quadratic Doppler effect which mix spectral distortions and anisotropies result [75].

We make the following assumptions in deriving the equations:

1. The Thomson limit applies, *i.e.* the fractional energy transfer $\delta p/p \ll 1$ in the rest frame of the background radiation.
2. The radiation is unpolarized and remains so.
3. The density of electrons is low so that Pauli suppression terms may be ignored.
4. The electron distribution is thermal about some bulk flow velocity determined by the baryons $\mathbf{v}_b$.

Approximations (1), (3), and (4) are valid for most situations of cosmological interest. The approximation regarding polarization is not strictly true. Polarization is generated at the last scattering surface by Compton scattering of anisotropic radiation However, since anisotropies themselves tend to be small, polarization is only generated at the $\sim 10\%$ level compared with temperature perturbations [93]. The feedback effect into the temperature only represents a $\sim 5\%$ correction to the temperature evolution and thus is only important for high precision calculations. We will consider its effects in greater detail in Appendix §A.3.1.



(a) Electron Rest Frame (b) Background Frame

Figure 2.1: Scattering Geometry

In the electron rest frame, scattering only transfers energy to order $\mathcal{O}(\tilde{p}/m_e)$ due to the recoil of the electron. The Doppler shift into the background frame however induces a dipole which is aligned with the electron velocity. Dash length represents the photon wavelength. Aside from the energy shift due to recoil, the quadratic Doppler effect transfers energy to the photons $\delta p/p = \mathcal{O}[v_e^2 = (q/m_e)^2]$. The change in scattering angle is due to relativistic beaming effects.

2.2.1 Collision Integral

Again employing a locally orthonormal, *i.e.* Minkowski, frame we may in general express the collision term as [11]

$$\begin{aligned} C[f] = & \frac{1}{2E(p)}\int DqDq'Dp'(2\pi)^4\delta^{(4)}(p+q-p'-q')|M|^2 \\ & \times\Big\{g(t,\mathbf{x},\mathbf{q}')f(t,\mathbf{x},\mathbf{p}')\left[1+f(t,\mathbf{x},\mathbf{p})\right] \\ & -g(t,\mathbf{x},\mathbf{q})f(t,\mathbf{x},\mathbf{p})\left[1+f(t,\mathbf{x},\mathbf{p}')\right]\Big\}, \end{aligned} \tag{2.29}$$

where $|M|^2$ is the Lorentz invariant matrix element, $f(t, \mathbf{x}, \mathbf{p})$ is the photon distribution function, $g(t, \mathbf{x}, \mathbf{q})$ is the electron distribution function and

$$Dq = \frac{d^3q}{(2\pi)^3 2E(q)}, \tag{2.30}$$

is the Lorentz invariant phase space element. The terms in equation (2.29) which contain the distribution functions are just the contributions from scattering into and out of the momentum state $\mathbf{p}$ including stimulated emission effects.

We will assume that the electrons are thermally distributed about some bulk flow velocity $\mathbf{v}_b$,

$$g(t, \mathbf{x}, \mathbf{q}) = (2\pi)^3 x_e n_e (2\pi m_e T_e)^{-3/2} \exp\left\{\frac{-[\mathbf{q} - m_e\mathbf{v}_b]^2}{2m_e T_e}\right\}, \tag{2.31}$$

where x_e is the ionization fraction, n_e is the electron number density, m_e is the electron mass, and we employ units with $c = \hbar = k_B = 1$ here and throughout. Expressed in the rest frame of the electron, the matrix element for Compton scattering summed over polarization is given by [113]

$$|M|^2 = 2(4\pi)^2\alpha^2 \left[\frac{\tilde{p}'}{\tilde{p}} + \frac{\tilde{p}}{\tilde{p}'} - \sin^2\tilde{\beta}\right], \tag{2.32}$$

where the tilde denotes quantities in the rest frame of the electron, α is the fine structure constant, and $\cos\tilde{\beta} = \tilde{\boldsymbol{\gamma}} \cdot \tilde{\boldsymbol{\gamma}}'$ is the scattering angle (see Fig. 2.1). The Lorentz transformation gives

$$\frac{p}{\tilde{p}} = \frac{\sqrt{1 - q^2/m_e^2}}{1 - \mathbf{p}\cdot\mathbf{q}/pm_e}, \tag{2.33}$$

and the identity $\tilde{p}_\mu \tilde{p}'^\mu = p_\mu p'^\mu$ relates the scattering angles.

We now expand in the energy transfer $p - p'$ from scattering. There are several small quantities involved in this expansion. It is worthwhile to compare these terms. To first order, there is only the bulk velocity of the electrons v_b. In second order, many more terms appear. The quantity T_e/m_e characterizes the kinetic energy of the electrons and is to be compared with p/m_e or essentially $T/m_e \simeq 5\times 10^{-10}(1+z_*)$, where T is the temperature of the photons. Before a redshift $z_{cool} \simeq 8.0(\Omega_0 h^2)^{1/5} x_e^{-2/5}$, where x_e is the ionization fraction (this corresponds to $z \gtrsim 500(\Omega_b h^2)^{2/5}$ for standard recombination), the tight coupling between photons and electrons via Compton scattering requires these two temperatures to be comparable (see §3.2.1). At lower redshifts, it is possible that $T_e \gg T$, which produces distortions in the radiation via the Sunyaev-Zel'dovich (SZ) effect as discussed in section §3.2.1. Note that the term T_e/m_e may also be thought of as the average thermal velocity squared $\langle v_{\rm therm}^2\rangle = 3T_e/m_e$. This is to be compared with the bulk velocity squared v_b^2 and will depend on the specific means of ionization. Terms of order $(q/m_e)^2$ contain both effects.

Let us evaluate the collision integral keeping track of the order of the terms. The matrix element expressed in terms of the corresponding quantities in the frame of the radiation is

$$|M|^2 = 2(4\pi)^2\alpha^2\left(\mathcal{M}_0 + \mathcal{M}_{q/m_e} + \mathcal{M}_{(q/m_e)^2} + \mathcal{M}_{(qp/m_e^2)} + \mathcal{M}_{(p/m_e)^2}\right) + h.o., \tag{2.34}$$

where

$$\begin{aligned}
\mathcal{M}_0 &= 1 + \cos^2\beta, \\
\mathcal{M}_{q/m_e} &= -2\cos\beta(1-\cos\beta)\left[\frac{\mathbf{q}\cdot\mathbf{p}}{m_e p} + \frac{\mathbf{q}\cdot\mathbf{p}'}{m_e p'}\right], \\
\mathcal{M}_{(q/m_e)^2} &= \cos\beta(1-\cos\beta)\frac{q^2}{m_e^2}, \\
\mathcal{M}_{qp/m_e^2} &= (1-\cos\beta)(1-3\cos\beta)\left[\frac{\mathbf{q}\cdot\mathbf{p}}{m_e p} + \frac{\mathbf{q}\cdot\mathbf{p}'}{m_e p'}\right]^2 \\
&\quad + 2\cos\beta(1-\cos\beta)\frac{(\mathbf{q}\cdot\mathbf{p})(\mathbf{q}\cdot\mathbf{p}')}{m_e^2 pp'}\Bigg\}, \\
\mathcal{M}_{(p/m_e)^2} &= (1-\cos\beta)^2\frac{p^2}{m_e^2}.
\end{aligned} \tag{2.35}$$

Notice that the zeroth order term gives an angular dependence of $1 + \cos^2\beta$ which is the familiar Thomson cross section result.

Likewise, the electron energies can be expressed as

$$\frac{1}{E_q E_q'} = \frac{1}{m_e^2}[1 - \mathcal{E}_{(q/m_e)^2} - \mathcal{E}_{qp/m_e^2} - \mathcal{E}_{(p/m_e)^2}], \tag{2.36}$$

where

$$\begin{aligned}
\mathcal{E}_{(q/m_e)^2} &= \frac{q^2}{m_e^2}, \\
\mathcal{E}_{qp/m_e^2} &= \frac{(\mathbf{p}-\mathbf{p}')\cdot\mathbf{q}}{m_e^2}, \\
\mathcal{E}_{(p/m_e)^2} &= \frac{(\mathbf{p}-\mathbf{p}')^2}{2m_e^2}.
\end{aligned} \tag{2.37}$$

The following identities are very useful for the calculation. Expansion to second order in energy transfer can be handled in a quite compact way by "Taylor expanding" the delta function for energy conservation in $\delta p = q - q'$,

$$\begin{aligned}
\delta(p+q-p'-q') &= \delta(p-p') + (\mathcal{D}_{q/m_e} + \mathcal{D}_{p/m_e})p\left[\frac{\partial}{\partial p'}\delta(p-p')\right] \\
&\quad + \frac{1}{2}(\mathcal{D}_{q/m_e} + \mathcal{D}_{p/m_e})^2 p^2\left[\frac{\partial^2}{\partial p'^2}\delta(p-p')\right] + h.o.,
\end{aligned} \tag{2.38}$$

where

$$\begin{aligned}
\mathcal{D}_{q/m_e} &= \frac{1}{m_e p}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{q}, \\
\mathcal{D}_{p/m_e} &= \frac{1}{m_e p}(\mathbf{p}-\mathbf{p}')^2.
\end{aligned} \tag{2.39}$$

This is of course defined and justified by integration by parts. Integrals over the electron distribution function are trivial,

$$\begin{aligned}
\int \frac{d^3\mathbf{q}}{(2\pi)^3} g(\mathbf{q}) &= x_e n_e, \\
\int \frac{d^3\mathbf{q}}{(2\pi)^3} q^i g(\mathbf{q}) &= m_e v_b^i x_e n_e, \\
\int \frac{d^3\mathbf{q}}{(2\pi)^3} q^i q^j g(\mathbf{q}) &= m_e^2 v_b^i v_b^j x_e n_e + m_e T_e \delta^{ij} x_e n_e.
\end{aligned} \tag{2.40}$$

Thus while the terms of $\mathcal{O}(q/m_e) \to \mathcal{O}(v_b)$, the $\mathcal{O}(q^2/m_e^2)$ terms give two contributions: $\mathcal{O}(v_b^2)$ due to the bulk velocity and $\mathcal{O}(T_e/m_e)$ from the thermal velocity.

The result of integrating over the electron momenta can be written

$$C[f] = \frac{d\tau}{dt}\int dp'\frac{p'}{p}\int\frac{d\Omega'}{4\pi}\frac{3}{4}\left[\mathcal{C}_0 + \mathcal{C}_{p/m_e} + \mathcal{C}_{v_b} + \mathcal{C}_{v_b^2} + \mathcal{C}_{T_e/m_e} + \mathcal{C}_{v_b p/m_e} + \mathcal{C}_{(p/m_e)^2}\right], \tag{2.41}$$

where we have kept terms to second order in $\delta p/p$ and the optical depth to Thomson scattering τ is defined through the scattering rate

$$\frac{d\tau}{dt} \equiv x_e n_e \sigma_T, \tag{2.42}$$

with

$$\sigma_T = 8\pi\alpha^2/3m_e^2, \tag{2.43}$$

as the Thomson cross section. Equation (2.41) may be considered as the source equation for all first and second order Compton scattering effects.

2.2.2 Individual Terms

In most cases of interest, only a few of the terms in equation (2.41) will ever contribute. Let us now consider each in turn. It will be useful to define two combinations of distribution functions

$$\begin{aligned} F_1(t,\mathbf{x},\mathbf{p},\mathbf{p}') &= f(t,\mathbf{x},\mathbf{p}') - f(t,\mathbf{x},\mathbf{p}), \\ F_2(t,\mathbf{x},\mathbf{p},\mathbf{p}') &= f(t,\mathbf{x},\mathbf{p}) + 2f(t,\mathbf{x},\mathbf{p})f(t,\mathbf{x},\mathbf{p}') + f(t,\mathbf{x},\mathbf{p}'), \end{aligned} \tag{2.44}$$

which will appear in the explicit evaluation of the collision term.

(a) Anisotropy Suppression: $\mathcal{C}_0$

Scattering makes the photon distribution isotropic in the electron rest frame. Microphysically this is accomplished via scattering into and out of a given direction. Since the electron velocity is assumed to be first order in the perturbation, to zeroth order scattering makes the radiation isotropic $\delta f \equiv f - f_0 \to 0$, where f_0 is the isotropic component of the distribution function.

Its primary function then is the suppression of anisotropies as seen by the scatterers. Since isotropic perturbations are not damped, *inhomogeneities* in the distribution persist. Inhomogeneities at a distance are seen as anisotropies provided there are no intermediate scattering events, *i.e.* they are on the last scattering surface. They are the dominant source of primary anisotropies (see §6) and an important contributor to secondary anisotropies (see §7.1.3).

Explicitly the suppression term is

$$\mathcal{C}_0 = \delta(p-p')\left(1+\cos^2\beta\right)F_1(t,\mathbf{x},\mathbf{p},\mathbf{p}'). \tag{2.45}$$

Inserting this into equation (2.41) for the integration over incoming angles and noting that $\cos\beta = \boldsymbol{\gamma}\cdot\boldsymbol{\gamma}'$, we obtain the contribution

$$C_0[f] = \frac{d\tau}{dt}[(f_0 - f) + \gamma_i\gamma_j f^{ij}], \tag{2.46}$$

where f_0 is the isotropic component of the distribution and the f^{ij} are proportional to the quadrupole moments of the distribution

$$f^{ij}(t,\mathbf{x},p) = \frac{3}{4}\int\frac{d\Omega}{4\pi}(\gamma^i\gamma^j - \frac{1}{3}\delta^{ij})f. \tag{2.47}$$

The angular dependence of Compton scattering sources a quadrupole anisotropy damp more slowly than the higher moments.[1] Even so C_0 vanishes only if the distribution is isotropic $f = f_0$. Furthermore, since the zeroth order effect of scattering is to isotropize the distribution, in most cases any anisotropy is at most first order in the perturbative expansion. This enormously simplifies the form of the other terms.

(b) Linear and Quadratic Doppler Effect: $\mathcal{C}_{v_b}$ and $\mathcal{C}_{v_b^2}$

Aside from the small electron recoil (see c), the kinematics of Thomson scattering require that no energy be transferred in the rest frame of the electron *i.e.* $\tilde{p}' = \tilde{p}$. Nevertheless, the transformation from and back into the background frame induces a Doppler shift,

$$\frac{\delta p}{p} = \frac{1-\mathbf{v}_b\cdot\boldsymbol{\gamma}'}{1-\mathbf{v}_b\cdot\boldsymbol{\gamma}} - 1 = \mathbf{v}_b\cdot(\boldsymbol{\gamma}-\boldsymbol{\gamma}') + (\mathbf{v}_b\cdot\boldsymbol{\gamma})\mathbf{v}_b\cdot(\boldsymbol{\gamma}-\boldsymbol{\gamma}') + \mathcal{O}(v_b^3). \tag{2.48}$$

Notice that in addition to the usual linear term, there is also a term quadratic in v_b. Furthermore, quadratic contributions do not disappear upon averaging over incoming and outgoing directions. They represent a net energy gain and/or loss by the CMB.

Let us first consider the case that scattering is rapid, *e.g.* before recombination, such that all CMB photons scatter before traversing a coherence scale of the velocity field. After averaging over incoming directions the net first order contribution is $\delta p/p = \boldsymbol{\gamma}\cdot\mathbf{v}_b$. As one might expect, this is just the Doppler shifted signal we expect from radiation that is isotropic in the electron rest frame. The spectrum therefore is a blackbody with a dipole signature $\mathbf{v}_\gamma$ in angle: $\delta T/T = \boldsymbol{\gamma}\cdot\mathbf{v}_\gamma = \boldsymbol{\gamma}\cdot\mathbf{v}_b$. To $\mathcal{O}(v_b^2)$, there is a net energy transfer. Scattering brings the photons into kinetic equilibrium with the electrons. This equalization amounts to an energy gain by the photons if $v_\gamma < v_b$, and a loss in the opposite case. The energy transfer occurs only until kinetic equilibrium is attained. In other words, once the photons are isotropic in the electron rest frame $\mathbf{v}_\gamma = \mathbf{v}_b$, scattering has no further effect.

On the other hand, if the mean free path of the photons due to Compton scattering is much greater than the typical coherence scale of the velocity, the photons are in the diffusion limit. This can occur in reionized scenarios. Scattering is not rapid enough to ever make the distribution isotropic in the local rest frame of the electrons. Say some fraction $d\tau = n_e\sigma_T dt$ of the CMB scatters within a coherence scale. Then the Doppler shift will be reduced to $\boldsymbol{\gamma}\cdot\mathbf{v}_b d\tau$ and the energy transfer will be of order $\mathcal{O}(v_b^2 d\tau)$. As the photons continue to scatter, the first order Doppler term vanishes since redshifts and blueshifts from regions with different orientations of the electron velocity will mainly cancel (see §7.1.4). The second order term will however be positive definite: $\mathcal{O}(\int v_b^2 d\tau)$.

[1] This can generate viscosity in the photon-baryon fluid and affects diffusion damping of anisotropies as we show in Appendix A.3.1.

Is the resultant spectrum also a blackbody? In averaging over angles and space above, we have really superimposed many Doppler shifts for individual scattering events. Therefore the resulting spectrum is a superposition of blackbodies with a range of temperatures $\Delta T/T = \mathcal{O}(v_b)$. Zel'dovich, Illarionov, & Sunyaev [182] have shown that this sort of superposition leads to spectral distortions of the Compton-y type with $y = \mathcal{O}(v_b^2)$ (see §3.2.1).

Now let us write down the explicit form of these effects. The linear term is given by

$$\begin{aligned}\mathcal{C}_{v_b} = \Big\{ & \left[\frac{\partial}{\partial p'}\delta(p-p')\right]\left(1+\cos^2\beta\right)\mathbf{v}_b\cdot(\mathbf{p}-\mathbf{p}') \\ & -\delta(p-p')2\cos\beta(1-\cos\beta)\left[\frac{\mathbf{v}_b\cdot\mathbf{p}}{p}+\frac{\mathbf{v}_b\cdot\mathbf{p}'}{p'}\right]\Big\}F_1(t,\mathbf{x},\mathbf{p},\mathbf{p}'). \qquad (2.49)\end{aligned}$$

Assuming that the anisotropy is at most first order in the perturbation $\delta f \equiv f - f_0 \lesssim \mathcal{O}(v_b)$, the contribution to the collision term can be explicitly evaluated as

$$C_{v_b}[f] = -\frac{d\tau}{dt}\left[(\boldsymbol{\gamma}\cdot\mathbf{v}_b)p\frac{\partial f_0}{\partial p} - \mathcal{O}(\delta f\, v_b)\right]. \qquad (2.50)$$

The $\mathcal{O}(\delta f\, v_b)$ term is not necessarily small compared with other second order terms. However, we already know its effect. If scattering is sufficiently rapid, the anisotropy δf will be a dipole corresponding to the electron velocity $\mathbf{v}_b$. In this case, its effects will cancel the $\mathcal{O}(v_b^2)$ quadratic term. Notice that to first order equilibrium will be reached between the zeroth and first order terms when

$$f_0 - f - p(\boldsymbol{\gamma}\cdot\mathbf{v}_b)\frac{\partial f}{\partial p} = \mathcal{O}(v_b^2), \qquad (2.51)$$

assuming negligible quadrupole. For a blackbody, $T(\partial f/\partial T) = -p(\partial f/\partial p)$. Thus the equilibrium configuration represents a temperature shift $\delta T/T = \boldsymbol{\gamma}\cdot\mathbf{v}_b$. This formally shows that the $\mathcal{O}(v_b)$ term makes the photons isotropic in the baryon rest frame.

The quadratic term, given explicitly by

$$\begin{aligned}\mathcal{C}_{v_b^2} = & \frac{1}{2}\left[\frac{\partial^2}{\partial p'^2}\delta(p-p')\right]\left(1+\cos^2\beta\right)[\mathbf{v}_b\cdot(\mathbf{p}-\mathbf{p}')]^2F_1(t,\mathbf{x},\mathbf{p},\mathbf{p}') \\ & -\left[\frac{\partial}{\partial p'}\delta(p-p')\right]2\cos\beta(1-\cos\beta)\left[\frac{\mathbf{v}_b\cdot\mathbf{p}}{p}+\frac{\mathbf{v}_b\cdot\mathbf{p}'}{p'}\right]\mathbf{v}_b\cdot(\mathbf{p}-\mathbf{p}')F_1(t,\mathbf{x},\mathbf{p},\mathbf{p}') \\ & +\delta(p-p')\Big\{-(1-2\cos\beta+3\cos^2\beta)v_b^2+2\cos\beta(1-\cos\beta)\frac{(\mathbf{v}_b\cdot\mathbf{p})(\mathbf{v}_b\cdot\mathbf{p}')}{pp'} \\ & +(1-\cos\beta)(1-3\cos\beta)\left[\frac{\mathbf{v}_b\cdot\mathbf{p}}{p}+\frac{\mathbf{v}_b\cdot\mathbf{p}'}{p'}\right]^2\Big\}F_1(t,\mathbf{x},\mathbf{p},\mathbf{p}'), \qquad (2.52)\end{aligned}$$

can also be evaluated under the assumption of small anisotropy,

$$C_{v_b^2}[f] = \frac{d\tau}{dt}\left\{\left[(\boldsymbol{\gamma}\cdot\mathbf{v}_b)^2+v_b^2\right]p\frac{\partial f}{\partial p}+\left[\frac{11}{20}(\boldsymbol{\gamma}\cdot\mathbf{v}_b)^2+\frac{3}{20}v_b^2\right]p^2\frac{\partial^2 f}{\partial p^2}\right\}. \qquad (2.53)$$

(c) Thermal Doppler Effect and Recoil: C_{T_e/m_e} and C_{p/m_e}

Of course, we have artificially separated out the bulk and thermal components of the electron velocity. The thermal velocity leads to a quadratic Doppler effect exactly as described above if we make the replacement $\langle v_b^2\rangle \to \langle v_{\rm therm}^2\rangle = 3T_e/m_e$. For an isotropic distribution of photons, this leads to the familiar Sunyaev-Zel'dovich (SZ) effect [162]. The SZ effect can therefore be understood as the second order spectral distortion and energy transfer due to the superposition of Doppler shifts from individual scattering events off electrons in thermal motion. It can also be naturally interpreted macrophysically: hot electrons transfer energy to the photons via Compton scattering. Since the number of photons is conserved in the scattering, spectral distortions must result. Low energy photons are shifted upward in frequency, leading to the Rayleigh-Jeans depletion and the Wien tail enhancement characteristic of Compton-y distortions. We will consider this process in more detail in §3.2.1.

If the photons have energies comparable to the electrons (*i.e.* the electron and photon temperatures are nearly equal), there is also a significant correction due to the recoil of the electron. The scattering kinematics tell us that

$$\frac{\tilde{p}'}{\tilde{p}} = \left[1+\frac{\tilde{p}}{m_e}(1-\cos\tilde{\beta})\right]^{-1}. \qquad (2.54)$$

Thus to lowest order, the recoil effects are $\mathcal{O}(p/m_e)$. Together with the thermal Doppler effect, these terms form the familiar Kompaneets equation in the limit where the radiation is isotropic and drive the photons toward kinetic equilibrium as a Bose-Einstein distribution of temperature T_e (see §3.2.2). A blackbody distribution cannot generally be established since Compton scattering requires conservation of the photon number.

Explicitly, the recoil term

$$\mathcal{C}_{p/m_e} = -\left[\frac{\partial}{\partial p'}\delta(p-p')\right]\left(1+\cos^2\beta\right)\frac{(\mathbf{p}-\mathbf{p}')^2}{2m_e}F_2(t,\mathbf{x},\mathbf{p},\mathbf{p}'), \qquad (2.55)$$

yields

$$C_{T_e/m_e}[f] = \frac{d\tau}{dt}\frac{p}{m_e}\left[4f(1+f)+(1+2f)f\frac{\partial f}{\partial p}\right]; \qquad (2.56)$$

whereas the thermal term

$$\begin{aligned}\mathcal{C}_{T_e/m_e} = \Big\{ & \left[\frac{\partial^2}{\partial p'^2}\delta(p-p')\right]\left(1+\cos^2\beta\right)\frac{(\mathbf{p}-\mathbf{p}')^2}{2} - \left[\frac{\partial}{\partial p'}\delta(p-p')\right]2\cos\beta \\ & \times(1-\cos^2\beta)(p-p')\delta(p-p')[4\cos^3\beta-9\cos^2\beta-1]\Big\}\frac{T_e}{m_e}F_1(t,\mathbf{x},\mathbf{p},\mathbf{p}'), \qquad (2.57)\end{aligned}$$

gives

$$C_{T_e/m_e}[f] = \frac{d\tau}{dt}\frac{T_e}{m_e}\left(4p\frac{\partial f}{\partial p}+p^2\frac{\partial^2 f}{\partial p^2}\right). \qquad (2.58)$$

(d) Higher Order Recoil Effects: $C_{v_b p/m_e}$ and $C_{(p/m_e)^2}$

These terms represent the next order in corrections due to the recoil effect. Explicit forms are provided in [75]. In almost all cases, they are entirely negligible. Specifically, for most cosmological models, the baryon bulk flow grows by gravitational instability and is small until relatively recently. On the other hand the photon energy redshifts with the expansion and is more important early on. Thus their cross term is never important for cosmology. Furthermore, since there is no cancellation in the $C_{p/m}$ term, $C_{(p/m)^2}$ will never produce the dominant effect. We will hereafter drop these terms in our consideration.

2.2.3 Generalized Kompaneets Equation

Even for an initially anisotropic radiation field, multiple scattering off electrons will have the zeroth order effect of erasing the anisotropy. Therefore when the optical depth is high, we can approximate the radiation field as nearly isotropic. Under the assumption of full isotropy, the individual effects from equations (2.50), (2.53), (2.56) and (2.58) combine to form the collision term

$$C[f] = \frac{d\tau}{dt}\Bigg\{ -\boldsymbol{\gamma}\cdot\mathbf{v}_b p\frac{\partial f}{\partial p} + \Bigg(\left[(\boldsymbol{\gamma}\cdot\mathbf{v}_b)^2 + v_b^2\right] p\frac{\partial f}{\partial p} + \left[\frac{3}{20}v_b^2 + \frac{11}{20}(\boldsymbol{\gamma}\cdot\mathbf{v}_b)^2\right] \tag{2.59}$$

$$\times\, p^2\frac{\partial^2 f}{\partial p^2}\Bigg) + \frac{1}{m_e p^2}\frac{\partial}{\partial p}\left[p^4\left\{T_e\frac{\partial f}{\partial p} + f(1+f)\right\}\right]\Bigg\}. \tag{2.60}$$

The first and second terms represent the linear and quadratic Doppler effects respectively. The final term is the usual Kompaneets equation. Notice that in the limit of many scattering regions, we can average over the direction of the electron velocity. The first order linear Doppler effect primarily cancels in this case. We can then reduce equation (2.59) to

$$C[f] = \frac{d\tau}{dt}\left\{\frac{\langle v_b^2\rangle}{3}\frac{1}{p^2}\frac{\partial}{\partial p}\left[p^4\frac{\partial f}{\partial p}\right] + \frac{1}{m_e p^2}\frac{\partial}{\partial p}\left[p^4\left\{T_e\frac{\partial f}{\partial p} + f(1+f)\right\}\right]\right\}. \tag{2.61}$$

Under the replacement $\langle v_{\rm therm}^2\rangle = 3T_e/m \to v_b^2$, the SZ (thermal Doppler) portion of the Kompaneets equation and quadratic Doppler equation have the same form. Thus, spectral distortions due to bulk flow have exactly the same form as SZ distortions and can be characterized by the Compton-y parameter (see §3.2.1) given in its full form by

$$y = \int \frac{d\tau}{dt}\left[\frac{1}{3}\langle v_b^2(t)\rangle + \frac{T_e - T}{m_e}\right]dt \qquad v_b \gg v_\gamma. \tag{2.62}$$

The appearance of the photon temperature T in equation (2.62) is due to the recoil terms in the Kompaneets equation.

The quadratic Doppler effect only contributes when the electron velocity is much greater than the photon dipole or bulk velocity. Just as the thermal term vanishes when the temperatures are equal, the "kinetic" part vanishes if the bulk velocities are equal. The effect therefore contributes only in the diffusion limit where the photons can be approximated

a weakly anisotropic distribution diffusing through independently moving baryons. However above redshift $z_d \simeq 160(\Omega_0 h^2)^{1/5} x_e^{-2/5}$ (see §5.3.1), Compton drag on the electrons keeps the electrons coupled to the photons and requires $v_b \sim v_\gamma$. For a fully ionized, *COBE* normalized CDM model, integrating (2.62) up until the drag epoch yields a quadratic Doppler contribution of CDM equal to $y(z_d) \simeq 5\times 10^{-7}$, almost two orders of magnitude below the current limits. Almost certainly the thermal effect in clusters will completely mask this effect. We will henceforth ignore its contributions when discussing spectral distortions.

2.2.4 Collisional Brightness Equation

We have shown that if the photons and baryons are in equilibrium, the effects which create spectral distortions vanish. In this case, we may integrate over the spectrum to form the temperature perturbation. Combining the collisional zeroth and $\mathcal{O}(v_b)$ parts, equations (2.46) and (2.50) respectively, with equation (2.26) for the collisionless part, we obtain for the temperature perturbation evolution in conformal time $\Theta(\eta, \mathbf{x}, \boldsymbol{\gamma})$

$$\dot\Theta + \gamma^i\frac{\partial}{\partial x^i}(\Theta + \Psi) + \dot\gamma^i\frac{\partial}{\partial\gamma^i}\Theta + \dot\Phi = \dot\tau(\Theta_0 - \Theta - \gamma_i v_b^i + \frac{1}{16}\gamma_i\gamma_j\Pi_\gamma^{ij}), \tag{2.63}$$

where

$$\begin{aligned}\Pi_\gamma^{ij} &= \frac{4}{\pi^2\rho_\gamma}\int p^3 dp f^{ij}(\eta,\mathbf{x}) \\ &= \frac{1}{\pi^2\rho_\gamma}\int p^3 dp \int\frac{d\Omega}{4\pi}(3\gamma^i\gamma^j - \delta^{ij})f(\eta,\mathbf{x},\boldsymbol{\gamma}) \\ &= \int\frac{d\Omega}{4\pi}(3\gamma^i\gamma^j - \delta^{ij})4\Theta(\eta,\mathbf{x},\boldsymbol{\gamma}).\end{aligned} \tag{2.64}$$

The quantities Π_γ^{ij} are the quadrupole moments of the energy distribution. Since the pressure $p_\gamma = \frac{1}{3}\rho_\gamma$, they are related to the anisotropic stress. To generalize this relation to open universes, merely replace the flat space metric δ^{ij} with γ^{ij}. Equation (2.63) is the fundamental equation for primary anisotropy formation (see §6). We will revisit second order effects in §7 when we discuss reionized scenarios.