Lecture VII: The early history of baryonic matter

(Dated: February 15, 2019)

I. INTRODUCTION

We continue our discussion of the thermal history of the Universe by considering baryonic matter. This lecture will treat the period from the early Universe up to $z \sim 70$. This includes:

- *Baryogenesis* – the origin of the baryon-to-antibaryon asymmetry in the Universe.
- *Big Bang Nucleosynthesis* (BBN: t ∼1–200 s) – the conversion of free protons and neutrons into the light elements, the earliest epoch in cosmic history where we have a direct record of what happened and where the key physical processes are understood.
- *Recombination* – the formation of atoms, $t \sim 400$ kyr, and the period we observe directly via the CMB because it marked the transition from an opaque to a transparent Universe.
- *The Dark Ages* – the period where the baryons were in the form of cold neutral gas, but before inhomogeneities could grow enough to form stars.

To study later periods in the history of baryonic matter, we will need to learn about perturbations and galaxy formation/evolution, which will come later in the course.

Comment – The historical reference for this material is Kolb & Turner, *The Early Universe* (1990). Our measurements of light element abundances and our detailed calculations of both BBN and recombination have improved dramatically since then, however many of the basic principles remain in place.

II. BARYOGENESIS

In the last lecture, we assumed that the baryons (and leptons) had zero chemical potential. Since baryon number is conserved at low energies, this is not an automatic consequence of thermalization but is actually an assumption about initial conditions: the net baryon number ($N_{\rm baryon} - N_{\rm antibaryon}$) of the Universe is assumed to be zero. In fact, we observe

$$\eta \equiv \frac{N_{\rm baryon} - N_{\rm antibaryon}}{N_{\gamma,\rm CMB}} = 2.68\times10^{-8}\Omega_b h^2 = 5.9\times10^{-10}. \tag{1}$$

The normalization to the density of CMB photons is conventional, and ensures constancy of the asymmetry parameter η as the Universe expands as long as it is measured after e^+e^- annihilation. (At earlier times, you can preserve the validity of this equation by inserting a factor of $3.36/g_\star$.)

In the "hot Big Bang" picture, if baryon number is exactly conserved, then η is simply a parameter of the initial conditions. However, there are reasons you might think this is not a complete picture of the early Universe:

- Inflation: If inflation occurred, any pre-existing baryon number density should have been exponentially diluted, and then additional photons created as the energy density of the inflaton thermalized. This results in $\eta \to 0$.
- SM baryon number violation: Perturbative (Feynman diagram) processes in the Standard Model preserve baryon number. But non-perturbative (in this case, topological) processes can change baryon number, preserving only baryon minus lepton number, $B - L$.
- Beyond standard model physics: Baryon number violating interactions are common in extensions to the Standard Model such as GUTs and some supersymmetric models.

Some of these reasons might go together. Note also that the lepton asymmetry $L/N_{\gamma,\rm CMB}$ is not nearly so well measured, since we have not measured the difference between the number of ν vs. $\bar\nu$; it may be of order η or not.

If the inflationary picture is correct, then the baryon-antibaryon asymmetry was produced during the hot expanding phase of the Universe, and BSM physics is required (the SM baryon number violation is not enough). Although we don't know the correct physics, we can identify 3 conditions – known as the *Sakharov conditions* – that are necessary in order to produce the asymmetry starting from symmetric (net $B = 0$) conditions.

- *Baryon number violation* (obviously). In the context of GUTs, this could arise from a gauge or Higgs boson that couples to a quark and a lepton, and acquires an enormous mass at GUT symmetry breaking.
- *C and CP violation* – The C violation is needed to distinguish particles and antiparticles. But we also need CP violation, since we can't just produce more right-handed baryons than left-handed antibaryons. CP violation exists in the Standard Model, and many extensions to the SM provide additional opportunities for CP-violating phases.
- *Departure from thermal equilibrium* – Since particle and antiparticle masses are the same in QFT, in thermal equilibrium the entropy will be maximized when $B = 0$, and no baryon asymmetry will develop. Non-equilibrium conditions are thus required.

III. BIG BANG NUCLEOSYNTHESIS

We now discuss the synthesis of the light elements. This is important as the abundances of D, He, and Li are key probes of the early Universe. We consider first the early stages of the Universe where baryonic matter is in protons and neutrons, and then consider nuclear reactions that occur at lower temperatures.

A. Protons vs. neutrons

At temperatures $T \ll m_\pi \sim 140$ MeV, the only form of strongly interacting matter in the Universe is the baryons that did not annihilate with anti-baryons. Their number density scales as $n_b \propto 1/a^3$. Moreover, heavy baryons can decay into lighter baryons, including via weak decays (for those that contain strange quarks), e.g., for the Λ^0 (quark content: *uds*):

$$\Lambda^0 \rightarrow \begin{cases} p^+ + \pi^- \\ n^0 + \pi^0 \end{cases}, \tag{2}$$

so at low temperature $T \sim$ few MeV, only the lightest baryons – the proton (uud) and neutron (udd) – are left over.

Protons and neutrons can be converted into each other at temperatures of a few MeV by weak interactions:

$$n^0 \leftrightarrow p^+ + e^- + \bar{\nu}_e, \quad n^0 + e^+ \leftrightarrow p^+ + \bar{\nu}_e, \quad n^0 + \nu_e \leftrightarrow p^+ + e^-. \tag{3}$$

If the lepton chemical potentials are small compared to T (which will happen if the lepton asymmetry is $\ll 1$), then we expect the equilibrium ratio of neutrons to protons to be

$$\frac{n_n}{n_p} = e^{-(m_n - m_p)/T}. \tag{4}$$

The neutron-proton mass difference is $m_n - m_p = 1.293$ MeV, so at $T \sim$ few MeV, we expect the baryons to be roughly a 50:50 mix of the two species. If thermal equilibrium remained valid, then at a temperature of 1 MeV the neutron:proton ratio would be $e^{-1.293}$:1 or 22:78. At still lower temperatures, the neutron abundance would decline exponentially, leaving only protons at the end.

However, we know that at some temperature $T_{\rm dec}$, the neutrino interactions that maintain the $n : p$ ratio in equilibrium become ineffective. If $T_{\rm dec} = 1$ MeV, then the fraction of baryons X_n that are neutrons when the neutrino interactions stop is 0.22. A detailed calculation gives a result closer to 0.15.

However, even after the interactions with neutrinos or electrons in the initial state in Eq. (3) stop, one reaction – the neutron decay reaction – continues, because it has only one body in the initial state and will run at any density. Thus, following weak decoupling, we expect the neutron fraction to be

$$X_n = 0.15 e^{-t/(880\,\mathrm{s})}, \tag{5}$$

where the neutron lifetime is 880 s. (*Note* – The Particle Data Group [`http://pdg.lbl.gov/`] lists the neutron lifetime as 880.2 ± 1.0 s. It should be noted that this value has been revised downward by 5.5 s since 2010, and the different measurement techniques disagree at this level. Since the rates of Eqs. 3 are all based on the neutron lifetime, this measurement is of direct importance to both the value of X_n at decoupling, and the exponential decay rate, and hence the final He abundance.)

If there were no nuclei, then Eq. (5) would be the end of the story: all neutrons would ultimately decay, and after a few hours the baryons in the Universe would be all protons. However, once the Universe cools enough to form nuclei, some neutrons can become bound into them, and depending on the binding energies the resulting nuclei may be stable against β decay.

B. Nuclei – equilibrium considerations

At sufficiently low temperatures, nuclei are the thermodynamically favored state of nuclear matter. To see this, we recall that a nucleus with Z protons, $A-Z$ neutrons, and A total baryons. Let us denote by $X_{Z,A}$ (or some similar notation) the fraction of baryonic mass in this type of nucleus. The chemical potential of that nucleus is

$$\mu_{Z,A} = m_{Z,A} + T\ln\frac{n_{Z,A}}{(2I_{Z,A})(m_{Z,A}T/2\pi)^{3/2}} = Zm_p + (A-Z)m_n - B_{Z,A} + T\ln\frac{(2\pi)^{3/2}n_bX_{Z,A}}{(2I_{Z,A}+1)A^{5/2}m_N^{3/2}T^{3/2}}, \tag{6}$$

where we approximate m_N to be the mass of the nucleon (where it does not matter n vs. p); $B_{Z,A}$ is the binding energy of the nucleus; and $I_{Z,A}$ is its spin. In *nuclear statistical equilibrium* (NSE), the chemical potential of a nucleus is equal to the sum of all the protons and neutrons needed to make it:

$$\mu_{Z,A} = Z\mu_p + (A-Z)\mu_n. \tag{7}$$

Now protons and neutrons also obey Eq. (6), with spin $\frac{1}{2}$, and so we find:

$$-B_{Z,A} + T\ln\frac{(2\pi)^{3/2}n_bX_{Z,A}}{(2I_{Z,A}+1)A^{5/2}m_N^{3/2}T^{3/2}} = ZT\ln\frac{(2\pi)^{3/2}n_bX_p}{2m_N^{3/2}T^{3/2}} + (A-Z)T\ln\frac{(2\pi)^{3/2}n_bX_n}{2m_N^{3/2}T^{3/2}}. \tag{8}$$

Solving for the abundance of the nucleus Z, A, we find:

$$X_{Z,A} = \frac{(2I_{Z,A}+1)A^{5/2}}{2^A}X_p^ZX_n^{A-Z}\left(\frac{2\pi}{m_NT}\right)^{3(A-1)/2}n_b^{A-1}e^{B_{Z,A}/T}. \tag{9}$$

We can see that the production of a heavy element is favored when the $e^{B_{Z,A}/T}$ factor wins out over the small baryon density n_b. At sufficiently low temperatures, this will always happen. We expect the transition temperature to be when $B_{Z,A}/T$ is a few tens so that this can overwhelm the factors outside the exponential; since nuclear binding energies are a few MeV, this suggests an order of magnitude of ~ 100 keV. The Universe reaches a temperature of 100 keV at $t = 130$ s (see Lecture VI), and so most of the primordial neutrons will still be around and will not have decayed.

In a bit more detail: the deuterium nucleus (D=^{2}H: $Z = 1$, $A = 2$, $B = 2.22$ MeV) becomes favored over $n + p$ at $T = 70$ keV. However, before then, at $T = 280$ keV, the very tightly bound nucleus ^{4}He ($Z = 2$, $A = 4$, $B = 28.3$ MeV) becomes thermodynamically favored over $n + p$. In fact, D is never favored, and if reactions were fast enough to preserve NSE, the Universe would pass from $p + n$, to $p+{}^4$He, to heavy elements, with only trace amounts of D.

C. Nuclei – reaction kinetics

In reality, nuclear reactions, especially between positively charged particles at low density ($\eta \ll 1$), are not fast enough to maintain equilibrium. A detailed prediction of the light element abundances requires writing down the rate equations $\dot{X}_{Z,A}$, taking into account the calculated or measured rates for all important reactions, and solving the system of ordinary differential equations. You can find public codes that do this [1]. I will summarize the main phenomenology here.

At early times, most of the baryons are in p or n, and small amounts of deuterium are present according to:

$$p + n \leftrightarrow \mathrm{D} + \gamma. \tag{10}$$

As the temperature drops to 70 keV at $t \sim 4$ minutes, there is enough deuterium that reactions converting it to ^{4}He become important, such as

$$\mathrm{D} + \mathrm{D} \to {}^3\mathrm{He} + p, \quad {}^3\mathrm{He} + \mathrm{D} \to {}^4\mathrm{He} + p \tag{11}$$

and

$$\mathrm{D} + \mathrm{D} \to {}^3\mathrm{H} + n, \quad {}^3\mathrm{H} + \mathrm{D} \to {}^4\mathrm{He} + n. \tag{12}$$

Since by this time ^{4}He is thermodynamically favored over D, these reactions run almost to completion: in combination with Eq. (10), they quickly "drain" all of the neutrons and convert them to ^{4}He. Since ^{4}He is half neutrons by mass, and we start with $X_n \sim 0.12$, we end with $X(^4\mathrm{He}) \sim 0.24$. The remaining 76% of the mass is still in protons, $X_p = 0.76$.

One might wonder if heavier elements can be created, but alas, the reactions $^4\mathrm{He} + p$, $^4\mathrm{He} + n$, and $^4\mathrm{He} +\, ^4\mathrm{He}$ cannot build heavier nuclei because there are no bound states of ^{5}Li, ^{5}He, and ^{8}Be. In stars, carbon can be made by the triple-alpha reaction:

$$^4\mathrm{He} +\, ^4\mathrm{He} \leftrightarrow\, ^8\mathrm{Be}, \quad ^8\mathrm{Be} +\, ^4\mathrm{He} \rightarrow\, ^{12}\mathrm{C}, \tag{13}$$

where the last reaction occurs before the ^{8}Be can decay, but at the low density and in the short time available in the early Universe this reaction is ineffective. Therefore BBN ends mostly with protons and ^{4}He nuclei. A small fraction of the baryons (a few $\times 10^{-5}$) remain as unburned D, ^{3}He, and ^{3}H (which decays later into ^{3}He) that did not find its way into ^{4}He before the temperature dropped too low for nuclear reactions. An even smaller amount (a few $\times 10^{-10}$) ends up as ^{7}Li via the reaction

$$^4\mathrm{He} +\, ^3\mathrm{He} \rightarrow\, ^7\mathrm{Be} + \gamma, \quad ^7\mathrm{Be} + e^- \rightarrow\, ^7\mathrm{Li} + \nu_e. \tag{14}$$

(The last step is an electron capture decay that occurs at recombination.) In standard cosmology, no other isotopes are made with significant abundances.

D. Predicted yields

I will take the computations in this section from Cyburt et al., Rev. Mod. Phys., 88, 015004 (2016).

The predicted helium abundance by mass, $Y \equiv X(^4\mathrm{He})$, is 0.247. This increases slowly (logarithmically) with η since at higher baryon density the prodcution of ^{4}He happens earlier (but only logarithmically so – see above discussion), and thus more of the neutrons will survive. The helium abundance also increases if there are extra species of neutrino or other invisible particle, since then $g_\star$ is increased and neutrino decoupling occurs at a higher temperature (see Lecture VI), hence with a $n:p$ ratio closer to 1:1.

The D and ^{3}He abundances are usually reported by number relative to ^{1}H: that is, D/H and ^{3}He/H. Since these are thermodynamically disfavored intermediate states in the conversion of neutrons into ^{4}He, their abundance decreases as the baryon density η increases, and thus they can in principle be used as probes of the baryon density. The ratios are roughly:

$$\mathrm{D/H} = 2.58 \times 10^{-5} \left(\frac{\eta}{6.1 \times 10^{-10}}\right)^{-1.6} \quad \text{and} \quad ^3\mathrm{He/H} = 1.00 \times 10^{-5} \left(\frac{\eta}{6.1 \times 10^{-10}}\right)^{-0.6}. \tag{15}$$

The D/H ratio is both more sensitive to the baryon abundance, and can be measured in the IGM with present instrumentation, so it is the preferred baryometer.

The predicted ^{7}Li abundance varies non-monotonically with η: at high η we have $d(^7\mathrm{Li/H})/d\eta > 0$ because high density favors production of ^{7}Be (see above), but at very low η, we have $d(^7\mathrm{Li/H})/d\eta < 0$ because some ^{7}Li can be produced directly and then evade destruction by $^7\mathrm{Li} + p \rightarrow 2\alpha$. The predicted ^{7}Li/H ratio based on the D/H or CMB-favored baryon density ($\eta \sim 6 \times 10^{-10}$) is 4.5×10^{-10}. An even smaller amount of ^{6}Li is produced, $^6\mathrm{Li/H} = 1.3 \times 10^{-13}$.

The Supplement to Lecture VII will discuss the observed abundances of the isotopes D, ^{3}He, ^{4}He, ^{6}Li, and ^{7}Li. Historically, measurements have been complicated both by observational limitations, and by late-Universe nuclear processes (e.g., fusion in stars) that may mean that "measured" abundances are not the primordial values. We will see that the agreement with D and ^{4}He is quite good; the measurements of ^{3}He are difficult to interpret; and the measured abundance of Li is lower than the predictions (the famous "lithium problem").

IV. THE TRANSITION TO MATTER DOMINATION

While not really a thermodynamic event, it is important for our discussion of timescales for us to consider when the Universe transitioned from being radiation-dominated to being matter-dominated. Recall that the density of the Universe varies as

$$\rho(z) = \rho_{\mathrm{crit},0} \left[\Omega_\Lambda + \Omega_\mathrm{m}(1+z)^3 + \Omega_\mathrm{r}(1+z)^4\right]. \tag{16}$$

The present-day density in the CMB is (see Lecture III):

$$\Omega_\mathrm{CMB} h^2 = 2.57 \times 10^{-5}. \tag{17}$$

We can thus see that – since the ratio of radiation to matter density scales as $1 + z$ – the ratio of CMB to matter density is

$$\frac{\rho_{\rm CMB}}{\rho_{\rm m}} = \frac{\Omega_{\rm CMB} h^2}{\Omega_{\rm m} h^2}(1+z) = \frac{2.57 \times 10^{-5}}{\Omega_{\rm m} h^2}(1+z). \tag{18}$$

At early times, before the neutrinos become non-relativistic, we learned in Lecture VI that the total energy density is related to the photon temperature via $g_{\star,\rm eff} = 3.38$: of this, there is a contribution of 2 from the photons and $6 \times \frac{7}{8} \times (\frac{4}{11})^{4/3} = 1.36$ (really, 1.38) from the neutrinos. This means that the ratio of total radiation to CMB is

$$\frac{\rho_{\rm r}}{\rho_{\rm CMB}} = \frac{3.38}{2} = 1.69, \tag{19}$$

and so

$$\frac{\rho_{\rm r}}{\rho_{\rm m}} = 1.69 \times \frac{2.57 \times 10^{-5}}{\Omega_{\rm m} h^2}(1+z). \tag{20}$$

There is *equality* of matter and radiation at a redshift

$$1 + z_{\rm eq} = \frac{\Omega_{\rm m} h^2}{1.69 \times 2.57 \times 10^{-5}} = 23000 \Omega_{\rm m} h^2 \approx 2800, \tag{21}$$

where the last step is for the value from CMB anisotropy measurements, $\Omega_{\rm m} h^2 = 0.12$. Thus at $z \sim 2800$ or $T \sim 7600$ K, the Universe becomes matter-dominated. Most of this matter is of course dark.

During the period of matter domination, we can write the age of the Universe as

$$t = \int \frac{da}{aH} = \int \frac{da}{aH_0\sqrt{\mathcal{E}(a)}} \approx \int \frac{da}{aH_0\sqrt{\Omega_{\rm m} a^{-3}}} = \frac{2}{3\Omega_{\rm m}^{1/2} H_0} a^{3/2} = 2.1 \times 10^{17} (\Omega_{\rm m} h^2)^{-1/2} (1+z)^{-3/2}\,{\rm s}. \tag{22}$$

V. COSMOLOGICAL RECOMBINATION

Cosmological recombination is the epoch where the nuclei captured the free electrons to form neutral atoms. This is also when the Universe went from opaque to transparent, so it is critical for CMB studies – indeed, this is the epoch we "see" when we look at the CMB. These reactions and their ionization energies are:

- $\mathrm{He}^{2+} + e^- \rightarrow \mathrm{He}^+$ ($I = 54.4$ eV)
- $\mathrm{He}^{+} + e^- \rightarrow \mathrm{He}$ ($I = 24.6$ eV)
- $\mathrm{H}^{+} + e^- \rightarrow \mathrm{H}$ ($I = 13.6$ eV)

(I won't focus on lithium as it is not abundant enough to affect the recombination epoch in an observable way.) Based on the ionization energies, you can guess that this is the order in which the recombinations take place. Hydrogen carries 86% of the electrons, and it recombines last, so in terms of overall importance for the CMB, it is dominant; but helium must be understood at the present-day levels of precision of the measurements.

A. Equilibrium physics

Considerations similar to NSE lead us in the case of ionization states of atoms to the *Saha equation*, which describes the ionization of an element in thermal equilibrium:

$$\frac{n(\mathrm{X}^{i+})}{n_e n(\mathrm{X}^{(i+1)+})} = \frac{g(\mathrm{X}^{i+})}{2g(\mathrm{X}^{(i+1)+})} \left(\frac{2\pi}{m_e T}\right)^{3/2} e^{I/T}, \tag{23}$$

where I is the ionization energy of X^{i+}, and g represents the statistical degeneracy of the ground state (4 for H^0 and 2 for H^+, if we include spins).

For hydrogen recombination, if we denote the ionization fraction by x, and we assume the mass fraction of hydrogen is $X_p = 0.76$, then the left-hand side can be written as $(1-x)/(x^2 X_p n_b)$. The number density of baryons is

$$n_b = \eta n_\gamma = \eta \times \frac{2\zeta(3)}{\pi^2} T^3, \tag{24}$$

where $\zeta(3) \approx 1.202$ is the Riemann ζ-function. The Saha equation then says

$$\frac{1-x}{x^2} = \frac{2\zeta(3)}{\pi^2} X_p \eta T^3 \left(\frac{2\pi}{m_e T}\right)^{3/2} e^{I/T} = 3.836 X_p \eta \left(\frac{T}{m_e}\right)^{3/2} e^{I/T}. \tag{25}$$

Thus the quantity $(1-x)/x^2$ starts small in the early Universe because η is small, but then increases as T drops and the exponential factor takes over. Accordingly, x starts from near 1 and then drops to near 0. For $\eta = 5.9 \times 10^{-10}$ and $X_p = 0.76$, we reach the half-ionized point ($x = 0.5$, or $(1-x)/x^2 = 2$) at a temperature $T = 0.32$ eV. This corresponds to $T = 3700$ K in usual units, and given the present-day CMB temperature of $T_0 = 2.725$ K, and using $T = T_0(1+z)$, this implies a redshift $z_{\rm rec} = 1360$.

B. Reaction kinetics

In reality, hydrogen recombination (and He and Li recombination) proceeds in a strongly out-of-equilibrium mode. The problem was first treated in the late 1960s by Peebles and Zel'dovich & Sunyaev, and was considered in great detail in the 2005–12 period by several groups (including your professor) in support of the *Planck* CMB mission. A complete treatment requires one to follow a system of differential equations describing the abundances of H atoms in every level (including the ground state 1s, and the excited states; modern calculations go up to n of a few hundred). Moreover, radiation emitted by the atoms can excite neighboring atoms, so one must also track the spectrum of radiation produced by H recombination and how this adds to the blackbody radiation. What follows is a simplified description of how the process takes place. An interested student may download the publicly available recombination codes [2].

Hydrogen atoms recombine according to the process

$$\mathrm{H}^+ + e^- \leftrightarrow \mathrm{H}(nl) + \gamma. \tag{26}$$

The rate coefficient for the forward process – known as the *recombination coefficient* – is $\alpha \sim 8 \times 10^{-19}$ m^3 s^{-1} at $T = 3700$ K (it varies as $\alpha \propto T^{-0.7}$ in this temperature range), if we sum over all levels nl. At $z = 1360$ or $T = 0.32$ eV, the baryon density from Eq. (24) is

$$n_b = \eta \times \frac{2\zeta(3)}{\pi^2} T^3 = 4.7 \times 10^{-12}\,\mathrm{eV}^3 = 6 \times 10^8\,\mathrm{m}^{-3}. \tag{27}$$

(the last step converted to m^{-3} by setting $\hbar c = 2 \times 10^{-7}$ eV m equal to 1). From Eq. (22), the age of the Universe at $z = 1360$ and $\Omega_{\rm m} h^2 = 0.12$ is $t = 1.2 \times 10^{13}$ s. Since

$$\alpha n_b t \sim (8 \times 10^{-19}\,\mathrm{m}^3\,\mathrm{s}^{-1})(6 \times 10^8\,\mathrm{m}^{-3})(1.2 \times 10^{13}\,\mathrm{s}) = 6000 \gg 1, \tag{28}$$

we expect that a hydrogen nucleus should be able to capture an electron in a time short compared to the age of the Universe. If this were all that happened, H recombination, Eq. (26) would proceed in equilibrium.

However, the real situation has an additional complication. If Eq. (26) produces an atom in the ground state H(1s), then the photon emitted has an energy exceeding 13.6 eV. Once a small fraction of the hydrogen is neutral, the Universe is opaque to these photons, and every photon emitted by this process results in an ionization. The only way to avoid this is to recombine first to an excited state, and then drop down to the ground state, e.g.,

$$\mathrm{H}^+ + e^- \leftrightarrow \mathrm{H}(n \geq 2) + \gamma, \quad \ldots \quad \left\{ \begin{array}{c} \mathrm{H}(2s) \leftrightarrow \mathrm{H}(1s) + \gamma + \gamma \\ \mathrm{H}(2p) \leftrightarrow \mathrm{H}(1s) + \gamma(h\nu = 10.2\,\mathrm{eV}). \end{array} \right. \tag{29}$$

Here 10.2 eV is the energy difference between the $n = 1$ and $n = 2$ levels of hydrogen. Excited H atoms can ultimately decay to the ground state by either the two-photon decay from H(2s) (which has a lifetime of 0.12 s, and results in two low-energy photons that are unlikely to excite another hydrogen atom) or by emission of the Lyman-α line with energy 10.2 eV from H(2p) (which has a lifetime of 1.6 ns). The latter channel is additionally suppressed because the Lyman-α photon will get absorbed by another H atom – in this case, there is one decay, one excitation, and overall nothing happens. But in an expanding Universe, there is a small probability that before this happens, the photon will redshift out of the line, resulting in a net production of H($1s$). These decay compete with the alternative possibility: that the excited H atom might get ionized by absorbing a photon from the CMB (which is in the optical/infrared at this epoch). A detailed calculation shows that recombination proceeds slowly, reaching 50% ionization at $z = 1280$, 10% ionization at $z = 1060$, and 1% ionization at $z = 880$.

It is a remarkable coincidence that the two channels in Eq. (29) have similar contributions. It is also remarkable that the fit to the CMB anisotropy data gives the best measurement of the H($2s$) inverse lifetime: $\tau^{-1} = 7.7 \pm 0.6$ s^{-1}, versus the theoretical value of 8.22 s^{-1} (see Planck 2015 XIII, A&A 594, A13).

Recombination is a second-order reaction, which means that the rate of the reaction is proportional to the abundance of reactants squared (we require one proton and one electron). A key feature of such reactions is that they do not run to completion. In this case, one the Universe gets cold enough that the CMB can't ionize a hydrogen atom from $n = 2$, the rate of reactions $-n_{\rm H}\,dx/dt$ is equal to α times the density of electrons squared, $(n_{\rm H}x)^2$:

$$-n_{\rm H}\frac{dx}{dt} = \alpha(n_{\rm H}x)^2 \quad \to \quad -\frac{dx}{x^2} = \alpha n_{\rm H}\,dt, \tag{30}$$

where $n_{\rm H}$ is the density of all forms of hydrogen (ionized or neutral). Integrating gives

$$\frac{1}{x} = \frac{1}{x_{\rm init}} + \int \alpha n_{\rm H}\,dt. \tag{31}$$

We argued that the last term is ~ 6000 (see Eq. 29), so you can correctly guess that hydrogen recombination will end with an ionization fraction of $\sim 1/6000 \approx 2\times10^{-4}$. This number is right (the factors of order unity are right, which is fortuitous). This phenomenon is called *residual ionization*.

VI. THE DARK AGES

The period following recombination is often known as the "Cosmic Dark Ages," meaning that this period is not directly seen via the CMB (the Universe is transparent), but there are not yet any self-luminous sources such as stars. Despite the lack of direct observations, this epoch is important both for setting the initial conditions for galaxy formation, and because we hope to observe it in the future via hydrogen 21 cm absorption against the CMB.

A. Thermal evolution

The temperature of the CMB during the Dark Ages continues to scale as $T_\gamma \propto 1 + z$. The baryons have negligible heat capacity compared to the CMB since $\eta \ll 1$. However, when the Universe is sufficiently dilute, the baryons may have a temperature T_b that is different from T_γ. If the baryons stopped interacting with the CMB entirely, then we would have

$$T_b \propto V^{-2/3} \propto (1+z)^2, \tag{32}$$

as appropriate for an adiabatically expanding monatomic gas. But in reality, the baryons can scatter CMB radiation via Compton scattering:

$$e^- + \gamma \leftrightarrow e^- + \gamma. \tag{33}$$

Compton scattering can both transfer energy from the gas to the CMB, and from the CMB to the gas. Transfer from the gas to the CMB occurs because an electron with velocity v sees more photons (and more energy per photon) coming from in front of it than behind (recall the Doppler effect and the relativistic aberration of light). The electron dumps its energy into the CMB after it scatters its own mass in photons, that is, after a time $\sim m_ec^2/(u_\gamma\sigma_{\rm T}c)$ (where $\sigma_{\rm T}$ is the Thomson cross section). Now if a fraction $f = x/(1 + x)$ of the particles are electrons, the timescale for Compton scattering to dump all of the energy of the gas into the CMB is

$$t_{\rm Compton} \sim \frac{m_ec}{fu_\gamma\sigma_{\rm T}} = \frac{(1+x)m_ec}{xu_\gamma\sigma_{\rm T}}. \tag{34}$$

On this timescale, the baryonic gas should come to thermal equilibrium with the CMB.

Using $\sigma_T = 6.65 \times 10^{-29}$ m^2, $u_\gamma = a_{\rm rad} T_\gamma^4 = 4.2 \times 10^{-14}(1+z)^4$ J m^{-3}, and $x = 2 \times 10^{-4}$, we find

$$t_{\rm Compton} \sim 4.9 \times 10^{23}(1+z)^{-4}\,{\rm s}. \tag{35}$$

Comparing to the age of the Universe, Eq. (22), with $\Omega_{\rm m} h^2 = 0.12$, we find

$$\frac{t_{\rm Compton}}{t} = 8 \times 10^5 (1+z)^{-5/2} = \left(\frac{1+z}{230}\right)^{-5/2}. \tag{36}$$

At early times – $z > 230$ – the Compton timescale is short, and the gas is in thermal equilibrium with the CMB. At later times, the Compton timescale is long, and the gas begins to cool as $(1+z)^2$ instead of $1+z$. (A more detailed calculation shows that the effective transition redshift is ~ 150, but our estimate of 230 was close.) Thus the gas in the Universe is colder than the CMB, until some other energy source (e.g., first stars, accreting black holes, etc.) heats it back up again.

If the gas has any lines, we expect to see them in absorption against the CMB, and then emission as early astrophysical sources heat it up. Given that it is H and He, the most likely choice will be the 21 cm hyperfine line of hydrogen – but, given that this is redshifted to 1.5 m (at $z = 6$) to 30 m ($z \sim 150$), and how bright our own Galaxy is at these wavelengths due to synchrotron emission, this is a very challenging experiment and the one candidate detection to date (EDGES) is still under debate.

B. Chemistry

As the Universe continues to expand and cool, one might wonder if the hydrogen converts from atomic to molecular form (which is thermodynamically favored at low temperatures). The answer, for the most part, is no. Because there is no dipole moment, we do not expect the reaction:

$$\mathrm{H} + \mathrm{H} \nrightarrow \mathrm{H}_2 + \gamma. \tag{37}$$

In the modern-day Universe, the production of H_2 is catalyzed by dust grains, but these are not present prior to the formation of heavy elements in stars. Instead, molecule production is catalyzed by the residual electrons:

$$\mathrm{H} + e^- \leftrightarrow \mathrm{H}^- + \gamma, \quad \mathrm{H}^- + \mathrm{H} \rightarrow \mathrm{H}_2 + e^-. \tag{38}$$

Detailed calculations (e.g., Hirata & Padmanabhan 2006) show that the total amount of molecular hydrogen produced in mean-density regions of the Universe is only $n(\mathrm{H}_2)/n(\mathrm{H}) \sim 6 \times 10^{-7}$, with production peaking at $z \sim 70$.

---

[1] See, for example, AlterBBN: `https://alterbbn.hepforge.org`
[2] See: HyRec [`https://cosmo.nyu.edu/yacine/hyrec/hyrec.html`] and
CosmoRec [`http://www.jb.man.ac.uk/~jchluba/Science/CosmoRec/CosmoRec.html`].