4 Big Bang Nucleosynthesis

4.1 Introduction

The theory describing the creation of light elements in the universe is, together with the properties of the CMB, one of the physical cornerstones of the Hot Big Bang model of the universe. Though the geometric properties of the universe are described using General Relativity via the Robertson–Walker metric and the Friedmann equation, the theories of the the CMB and Big Bang Nucleosynthesis (BBN) tackle the physical evolution of the contents of the universe and their relationship to cosmological parameters. The set of required cosmological parameters has also begun to expand in response to a more comprehensive model, i.e. $\Omega_M = \Omega_b + \Omega_{DM} + \Omega_?$ and $\eta = n_\gamma / n_b$.

We have discussed in Lecture 3 that at the time of recombination the universe consisted of photons, neutrinos, dark matter particles, protons and electrons together with small proportions of light elements. The creation of these light elements is the topic of this lecture.

We also discussed how temperature evolution of the universe as a function of cosmic time, developing the equations that allow us to predict the radiation density/temperature as a function of cosmic epoch.

Photons are the principal energy carriers in the universe; the radiation temperature therefore defines the physical reactions that will occur (alternative, those that will be in thermodynamic equilibrium) at each Universal epoch – essentially one compares the 'typical' photon energy to the energy required for a given reaction to occur.

In this lecture we will discuss the production of light nuclear isotopes in the early universe: hydrogen, deuterium, helium-3, helium-4, and lithium-7. Heavier elements, we believe, were synthesized much later, in the interior of stars, and were then explosively ejected into interstellar space.

4.2 A short cosmic history

The energy density in the early universe is dominated by radiation. Therefore,we may write

$$a(t) \propto t^{1/2}; \quad T(t) \propto a(t)^{-1}; \quad T(t) \propto t^{-1/2}. \tag{1}$$

Ryden uses an analysis of the conditions at the Planck time to obtain the constant of proportionality. Using radiation-matter equality also provides a reasonable answer, i.e.

$$T(t) \simeq 10^{10}\ \mathrm{K} \left(\frac{t}{1\mathrm{s}}\right)^{-1/2} \tag{2}$$

$$kT(t) \simeq 1\ \mathrm{MeV} \left(\frac{t}{1\mathrm{s}}\right)^{-1/2} \tag{3}$$

$$E_{mean} = 2.7kT(t) \simeq 3\ \mathrm{MeV} \left(\frac{t}{1\mathrm{s}}\right)^{-1/2} \tag{4}$$

These formulae provide us with the energy scale as a function of epoch in the early universe. A comparison to the energy associated with particle interactions provides an understanding the physical conditions associated with each epoch (Figure 1).

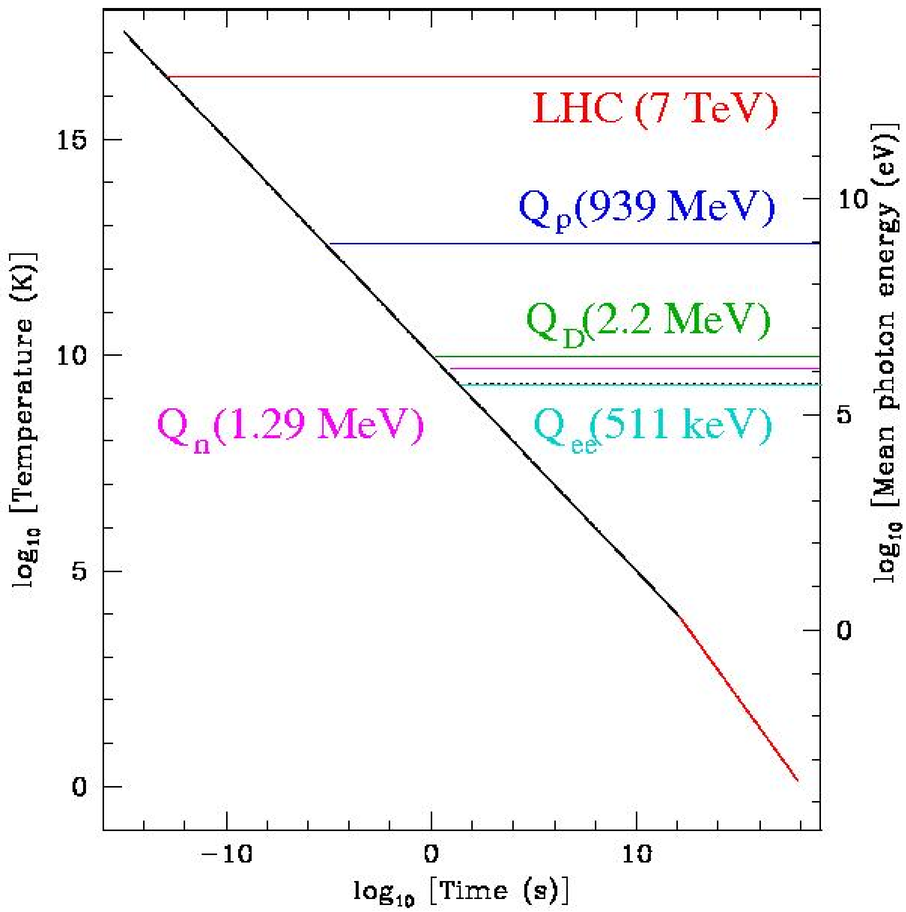

Figure 1: The radiation temperature/energy as a function of cosmic time.

1. **The hadron era:** $t_{min} < t \lesssim 10^{-5}$ sec. This marks the earliest Universal epoch where experimental physics can be applied with any confidence. Almost all matter, including electrons, protons, neutrons, neutrinos and their associated anti–particles are in thermal equilibrium with the photon radiation field. The disparity between particles and anti–particles is thought to be 1 part in $\gtrsim 10^7$ and is ultimately responsible for all matter in the present day universe. The exact physics (e.g. equation of state) of this epoch is not known. Thus the dependence of the scale factor $a(t)$ and the temperature $T(t)$ upon cosmic time is not known.

2. **The lepton era:** 10^{-5} sec $\lesssim t \lesssim 10$ sec. The temperature decreases such that kT is significantly lower than the rest mass energy of the proton ($m_p = 938$ MeV). Proton–antiproton pairs, in addition to other hadrons present, annihilate. The lepton era begins with photons in thermal equilibrium with electrons and positrons, muons, neutrinos and antineutrinos. The energy released by hadron annihilation is thus shared between all of these particle families (see Rees for an interesting discussion of later epochs where the energy partition is not equal). Each of the relativistic particles (photons, neutrinos and electrons – plus antiparticles) contributes an energy density $\propto T^4$. The lepton era ends when the radiation temperature drops significantly below $T \simeq 5 \times 10^9$ K (i.e. $kT \simeq m_e c^2 = 511$ keV). Electron–positron pairs annihilate and temperatures begin to decrease to levels where a protons fall out of equilibrium with neutrons.

3. **The plasma era:** 10 sec $\lesssim t \lesssim 10^{13}$ sec. The universe consists of photons, neutrinos, electrons, protons and neutrons (the discussion assumes that at this stage Dark Matter particles do not interact). The early stages of the plasma era remain sufficiently hot and dense to produce light nuclear elements from hydrogen. Matter and radiation are coupled to form a photon–baryon fluid discussed in Lecture 3: photons are coupled to electrons via Thompson scattering, and electrons are coupled to protons via Coulomb interactions. Matter and radiation remain in thermal equilibrium until the photon temperature drops below $T \simeq 3 \times 10^3$ K where electrons combine with protons to form atomic hydrogen. The radiation field continues unimpeded to the present day where it is observed as the Cosmic Microwave Background.

4. **The post–recombination era:** $t \gtrsim t_{rec} \simeq 10^{13}$ sec. Various astrophysical processes combine to produce the present day universe.

4.3 Neutron production

Neutrons existed in thermodynamic equilibrium with protons, electrons and neutrinos in the early universe, i.e.

$$p + e^- \rightleftharpoons n + \nu_e \tag{5}$$

$$p + \bar{\nu}_e \rightleftharpoons n + e^+. \tag{6}$$

Electrons and positrons are also in thermal equilibrium at this epoch, i.e.

$$\gamma + \gamma \rightleftharpoons e^- + e^+. \tag{7}$$

Finally, isolated neutrons decay via beta–decay with a decay time of $\tau_n = 890$ seconds ($n_n(t) = n_n(t=0)\exp(-t/\tau_n)$),

$$n \rightarrow p + e^- + \bar{\nu}_e. \tag{8}$$

In thermal equilibrium, the numbers of neutrons and protons are given by the Maxwell-Boltzmann distribution (a version of which we have already used for Saha's equation in Lecture 3)

$$n_n = g_n \left(\frac{m_n kT}{2\pi\hbar^2}\right)^{3/2} \exp\left(-\frac{m_n c^2}{kT}\right) \tag{9}$$

$$n_p = g_p \left(\frac{m_p kT}{2\pi\hbar^2}\right)^{3/2} \exp\left(-\frac{m_p c^2}{kT}\right) \tag{10}$$

$$\frac{n_n}{n_p} = \frac{g_n}{g_p} \left(\frac{m_n}{m_p}\right)^{3/2} \exp\left(-\frac{(m_n - m_p)c^2}{kT}\right). \tag{11}$$

We can employ a number of simplifications

- $g_n = g_p = 2$.
- $(m_n/m_p)^{3/2} = 1.002$.
- $(m_n - m_p)c^2 = Q_n = 1.29$ MeV.

Therefore, in equilibrium the neutron to proton ratio is

$$\frac{n_n}{n_p} = \exp\left(-\frac{Q_n}{kT}\right). \tag{12}$$

As $Q_n = 1.29$MeV, this implies a corresponding value of $kT \approx 1.5 \times 10^{10}$K. Therefore, at $T \gg 1.5 \times 10^{10}$K we expect $n_n = n_p$ and at $T \ll 1.5 \times 10^{10}$K we expect $n_n \ll n_p$, with the changeover between these two epochs occurring at $t \sim 1$ second.

In order to observe any neutrons at all in the present day universe, two events must have occurred:

1. Neutron-proton equilibrium must break down.
2. Free neutrons must be bound up into atomic nuclei.

Equations 5 and 6 demonstrate that neutron-proton equilibrium is maintained by the exchange of neutrinos. Neutrinos interact via the weak nuclear force. At the energy scale we are dealing with, the cross-section for a neutrino to interact with another particle via the weak nuclear force is

$$\sigma_w \sim 10^{-47}\mathrm{m}^2 \left(\frac{kT}{1\mathrm{MeV}}\right)^2 \propto t^{-1}. \tag{13}$$

As the number density of neutrinos $n_\nu \propto a^{-3} \propto t^{-3/2}$, the interaction rate may be written as

$$\Gamma = n_\nu \sigma_w c \propto t^{-5/2}. \tag{14}$$

In the radiation dominated era, the expansion rate $H \propto t^{-1}$ and neutrinos begin to decouple from neutrons and protons when $\Gamma \sim H$. The exact epoch at which neutrinos "freeze out" is set by $\sigma_{w,n}$ and $\sigma_{w,p}$. Using values determined from particle physics experiments yields the values $kT_{freeze} = 0.8$MeV, $T_{freeze} = 9 \times 10^9$K, or $t_{freeze} \sim 1$ second. The neutron to proton ratio is fixed at this epoch

$$\frac{n_n}{n_p} = \exp\left(-\frac{Q_n}{kT}\right) = \exp\left(-\frac{1.29\text{MeV}}{0.8\text{MeV}}\right) \simeq 0.2, \tag{15}$$

i.e. $n_n : n_p$ equals $1 : 5$ for $t_{freeze} < t \ll \tau_n$.

4.4 The deuterium bottleneck

As mentioned earlier, free neutrons decay with a well–defined decay time. This decay time becomes effectively infinite (i.e. they do not decay) once the neutrons are bound within an atomic nucleus. The principal reaction that binds neutrons in this way is deuterium creation,

$$p + n \rightleftharpoons \text{D} + \gamma. \tag{16}$$

However, at $T \simeq 10^{10}$K the typical photon energy is 3 MeV. The value Q_D or binding energy of the deuterium nucleus is 2.2 MeV. Therefore, at $T \simeq 10^{10}$ there exists $n_\gamma(kT > 2.2\,\text{MeV}) \gg n_D$ and deuterium is rapidly photo–dissociated. The universe must therefore expand further until the radiation temperature has decreased to a level where deuterium can survive for long enough for further nuclear reactions to proceed. To define the epoch of stable deuterium production we once again return to the Maxwell-Boltzmann distribution describing the relative populations in each physical state, i.e.

$$\frac{n_D}{n_p n_n} = \frac{g_D}{g_p g_n}\left(\frac{m_D}{m_p m_n}\right)^{3/2}\left(\frac{kT}{2\pi\hbar^2}\right)^{-3/2}\exp\left(\frac{(m_p + m_n - m_D)c^2}{kT}\right). \tag{17}$$

We can once again employ a number of simplifications

- $g_D = 3$, $g_n = g_p = 2$.
- $m_p = m_n = m_D/2$
- $(m_p + m_n - m_D)c^2 = Q_D = 2.2$ MeV,

to obtain the ratio of deuterium to neutrons and protons

$$\frac{n_D}{n_p n_n} = 6\left(\frac{m_n kT}{\pi\hbar^2}\right)^{-3/2}\exp\left(\frac{Q_D}{kT}\right). \tag{18}$$

Therefore, deuterium is favoured at $kT \to 0$ and free protons and neutrons are favoured at $kT \to \infty$. We define the epoch at which deuterium nucleosynthesis occurs via the ratio $n_D/n_n = 1$, i.e. when one-half of the neutrons have been converted to deuterium, i.e.

$$\frac{n_D}{n_n} = 6n_p \left(\frac{m_n kT}{\pi\hbar^2}\right)^{-3/2} \exp\left(\frac{Q_D}{kT}\right) = 1. \tag{19}$$

We can simplify this expression further by removing the dependence upon the proton density. At the beginning of Deuterium nucleosynthesis we know that $n_n : n_p$ equals $1:5$ and thus

$$n_p \simeq 0.8 n_b = 0.8\eta n_\gamma = 0.8\eta \left[0.243 \left(\frac{kT}{\hbar c}\right)^3\right]. \tag{20}$$

The Deuterium to neutron ratio can now be written as

$$\frac{n_D}{n_n} \simeq 6.5\eta \left(\frac{kT}{m_n c^2}\right)^{3/2} \exp\left(\frac{Q_D}{kT}\right) = 1. \tag{21}$$

Taking $\eta = 5.5 \times 10^{-10}$ yields $kT_{nuc} = 0.066$MeV, or $T_{nuc} = 7.6 \times 10^8$K, or $t_{nuc} \simeq 200$s. The neutron to proton ratio at t_{nuc} will have fallen to

$$\frac{n_n}{n_p} \simeq \frac{\exp(-t_{nuc}/\tau_n)}{5 + [1 - \exp(-t_{nuc}/\tau_n)]} \simeq \frac{0.8}{5.2} \simeq 0.15 \simeq \frac{1}{7}. \tag{22}$$

Deuterium creates an effective 'bottleneck' in the production of heavier atoms. The binding energy of deuterium is lower than T (6.257 MeV) or He^3 (5.494 MeV). However, these more stable nuclei can only be created by the combination of $D + p/n \to He^3/T + \gamma$ or three particle interactions (much rarer).

With the beginning of the creation of 'stable' deuterium the universe enters a brief phase where conditions of temperature and density permit the creation of heavier atomic elements via interactions governed by basic statistical mechanics and nuclear interactions. However, the expanding universe may be thought of as the 'ticking clock' – the universe is expanding and cooling rapidly. The interplay of these two effects governs the production of light nuclear elements in the primordial fireball (i.e. photon energies and particle densities decrease – reaction rates decrease).

4.5 Light element production

A number of two–particle reactions contribute to the creation of light nuclear elements – see Wagoner et al. (1967) for a full list – some of the principal reactions are,

$$n + p \rightleftharpoons D + \gamma$$

$$\begin{aligned}
p + \mathrm{D} &\rightleftharpoons \mathrm{He}^3 + \gamma \\
n + \mathrm{D} &\rightleftharpoons \mathrm{T} + \gamma \\
n + \mathrm{He}^3 &\rightleftharpoons p + \mathrm{T} \\
p + \mathrm{T} &\rightleftharpoons \mathrm{He}^4 + \gamma \\
n + \mathrm{He}^3 &\rightleftharpoons \mathrm{He}^4 + \gamma \\
p + \mathrm{Li}^7 &\rightleftharpoons \mathrm{He}^4 + \mathrm{He}^4 \\
n + \mathrm{Be}^7 &\rightleftharpoons \mathrm{He}^4 + \mathrm{He}^4
\end{aligned} \tag{23}$$

The complete calculation of the relative abundance of all atomic elements as a function of time and/or baryon density consists of a series of linked 1st order differential equations. Consider the rate of change of the number density of nucleus i, i.e. $\mathrm{d}n_i/\mathrm{d}t$. For convenience we consider the changing mass fraction as a function of time, i.e. $\mathrm{d}X_i/\mathrm{d}t$ where,

$$X_i = \frac{A_i n_i}{\rho_b N_A}. \tag{24}$$

In this case, A_i is the atomic number, ρ_b is the baryon density and N_A is Avogadro's number. Mass conservation requires $\Sigma_i X_i = 1$.

Consider a series of reactions that create or destroy nucleus i, i.e.

$$i + \gamma/e^{\pm}/\nu/\bar{\nu} \rightarrow j + k \quad \text{Photo–disintegration or weak interactions} \tag{25}$$

$$j + k \rightarrow i + l \quad \text{Two particle interactions} \tag{26}$$

$$j + k + l \rightarrow i + m \quad \text{Three particle interactions} \tag{27}$$

Note that in Equations 26 and 27, the additional products, l and m, can be either leptons, photons or baryons.

Consider Equation 25. The rate of change of nucleus i may be written as

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = \text{number of reactions per unit time} = n_i \times \frac{1}{\tau_l(i)}, \tag{28}$$

where $\tau_l(i)$ is the time–scale for the photon/lepton field, l, to dissociate nucleus i. One may therefore define

$$\lambda_l(i) = 1/\tau_l(i) = \int_0^\infty n_l(E)\,\sigma_l(E)\,v_l(E)\,\mathrm{d}E \tag{29}$$

as the interaction rate, where $n_l(E)$, $\sigma_l(E)$, $v_l(E)$ are the number, cross–section and velocity respectively of photons/leptons per unit energy interval. Returning to Equation 28 one may use Equation 24 to write

$$\begin{aligned}
\frac{1}{A_i}\rho_b N_A \frac{\mathrm{d}X_i}{\mathrm{d}t} &= \pm\frac{1}{A_i}\rho_b N_A X_i \lambda_l(i) \\
\frac{1}{A_i}\frac{\mathrm{d}X_i}{\mathrm{d}t} &= \pm\frac{1}{A_i} X_i \lambda_l(i),
\end{aligned} \tag{30}$$

where the sign of the relation describes whether nucleus i is created or destroyed in the reaction.

Consider now Equation 26. The rate of change of nucleus i may now be written as

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = n_j n_k \times \frac{1}{\tau_l(jk)} \tag{31}$$

In this case $\tau_l(jk)$ is described by the joint distribution function

$$\frac{1}{\tau_l(jk)} = \langle \sigma v \rangle_{jk} = \int_0^\infty f(v,T)\,\sigma(v)\,v\,\mathrm{d}v, \tag{32}$$

where $f(v,T)$ is the Boltzmann distribution for nuclei at a temperature T and relative velocity v and $\sigma(v)$ is the cross–section as a function of relative velocity. Therefore one may write

$$\frac{1}{A_i}\frac{\mathrm{d}X_i}{\mathrm{d}t} = \pm \frac{X_j}{A_j}\frac{X_k}{A_k}\,[jk]_l, \tag{33}$$

where $[jk]_l$ is defined as

$$[jk]_l = \rho_b N_A \langle \sigma v \rangle_{jk}. \tag{34}$$

A similar relation may be defined for three particle interactions. The total rate of change of abundance of nucleus i is therefore

$$\frac{1}{A_i}\frac{\mathrm{d}X_i}{\mathrm{d}t} = \pm \sum_i \frac{1}{A_i} X_i \lambda_l(i) \pm \sum_{j \geq k} \frac{X_j}{A_j}\frac{X_k}{A_k}\,[jk] \pm \sum_{j \geq k \geq l} \frac{X_j}{A_j}\frac{X_k}{A_k}\frac{X_l}{A_l}\,[jkl] \tag{35}$$

Solution of these linked equations for all nuclei as a function of time (which implies changing temperature and density) generates the nuclear abundance as a function of time. Computing the final values at large time for a range of baryon densities generates the nuclear abundance as a function of baryon density. **See Figures 3 and 4 in Wagoner et al.** Note that the only free parameter is η or, given that one knows the radiation density, the baryon density. Some of the main points from the BBN calculation are as follows:

1. To 1st order, all neutrons end up in He^4. The He^4 mass fraction may therefore be written as $Y = 2n/(n+p)$, or twice the neutron mass fraction. For a neutron–to–proton ratio of 1:7, the He^4 mass fraction is $Y = 0.25$ – exceptionally close to the observed value (see later).

2. Other light elements survive as intermediate nuclear products. The relative amount of neutron carrying by-products constrains the Universal baryon density. For example:

    (a) Deuterium is photo–dissociated and destroyed in collisions. Increasing the baryon density increases the collision rate and consumes the available deuterium at an increased rate.

   (b) Tritium is not observed a primordial nuclear element. It is converted via beta–decay to He^3 with a half–life of 12.41 years.
   (c) He^3 is inversely dependent upon the baryon density but to a lesser extent than deuterium.
   (d) He^4 depends upon the total baryon density. Increased particle density breaks the deuterium bottleneck at higher temperatures (earlier times). Less free neutrons have decayed and thus slightly more He^4 can be produced.
   (e) Li^7 is another density enhanced product. Increased particle density permits more He^3 and He^4 to be converted to Li^7 and Be^7.

3. Very few elements heavier than He^4 are produced – only Li^7 via direct production and Be^7 decay. There are no stable nuclei with $A = 5$ or 8. The collision of two helium nuclei can bridge the $A = 5$ gap - though only rarely. To jump the $A = 8$ gap directly requires the triple–alpha reaction – $3He^4 \rightleftharpoons C^{12} + \gamma, \gamma$ – this reaction requires sufficiently high densities that the predicted amounts are tiny (if any). Thus the linked differential equations are limited to a relatively small subset of the available reactions presented in Wagoner et al.

Note that Weinberg provides an excellent historical discussion of the interplay between predictions based upon the Hot Big Bang (HBB = CMB + BBN) model and (particularly) CMB observers. In contrast to the Wagoner et al. mode,. the first paper to consider the production of universal nuclear elements was Alpher, Bethe and Gamow (1948). They considered a universe where baryons were dominated by neutrons. In their model nuclear elements were created via successive neutron decay and capture events to create all known universal elements. Though inaccurate in many respects this paper predicted the presence of a few Kelvin radiation background and was the first consideration of a HBB model.

4.6 Observations of 'primordial' light elements

Current observational techniques employed to determine the abundances of primordial nuclear elements have to overcome two difficulties, 1) that the elements really are 'primordial' and have not been created/destroyed by secondary events and , 2) the tremendous dynamic range covered by the predicted abundances. Techniques for each element are summarised below.

1. **He**^{4} – is observed in Galactic and extra–galactic HII (star–forming) regions. Primordial interstellar gas is ionised by energetic photons emitted from young stars. The gas then cools via a number of strong emission lines (e.g. HeII). The measurement of Y_p involves three steps.

   (a) Emission line flux ratios must be measured to high accuracy, which requires good detector linearity and flux calibration, and corrections for reddening and stellar He I absorption.

(b) These fluxes must be converted to an abundance, which requires correction for collisional ionization and neutral He. The correction for unseen neutral He depends on the spectral energy distribution adopted for the ionizing radiation and might change Y by 1 - 2 percent.

(c) Then the primordial abundance must be deduced from the Y values in different galaxies. He abundances are plotted against a heavy element 'marker', e.g. O or N. By plotting Y versus O,N one may extrapolate the relation to zero metals to obtain Y_p, the predicted primordial He abundance. Note – zero metals provides the best approximation to zero star–formation. He is created in stars.

2. **D** – is observed in Galactic and extra–galactic Hydrogen clouds. Galactic clouds are problematic because they are contaminated by heavy elements – indicative of star–formation and D is always destroyed in stellar interiors (i.e. they are sufficiently dense fusion reactors that all D is processed into heavier elements). Observations of Hydrogen in distant gas clouds back lit by QSOs provides a probe of extremely low metallicity environments. D is observed as a weak absorption doublet of Hydrogen with a characteristic velocity offset of 82 kms^{-1} in the absorption rest–frame. However, a considerable drawback is that only three reliable D systems have been identified and studied.

3. **He^3** – The primordial abundance of He^3 has not been measured. This is most unfortunate, since it is nearly as sensitive as D to the baryon density during BBN. He^3 is harder to measure than D because the difference in wavelength of He^3 and He^4 lines is smaller than for D, and the Lyman series lines of He II, the main absorption lines of He in the IGM, are in the far ultraviolet at 228 - 304Å which is hard to observe because of absorption in the Lyman continuum of HI at $\lambda < 912$Å.

4. **Li^7** – is observed in stellar atmospheres. However, accurate estimation of primordial abundances requires low metallicity stars and a good understanding of element production and distribution (dredge–up) rates in stellar interiors. In addition, Li^7 is produced in cosmic ray collisions with heavier elements.

When observed primordial abundance estimates are compared to the predicted values as a function of baryon density, only a poor agreement at best is achieved (and that by inclusion of uncertain estimates of systematic errors). Clearly, the chain of assumptions required to compare observed and predicted abundances must be improved. Abundance predictions based upon BBN my be compared to the baryon density estimated from observations the CMB. This comparison further supports the conclusion that current BBN techniques require improvement to be considered a 'precision' test of the HBB model, e.g.

1. Reduce current uncertainties on nuclear interaction rates employed in the abundance computation.

2. Observe more primordial abundance traces in order to highlight sources of intrinsic variation.
3. Improved models of stellar nuclear yields and their subsequent distribution thought the star.