1 Recombination - Overview

This lecture will cover the subject of cosmological recombination. The major topics will be:

- Expansion history.
- Hydrogen recombination.
- Helium and lithium recombination.
- Compton heating and the gas temperature.

2 Expansion history

During our studies of the radiation-dominated era after e^+e^- annihilation, we learned that the photon temperature T related to cosmic time t via:

$$T = 1.56 g_{\star,eff}^{-1/4} \sqrt{\frac{1\,\mathrm{s}}{t}}\,\mathrm{MeV}. \tag{1}$$

Since $g_{\star,eff} = 3.36$ we have

$$T = 1.15 \sqrt{\frac{1\,\mathrm{s}}{t}}\,\mathrm{MeV}, \tag{2}$$

so the temperature drops to:

- 100 keV at 2 minutes
- 10 keV at 4 hours
- 1 keV at 2 weeks
- 100 eV at 4 yr
- 10 eV at 400 yr
- 1 eV at 40 kyr.

The redshift is related to temperature by

$$z = \frac{T}{T_{\gamma 0}} - 1, \tag{3}$$

where $T_{\gamma 0} = 2.73$ K= 0.235 milli eV is the CMB temperature today. So e.g. the redshift at BBN ($T = 100$ keV) is 4×10^8.

Matter-radiation equality. However we've only considered the radiation energy density (photons+neutrinos) so far in the Friedmann equations, whereas matter (baryons and dark matter) is also present. The energy density of radiation is $\rho_r \propto a^{-4}$ whereas for the matter $\rho_m \propto a^{-3}$. At some point the matter and radiation will have the same density, $\rho_m = \rho_r$. This epoch is called *matter-radiation equality*. We'll need to calculate it next since it determines the expansion dynamics, and will also be important for CMB anisotropies.

The ratio of matter to radiation energy density is proportional to scale factor:

$$\frac{\rho_m}{\rho_r} \propto \frac{a^{-3}}{a^{-4}} = a, \tag{4}$$

so we can find the epoch of equality if we know the matter and radiation densities today:

$$a_{\text{eq}} = \frac{\rho_{r0}}{\rho_{m0}}. \tag{5}$$

The radiation energy density today is that of the CMB and the neutrinos:

$$\rho_\gamma = \frac{\pi^2}{30} g_{\star,eff} T_{\gamma 0}^4 = 7.8 \times 10^{-34}\,\text{g/cm}^3. \tag{6}$$

A fraction 2/3.36 (60%) of this is in photons and 1.36/3.36 (40%) in neutrinos. (I haven't included starlight, which wasn't around at matter-radiation equality so we don't consider it.)

The matter density is much harder to measure, but we can parameterize it by the number $\Omega_m h^2$, where h is the Hubble constant in units of 100 km/s/Mpc. (We'll discuss ways to measure this later.) The matter density today is

$$\rho_{m0} = \frac{3\Omega_m H_0^2}{8\pi G} = 1.879 \times 10^{-29} \Omega_m h^2\,\text{g/cm}^3. \tag{7}$$

Remark: $\Omega_m h^2$ includes all nonrelativistic matter, both baryons and dark matter: $\Omega_m h^2 = \Omega_b h^2 + \Omega_{dm} h^2$. From WMAP: $\Omega_m h^2 = 0.128 \pm 0.008$.

Taking the ratio gives us the scale factor at equality,

$$a_{\text{eq}} = 4.15 \times 10^{-5} (\Omega_m h^2)^{-1}. \tag{8}$$

This can be related to a redshift using $1 + z = a^{-1}$:

$$1 + z_{\text{eq}} = 2.4 \times 10^4 \Omega_m h^2, \tag{9}$$

or a photon temperature:

$$T_{\gamma,\text{eq}} = 6.6 \times 10^4 \Omega_m h^2\,\text{K} = 5.7 \Omega_m h^2\,\text{eV}. \tag{10}$$

This is of the same order of magnitude as atomic physics energies is an accident – so far as we know!

Detailed expansion history. We'll need to understand how the expansion of the Universe changes during the matter-radiation equality. The Hubble rate is given by the Friedmann equation; neglecting curvature (which we can do early in the Universe's history), we find:

$$\begin{aligned} H^2 &= \frac{8}{3}\pi G(\rho_m + \rho_r) \\ &= \frac{8}{3}\pi G \rho_m \left(1 + \frac{1}{y}\right), \end{aligned} \tag{11}$$

where $y = a/a_{\rm eq} = \rho_m/\rho_r$; this is then:

$$H^2 = \frac{8}{3}\pi G\rho_{m0}(a_{\rm eq}y)^{-3}\left(1+\frac{1}{y}\right) = \Omega_m H_0^2 a_{\rm eq}^{-3} y^{-3}\left(1+\frac{1}{y}\right). \tag{12}$$

The time can then be computed (homework exercise):

$$t = \frac{2.6\,{\rm kyr}}{(\Omega_m h^2)^2}\left[\frac{2}{3}(1+y)^{3/2} - 2(1+y)^{1/2} + \frac{4}{3}\right]. \tag{13}$$

Note that at late times, $y \gg 1$, we have $t \propto y^{3/2} \propto a^{3/2}$, or $a \propto t^{2/3}$, as appropriate for a matter dominated universe.

3 Hydrogen recombination – equilibrium theory

Now we consider the *recombination* process – how the ionized plasma of protons and electrons turns into hydrogen atoms. (We'll discuss what happens to He and Li later.) This will be important to us because the CMB photons can Thomson scatter off of electrons, but not atoms, so recombination is accompanied by a transition from an opaque to a transparent universe.

Saha equation. Just as for BBN, we'll begin our discussion by assuming thermal equilibrium, and then come back and revisit the more detailed non-equilibrium physics.

The formation of hydrogen atoms occurs via the reaction:

$$p^+ + e^- \leftrightarrow {\rm H}(1s) + {\rm photons}. \tag{14}$$

I've labeled the ground state $1s$ of the hydrogen atom because we'll consider excited states later. Since the photons have zero chemical potential (a blackbody!) we have, in equilibrium,

$$\mu(p+) + \mu(e^-) = \mu({\rm H}, 1s). \tag{15}$$

Recall that the chemical potential is:

$$\mu_X = m_X + T\ln\left[\frac{n_X}{g_X}\left(\frac{2\pi}{m_X T}\right)^{3/2}\right]. \tag{16}$$

The degeneracy of the proton is 2 (2 spin states), the electron is 2 (same reason), and the hydrogen atom is 4 (2×2). Equating the chemical potentials gives us:

$$m_p + T\ln\frac{n_p}{2} + \frac{3}{2}T\frac{2\pi}{m_p T} + m_e + T\ln\frac{n_e}{2} + \frac{3}{2}T\frac{2\pi}{m_e T} = m_{\rm H} + T\ln\frac{n({\rm H},1s)}{4} + \frac{3}{2}T\frac{2\pi}{m_{\rm H}T}. \tag{17}$$

The mass difference between the hydrogen atom and the proton+electron is called the *ionization energy* ϵ_0:

$$\epsilon_0 = m_p + m_e - m_{\rm H}, \tag{18}$$

and for hydrogen it is 13.6 eV or 1.58×10^5 K.

Collecting the above terms, we find

$$\frac{n_p n_e}{n({\rm H}, 1s)} = \left(\frac{m_p m_e T}{2\pi m_{\rm H}}\right)^{3/2} e^{-\epsilon_0/T}. \tag{19}$$

The ratio of the proton and hydrogen atom mass is essentially 1, so

$$\frac{n_p n_e}{n({\rm H}, 1s)} = \left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\epsilon_0/T}. \tag{20}$$

This is called the *Saha equation.* It is also applicable to stellar interiors.

Application to cosmology. It is conventional in recombination studies to define the *ionization fraction* x_e, which is the number of free electrons per hydrogen nucleus:

$$x_e = \frac{n_e}{n_{\rm H,tot}}, \tag{21}$$

where in the denominator both neutral and ionized H are counted: $n_{\rm H,tot} = n_p + n({\rm H}, 1s)$. The abundance of p^+, e^-, and neutral H atoms (assumed all in $1s$) is given by:

$$n_e = n_p = x_e n_{\rm H,tot} \quad \text{and} \quad n({\rm H}, 1s) = (1 - x_e) n_{\rm H,tot}. \tag{22}$$

Since BBN produced a fraction $X_{\rm H} = 0.76$ of the initial mass in hydrogen, and it is almost all ^{1}H, we can write:

$$n_{\rm H,tot} = 0.76 n_b = 8.6 \times 10^{-6} \Omega_b h^2 a^{-3}\,{\rm cm}^{-3} = 4.2 \times 10^5 \Omega_b h^2 T_4^3\,{\rm cm}^{-3}, \tag{23}$$

where T_4 is the temperature in units of 10^4 K.

Substituting this into the Saha equation gives:

$$n_{\rm H,tot} \frac{x_e^2}{1 - x_e} = \left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\epsilon_0/T}; \tag{24}$$

or, moving $n_{\rm H,tot}$ to the right and plugging in numbers:

$$\frac{x_e^2}{1 - x_e} = \frac{5.8 \times 10^{15}}{\Omega_b h^2 T_4^{3/2}} e^{-15.8/T_4}. \tag{25}$$

With this equation and the WMAP value for $\Omega_b h^2$ we can track the equilibrium recombination history:

- Half of the hydrogen recombines ($x_e = 0.5$) by $T_4 = 0.374$ or 3740 K; $z = 1370$.

- 90% of the hydrogen recombines ($x_e = 0.1$) by $T_4 = 0.342$ or 3420 K; $z = 1250$.
- 99% of the hydrogen recombines ($x_e = 10^{-2}$) by $T_4 = 0.310$ or 3100 K; $z = 1140$.

But this calculation is no guarantee that thermal equilibrium is achieved, and indeed we'll see that it is not.

4 Hydrogen recombination – the Peebles model.

In the equilibrium calculation we showed that hydrogen recombines very quickly and in the redshift range $z \sim 1370$. But this was based on thermal equilibrium and in order to see if this is what happens we need to do reaction kinetics. The basic treatment of hydrogen recombination that we'll follow here was by Peebles (1968), and is often known as the "Peebles" model or the *three-level atom* approximation.

The setup. The Peebles model considers three types of hydrogen:

- Hydrogen in the ground state, $1s$ (a fraction x_1 of the total hydrogen).
- Hydrogen in an excited state, most likely $2s$ or $2p$. The excited levels of hydrogen are assumed to be in thermal equilibrium with each other,

$$x_{nl} \propto g_{nl} e^{-E_{nl}/T}, \tag{26}$$

  since radiative excitation and decay are very fast. Note that this proportionality does not apply to the ground state $1s$ (which is special!). We'll parameterize this by considering the total number of hydrogen atoms x_2 in the excited states.
- Ionized hydrogen (fraction $x_p = x_e$ of total hydrogen).

Peebles considers the following processes:

- Radiative recombination to the ground state, and its inverse, photoionization:

$$p^+ + e^- \leftrightarrow \mathrm{H}(1s) + \gamma. \tag{27}$$

- Radiative recombination of a hydrogen atom to an excited state, and the inverse photoionization:

$$p^+ + e^- \leftrightarrow \mathrm{H}(nl, n \geq 2) + \gamma. \tag{28}$$

- Decay of an excited atom (probably in $2p$) to the ground state, by emission of a *Lyman-α* photon at a wavelength of 1216Å, and the inverse absorption process:

$$\mathrm{H}(2p) \leftrightarrow \mathrm{H}(1s) + \gamma(1216\text{Å}). \tag{29}$$

- Decay of an atom in the $2s$ state to $1s$ by simultaneous emission of two photons, and the inverse two-photon absorption process:

$$\mathrm{H}(2s) \leftrightarrow \mathrm{H}(1s) + \gamma + \gamma. \tag{30}$$

Direct recombination to the ground state. The first reaction (#4) has little effect because the photon's energy is always greater than $\epsilon_0 = 13.6$ eV. The cross section for absorbing photons just above 13.6 eV is

$$\sigma_{pi} = 6 \times 10^{-18}\,\mathrm{cm}^2, \tag{31}$$

whereas at redshift 1300 the density of hydrogen nuclei is 400 cm^{-3}. Thus the mean free path for one of these photons is

$$L_{\mathrm{mfp}} = \frac{1}{n_{1s}\sigma_{pi}} = \frac{1}{n_{\mathrm{H}} x_1 \sigma_{pi}} = 4 \times 10^{14} x_1^{-1}\,\mathrm{cm}, \tag{32}$$

and the photon is re-absorbed in time

$$\frac{L_{\mathrm{mfp}}}{c} \sim 10^4 x_1^{-1}\,\mathrm{s}. \tag{33}$$

If the neutral fraction x_1 exceeds 10^{-9} then the direct recombination photons are absorbed in much less than a Hubble time, and hydrogen cannot recombine this way. Hereafter we'll leave this process out.

Recombination to the excited states. Now let's consider the second reaction. The rate of production of hydrogen atoms in excited states is given by $\alpha n_e n_p$, where

$$\alpha = \sum_{n=2}^{\infty} \sum_{l=0}^{n-1} \langle \sigma[p^+ + e^- \to \mathrm{H}(nl) + \gamma] v \rangle \tag{34}$$

and the average is over the thermal velocity distribution. Note that we've left out the ground state ($n = 1$). This number is called the *Case B recombination coefficient* (the name comes from ISM research, in which "Case A" includes the ground state). As long as the excited states of hydrogen get de-populated (we'll see how soon), the photon produced here is free to escape.

Once an atom is in an excited state it rapidly decays to $n = 2$. (The decays to $n = 1$ are slow, as we'll find out soon.) Thus:

$$\dot{x}_2 = \alpha n_{\mathrm{H,tot}} x_e^2 - \beta x_2, \tag{35}$$

where β_2 is the thermal photoionization rate from the excited levels. (Technically a Boltzmann average.) We can calculate it from the principle of detailed balance: in Saha equilibrium, we would have:

$$x_2 = 4 n_{\mathrm{H,tot}} x_e^2 \left(\frac{2\pi}{m_e T}\right)^{3/2} e^{\epsilon_0/4T}, \tag{36}$$

where $\epsilon_0/4$ is the binding energy of the second level of hydrogen, and the 4 comes from the fact that the $n = 2$ level has 16 states (4 electron orbital $\times$ 2 electron spin $\times$ 2 proton spin) versus 2 each for the electron and proton (16/2/2=4). This means that:

$$\beta = \frac{\alpha}{4}\left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\epsilon_0/4T}. \tag{37}$$

Thus the rate of production of excited hydrogen atoms from process #4 is:

$$\dot{x}_2 = \alpha\left[n_{\rm H,tot}x_e^2 - \frac{1}{4}\left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\epsilon_0/4T} x_2\right]. \tag{38}$$

There's a corresponding rate of loss of free electrons, $\dot{x}_e = -\dot{x}_2$.

The Case B recombination coefficient can be approximated by

$$\alpha = \frac{4.309\times 10^{-13} T_4^{-0.6166}}{1 + 0.6703 T_4^{0.53}}\ {\rm cm^3\,s^{-1}}. \tag{39}$$

(Péquignot et al 1991 A&A 251,680).

Lyman-α decay. We next come to the decay of hydrogen atoms in $2p$ by Lyman-α emission. The most obvious approximation to make is that we can write a simple decay rate,

$$\dot{x}_2 = -\frac{3}{4}A_{2p}x_2, \tag{40}$$

where the factor of 3/4 is the fraction of the hydrogen atoms in $n = 2$ in the p orbital (statistical), and $A_{2p} = 6.2 \times 10^8\ {\rm s}^{-1}$ is the $2p$ decay rate. But this is wrong because the Lyman-α photons are quickyl re-absorbed. To see this, let's calculate the optical depth of Lyman-α photons. This is:

$$\tau = \int n({\rm H}, 1s)\sigma\, dt = n_{\rm H,tot}x_1 \int \frac{\sigma(\omega)}{|\dot{\omega}|} d\omega, \tag{41}$$

where we've written the cross section σ for Lyman-α absorption as a function of photon (angular) frequency ω. The optical depth is finite because $\dot{\omega} \neq 0$, i.e. the photon eventually redshifts out of the line. The rate of change of the frequency is $\dot{\omega} = -H\omega$, and the Lyman-α line is narrow enough that we will approximate the cross section as a δ-function:

$$\sigma(\omega) = S\delta(\omega - \omega_{{\rm Ly}\alpha}). \tag{42}$$

Then

$$\tau = \frac{n_{\rm H,tot}x_1}{H\omega_{{\rm Ly}\alpha}} S. \tag{43}$$

The integral of the cross section for absorption of Lyman-α radiation can be obtained from the principle of detailed balance. If we put an H atom in a blackbody, then in equilibrium the rate of decay from $2p \to 1s$ must be compensated by the rate of absorption $1s \to 2p$. That is,

$$A_{2p}[1 + f(\omega_{{\rm Ly}\alpha})]n({\rm H}, 2p) = n({\rm H}, 1s)\int_0^\infty \frac{dn_\gamma}{d\omega}\sigma(\omega)\, d\omega, \tag{44}$$

where $dn_\gamma/d\omega$ is the number of photons per unit volume per unit frequency. Putting in the values:

$$f(\omega_{\rm Ly\alpha}) = \frac{1}{e^{\omega_{\rm Ly\alpha}/T}-1} \tag{45}$$

and

$$\frac{dn_\gamma}{d\omega} = \frac{\omega^2}{\pi^2(e^{\omega/T}-1)}, \tag{46}$$

we find:

$$A_{2p}\frac{e^{\omega_{\rm Ly\alpha}/T}}{e^{\omega_{\rm Ly\alpha}/T}-1}n({\rm H},2p) = n({\rm H},1s)\frac{S\omega_{\rm Ly\alpha}^2}{\pi^2(e^{\omega_{\rm Ly\alpha}/T}-1)}. \tag{47}$$

Solving for S, we find:

$$S = \pi^2 A_{2p}\omega_{\rm Ly\alpha}^{-2}e^{\omega_{\rm Ly\alpha}/T}\frac{n({\rm H},2p)}{n({\rm H},1s)}. \tag{48}$$

In thermal equilibrium, the last factor takes on its Boltzmann ratio $3e^{-\omega_{\rm Ly\alpha}/T}$, so

$$S = 3\pi^2 A_{2p}\omega_{\rm Ly\alpha}^{-2}. \tag{49}$$

The optical depth is then

$$\tau = \frac{3\pi^2 A_{2p}n_{\rm H,tot}x_1}{H\omega_{\rm Ly\alpha}^3}. \tag{50}$$

This is called the *Sobolev optical depth* and plays a key role in line foormation in expanding media (the Universe; stellar winds; supernova remnants). It is typically a few$\times 10^8$ during recombination, so Lyman-α photons are immediately re-absorbed.

We can determine the probability P of re-absorption of the Lyman-α photon by reparameterizing the line to run from 0 optical depth (at the red side of the line) to τ (the blue side). If the emission frequency distribution equals that of absorption (detailed balance!) then a Lyman-α photon is equally likely to be emitted at any optical depth value τ' between 0 and τ. Thus the escape probability is:

$$P = \langle e^{-\tau'}\rangle = \frac{1}{\tau}\int_0^\tau e^{-\tau'}\,d\tau' = \frac{1-e^{-\tau}}{\tau} \approx \frac{1}{\tau}, \tag{51}$$

where the last approximation holds for an optically thick line. To get net Lyman-α emission, we should be multiplying our original Eq. (40) by P:

$$\dot{x}_2 = -\frac{3}{4}A_{2p}x_2P = -\frac{3A_{2p}}{4\tau}x_2 = -\frac{H\omega_{\rm Ly\alpha}^3}{4\pi^2 n_{\rm H,tot}x_1}x_2. \tag{52}$$

To this rate we need to include a detailed balance correction for thermal excitation by blackbody photons that redshift into Lyman-α:

$$\dot{x}_2 = -\frac{H\omega_{\rm Ly\alpha}^3}{4\pi^2 n_{\rm H,tot}x_1}\left(x_2 - 4x_1e^{-\omega_{\rm Ly\alpha}/T}\right), \tag{53}$$

where we've inserted a 4 in front of the last term because of the ratio of number of states 16/4=4. Simplifying gives:

$$\dot{x}_2 = -\frac{H\omega_{\mathrm{Ly}\alpha}^3}{\pi^2 n_{\mathrm{H,tot}}}\left(\frac{x_2}{4x_1} - e^{-\omega_{\mathrm{Ly}\alpha}/T}\right). \tag{54}$$

Of course there's a corresponding contribution to the ground state abundance, $\dot{x}_1 = -\dot{x}_2$.

Two-photon decay. The Lyman-α process is slow because the photon usually gets re-absorbed. Under such circumstances, rare decays of the excited hydrogen atom become important, such as two-photon decay:

$$\mathrm{H}(2s) \leftrightarrow \mathrm{H}(1s) + \gamma + \gamma, \tag{55}$$

where neither photon has enough energy to excite a hydrogen atom. This decay has a rate of $\Lambda = 8.2\,\mathrm{s}^{-1}$, and applies to 1/4 of the $n = 2$ H atoms. In accordance with our previous discussion:

$$\dot{x}_2 = -\Lambda\left(\frac{x_2}{4} - x_1 e^{-\omega_{\mathrm{Ly}\alpha}/T}\right). \tag{56}$$

Putting it all together. We can find the net rate of production of excited hydrogen atoms from the above equations:

$$\dot{x}_2 = \alpha n_{\mathrm{H,tot}} x_e^2 - \beta x_2 - \left(\Lambda + \frac{H\omega_{\mathrm{Ly}\alpha}^3}{\pi^2 n_{\mathrm{H,tot}} x_1}\right)\left(\frac{x_2}{4} - x_1 e^{-\omega_{\mathrm{Ly}\alpha}/T}\right). \tag{57}$$

We assume that the excited hydrogen atoms are short-lived so that the rate of production and rate of destruction at any given time are about the same. Then we can set $\dot{x}_2 = 0$ and find:

$$x_2 = 4\frac{\alpha n_{\mathrm{H,tot}} x_e^2 + (\Lambda + \Lambda_\alpha) e^{-\omega_{\mathrm{Ly}\alpha}/T} x_1}{\Lambda + \Lambda_\alpha + 4\beta}, \tag{58}$$

where

$$\Lambda_\alpha \equiv \frac{H\omega_{\mathrm{Ly}\alpha}^3}{\pi^2 n_{\mathrm{H,tot}} x_1}. \tag{59}$$

The net rate of loss of electrons is then:

$$\dot{x}_e = -\alpha n_{\mathrm{H,tot}} x_e^2 + \beta x_2, \tag{60}$$

or

$$\dot{x}_e = -\alpha n_{\mathrm{H,tot}} x_e^2 + 4\beta\frac{\alpha n_{\mathrm{H,tot}} x_e^2 + (\Lambda + \Lambda_\alpha) e^{-\omega_{\mathrm{Ly}\alpha}/T} x_1}{\Lambda + \Lambda_\alpha + 4\beta}. \tag{61}$$

This simplifies to:

$$\dot{x}_e = -\frac{\Lambda + \Lambda_\alpha}{\Lambda + \Lambda_\alpha + 4\beta}(\alpha n_{\mathrm{H,tot}} x_e^2 - 4\beta x_1 e^{-\omega_{\mathrm{Ly}\alpha}/T}). \tag{62}$$

Using the relation for β in terms of α:

$$\dot{x}_e = -\frac{\Lambda + \Lambda_\alpha}{\Lambda + \Lambda_\alpha + 4\beta}\alpha\left[n_{\rm H,tot}x_e^2 - \left(\frac{m_e T}{2\pi}\right)^{3/2} x_1 e^{-\epsilon_0/T}\right]. \tag{63}$$

The term in square brackets is zero in thermal equilibrium (Saha equation!) so this equation satisfies detailed balance. It's called the *Peebles equation* and you'll use it on the homework.

Results. The Peebles equation predicts that recombination is delayed relative to the Saha equation prediction:

- Half of the hydrogen recombines ($x_e = 0.5$) by $z = 1210$ ($z = 1370$ for Saha).
- 90% of the hydrogen recombines ($x_e = 0.1$) by $z = 980$ ($z = 1250$ for Saha).
- 99% of the hydrogen recombines ($x_e = 10^{-2}$) by $z = 820$ ($z = 1140$ for Saha).

Unlike the Saha prediction, the electron abundance freezes out at a nonzero value; modern estimate is 2×10^{-4}.

Improvements. Since the Peebles paper a number of improvements have been suggested and/or included in recombination codes, such as:

- Follow all excited levels of H, rather than lumping them all into a single level. (Speeds up recombination!)
- Radiative feedback: Lyman-β ($1s-3p$; 1026Å) photons are produced early and redshift to Lyman-α (1216Å).
- Collisions.
- Two-photon decays from other excited levels ($3s$, $3d$, etc.)
- Finite width of Lyman-α line.

These effects change recombination at the few percent or less level, but are necessary for high-precision experiments such as *Planck*.

5 Helium and lithium recombination

Helium. Helium has two electrons and recombines in two stages: $He^{2+} \rightarrow He^+$, and $He^+ \rightarrow He$. The ionization energies are 54.4 and 24.6 eV. These are larger than for H so He recombines first.

The first helium recombination obeys a Saha equation,

$$n_{\rm H,tot}\frac{x_e x({\rm He}^{2+})}{x({\rm He}^+)} = \left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-54.4{\rm eV}/T}; \tag{64}$$

the difference from hydrogen is the species on the left side. Half of the He^{2+} recombines to He^+ by $z = 5800$. The reaction rates are fast enough that this recombination proceeds in equilibrium.

The second helium recombination ($He^+ \to He^0$) is trickier. The Saha equation is

$$n_{\rm H,tot} \frac{x_e x({\rm He}^+)}{x({\rm He}^0)} = 4 \left(\frac{m_e T}{2\pi} \right)^{3/2} e^{-24.6 {\rm eV}/T}, \tag{65}$$

with a 4 because to form the ground state of He^0 the two electrons must combine to form a spin singlet. The Saha equation predicts that half of the He^+ recombines to He^0 by $z = 2500$.

However this recombination does not proceed according the Saha equation. One can draw a network of excited levels just as for hydrogen. The ground level is $1s^2\,{}^1S_0$, and the excited levels can be divided into two classes: spin singlets $1snl\,{}^1L_L$ and spin triplets $1snl\,{}^3L_{L-1,L,L+1}$. (Doubly excited states such as $2s2p\,{}^1P_1$ are unimportant.) In the nonrelativistic quantum theory radiative transitions can't flip the electron spins so singlets and triplets are separate. In the full relativistic theory electron spins can flip, but the associated rates are a factor of $\sim \alpha^2$ slower.

The routes to the ground state for He are:

- Two-photon decay from the $1s2s\,{}^1S_0$ level.
- Redshifting of the singlet spectral line $1s2p\,{}^1P_1 \to 1s^2\,{}^1S_0$, 584Å.
- Redshifting of the spin-forbidden spectral line $1s2p\,{}^3P_1 \to 1s^2\,{}^1S_0$, 591Å.
- Absorption of 584Å radiation by the small amount of neutral H present.

The last effect is hard to calculate because it requires analysis of the line shapes and scattering of radiation. We won't do the calculation here but the answer is (see e.g. Switzer & Hirata 2007):

- Half of the He^+ recombines to He^0 by $z = 2000$.
- 90% is recombined to He^0 by $z = 1830$.

Lithium. In principle, Li captures its electrons in 3 stages:

- $Li^{3+} \to Li^{2+}$, ionization energy 122.4 eV.
- $Li^{2+} \to Li^+$, ionization energy 75.6 eV.
- $Li^+ \to Li^0$, ionization energy 5.4 eV.

The first two reactions occur in Saha equilibrium, but have no observable consequences because the Li doesn't contribute significantly to the electron density and hence the opacity.

From the Saha equation, the final stage of lithium recombination should occur at $z \sim 500$ and there was some excitement about this possibility because it was hoped that the Li^0 resonance lines could be observed via scattering of the CMB. However it turns out that the final stage of lithium recombination never happens because the Lyman-α photons from hydrogen recombination (energy 10.2 eV) are sufficient to keep lithium singly ionized. So it's believed that the lithium never captured its last electron.

6 Matter temperature

So far we've considered the ionization state of the gas, but we haven't said much about the temperature. We've implicitly assumed thermal equilibrium between the baryonic matter and the photons, $T_m = T_\gamma$. This is true early on in the Universe but not later (and certainly not today)! This is because the gas in the Universe is subject to several major sources of heating and cooling:

- *Adiabatic expansion* – the gas is expanding so it cools. (It heats later if it falls into a galaxy and is compressed.)
- *Compton heating/cooling* – scattering of CMB photons can heat or cool the gas.
- *Photoionization/recombination* – when an atom is ionized, the energy of the absorbed photon is given to the gas; when it recombines and radiates a photon, the photon's energy is lost.
- *Line emission* – when a collision excites an atom and then it radiates a line photon, the gas is cooled.
- *Shocks* – shocking a gas increases the temperature (and the entropy – not an adiabatic process!)
- *Bremsstrahlung* – collisions of charged particles in a gas cause it to radiate electromagnetic waves.

All of these are important in intergalactic gas (and in galaxies even more effects are important!) But we'll focus on adiabatic expansion and the Compton effect as these dominate at $z > 50$.

The matter has much less heat capacity than the photons (more degrees of freedom) so the CMB redshifts as $T_\gamma \propto a^{-1}$ almost unaffected by the matter temperature. (Exception is Sunyaev-Zel'dovich effect.)

Adiabatic expansion. The particles in a gas are slowed down by the expansion of the Universe because their de Broglie wavelength $\lambda = 2\pi/p$ is redshifted as $\lambda \propto a$. Therefore the momentum of the particle goes as

$$p \propto a^{-1}. \tag{66}$$

Their kinetic energy E_K goes as

$$E_K = \frac{p^2}{2m} \propto a^{-2}, \tag{67}$$

since they are nonrelativistic. Therefore the temperature declines as $T_m \propto a^{-2}$, or

$$\dot{T}_m|_{\rm adiabatic} = -2\frac{\dot{a}}{a}T_m = -2HT_m. \tag{68}$$

If there were no interactions of the matter and photons, the temperature of the matter would decline *faster* than the CMB and thus it would be colder than the CMB.

Compton effect. The main mode of coupling between the baryons and photons is Compton scattering:

$$e^- + \gamma \to e^- + \gamma. \tag{69}$$

This process can change the electron's energy and hence the matter temperature. So let's calculate this in two parts: the heating due to recoil of the electron during scattering of a photon, and the cooling due to energetic electrons giving some of their energy back to the CMB. We'll do this in the nonrelativistic limit (Thomson scattering).

Heating first: when a photon of energy ω scatters off an electron, and is deflected by an angle θ, the momentum transfer is

$$q = |\omega(1,0,0) - \omega(\cos\theta, \sin\theta, 0)| = \omega\sqrt{(1-\cos\theta)^2 + \sin^2\theta}. \tag{70}$$

The energy delivered to the electron is

$$\begin{aligned}\Delta E_{\rm recoil} &= \frac{q^2}{2m_e} = \frac{\omega^2}{2m_e}[(1-\cos\theta)^2 + \sin^2\theta]. \\ &= \frac{\omega^2}{2m_e}[1 - 2\cos\theta + \cos^2\theta + \sin^2\theta] \\ &= \frac{\omega^2}{m_e}(1-\cos\theta). \end{aligned} \tag{71}$$

In order to get a heating rate we need the average energy, which is

$$\langle \Delta E_{\rm recoil} \rangle = \frac{\omega^2}{m_e} \tag{72}$$

because Thomson scattering is equally likely to be forward as backward, hence $\langle \cos\theta \rangle = 0$.

Thus the heating rate in erg/cm^3/s is

$$\Gamma = n_e n_\gamma \sigma_T \langle \Delta E_{\rm recoil} \rangle = n_e n_\gamma \sigma_T \frac{\langle \omega^2 \rangle}{m_e}, \tag{73}$$

where σ_T is the Thomson cross section, and the average is taken with equal weighting of each photon. (Speed of light is 1.)

This is the heating rate for a perfectly cold gas ($T_m = 0$) where all electrons are initially at rest. But the real gas isn't perfectly cold, instead the electrons are moving and will give some of their energy to the CMB via so-called "Compton drag." Consider an electron is moving at velocity v_e in the x-direction, and suppose that the CMB is isotropic in the comoving frame. However in the electron's rest frame, the CMB has a net momentum density: the electron sees more photons going "backward" than forward. To calculate this momentum density, recall that in the comoving frame the CMB stress-energy tensor is

$$T^{\mu\nu} = \begin{pmatrix} \rho_\gamma & 0 & 0 & 0 \\ 0 & \frac{1}{3}\rho_\gamma & 0 & 0 \\ 0 & 0 & \frac{1}{3}\rho_\gamma & 0 \\ 0 & 0 & 0 & \frac{1}{3}\rho_\gamma \end{pmatrix}. \tag{74}$$

The electron's 4-velocity is

$$u^\mu = \frac{1}{\sqrt{1-v_e^2}}(1, v_e, 0, 0), \tag{75}$$

and it carries three spatial vectors:

$$\begin{aligned} (\mathbf{e}_{\hat{1}})^\mu &= \frac{1}{\sqrt{1-v_e^2}}(v_e, 1, 0, 0) \\ (\mathbf{e}_{\hat{2}})^\mu &= (0, 0, 1, 0) \\ (\mathbf{e}_{\hat{3}})^\mu &= (0, 0, 0, 1). \end{aligned} \tag{76}$$

In the electron frame the net momentum density of the photons is (we'll prime the electron frame)

$$j'_\gamma = -T_{\mu\nu}u^\mu(\mathbf{e}_{\hat{1}})^\nu = -\frac{4v_e}{3(1-v_e^2)}\rho_\gamma \approx -\frac{4}{3}v_e\rho_\gamma, \tag{77}$$

where we've taken the nonrelativistic electron limit at the end of the calculation (small v_e). Now since in Thomson scattering a photon is as likely to be reradiated forward as backward, the electron picks up the momentum of each photon it scatters. The force on the electron is then:

$$F = n_\gamma \sigma_T \langle p'_\gamma \rangle, \tag{78}$$

where $n_\gamma \sigma_T$ is the scattering rate (in scatterings per electron per second), and the average is the mean momentum of the photons. But $n_\gamma \langle p'_\gamma \rangle$ is the photon momentum density j'_γ, so:

$$F = \sigma_T j'_\gamma = -\frac{4}{3}\sigma_T \rho_\gamma v_e = -\frac{4}{3}\sigma_T n_\gamma \langle \omega \rangle v_e. \tag{79}$$

The electron's loss of energy is $-\mathbf{F} \cdot \mathbf{v}_e$, so we can calculate net energy loss per unit volume per unit time:

$$\Lambda = -n_e \langle \mathbf{F} \cdot \mathbf{v}_e \rangle = \frac{4}{3}\sigma_T n_e n_\gamma \langle \omega \rangle \langle v_e^2 \rangle. \tag{80}$$

The RMS velocity in a Maxwellian distribution is $\sqrt{3T_m/m_e}$, so

$$\Lambda = 4\sigma_T n_e n_\gamma T_m \frac{\langle\omega\rangle}{m_e}. \tag{81}$$

We can now calculate the net difference of heating and cooling:

$$\Gamma - \Lambda = n_e n_\gamma \sigma_T \frac{\langle\omega^2\rangle - 4T_m\langle\omega\rangle}{m_e}. \tag{82}$$

This is the "textbook" formula for the Compton effect. It is useful in e.g. AGN coronae. For a blackbody radiation field, we can simplify this by using the known equations for the frequency distribution of a blackbody:

$$\begin{aligned} n_\gamma &= \frac{2\zeta(3)}{\pi^2}T_\gamma^3 \\ \langle\omega\rangle &= \frac{\pi^4}{30\zeta(3)}T_\gamma \\ \langle\omega^2\rangle &= \frac{12\zeta(5)}{\zeta(3)}T_\gamma^2. \end{aligned} \tag{83}$$

Here ζ is the Riemann ζ-function; $\zeta(3) = 1.202057$, $\zeta(5) = 1.036928$. So we find that: We can now calculate the net difference of heating and cooling:

$$\Gamma - \Lambda = \frac{4\pi^2}{15} n_e T_\gamma^4 \sigma_T \frac{0.958057T_\gamma - T_m}{m_e}. \tag{84}$$

This is the net heating rate (in the matter) per unit volume per unit time. But there's a big problem: if $0.958057T_\gamma < T_m < T_\gamma$, this formula predicts that the matter *loses* energy to the photons – even though the photon temperature is hotter than the matter! This contradicts the second law of thermodynamics so we must have done something wrong.

It turns out the resolution of the paradox is to consider an additional source of heating: the stimulated-Compton effect,

$$e^- + \gamma \rightarrow e^- + \gamma, \tag{85}$$

in which the emission of the second photon is stimulated. For a blackbody distribution, you will prove (see exercises!) that there is an additional heating source:

$$\Gamma_{\rm stim} = \frac{4\pi^2}{15} n_e T_\gamma^4 \sigma_T \frac{0.041943T_\gamma}{m_e}, \tag{86}$$

which augments the usual Compton terms to give a net heating:

$$\Gamma + \Gamma_{\rm stim} - \Lambda = \frac{4\pi^2}{15} n_e T_\gamma^4 \sigma_T \frac{T_\gamma - T_m}{m_e}. \tag{87}$$

Temperature evolution. In order to finish our calculation, we need the heat capacity per unit volume. For a monatomic gas, this is

$$C_v = \frac{3}{2}n, \tag{88}$$

where n is the total density of all species (electrons, protons, atoms). Considering only hydrogen this is $n_{\mathrm{H,tot}}(1+x_e)$. Thus:

$$\dot{T}_m|_{\mathrm{Compton}} = \frac{\Gamma + \Gamma_{\mathrm{stim}} - \Lambda}{C_v} = \frac{8\pi^2 x_e T_\gamma^4 \sigma_T (T_\gamma - T_m)}{45 m_e (1+x_e)}. \tag{89}$$

The total matter temperature equation is:

$$\dot{T}_m = -2HT_m + \frac{8\pi^2 x_e T_\gamma^4 \sigma_T}{45 m_e (1+x_e)} (T_\gamma - T_m). \tag{90}$$

The behavior of this equation depends on the ratio:

$$R = \frac{4\pi^2 x_e T_\gamma^4 \sigma_T}{45 m_e (1+x_e) H}. \tag{91}$$

If $R \gg 1$ then the coefficient of the second term in Eq. (90) is much larger than the Hubble constant, so to a good approximation we have $T_m \approx T_\gamma$. Once $R \ll 1$ the second term is negligible and the matter temperature decouples from the CMB and redshifts as $T_m \propto a^{-2}$. Since $T_\gamma \propto a^{-4}$ and $H \propto a^{-3/2}$ in matter domination, R is a decreasing function of a (or t). So the matter temperature is coupled to the CMB temperature at early times, and then decouples.

This transition turns out to occur at $z \sim 150$, when $T_\gamma \sim 400$ K. After this the matter temperature declines as $T_m \propto a^{-2}$. By $z = 50$ the CMB has cooled to 135 K, while the matter is at only 50 K. The matter temperature is eventually increased by photoionization heating when the first stars and galaxies turn on.