Theoretical cosmology II (NTMF333)

Lectures 1–6: Hot early Universe

Michal Malinský

Institute of Particle and Nuclear Physics, Charles University

Version 3 as of May 25, 2023

Contents

1 The starting point **3**

2 The hot early Universe **3**
2.1 The early Universe thermal bath 3
2.2 Equilibrium thermodynamics & thermal distributions' recap 4
2.3 Few basic conceptual remarks on the chemical potentials 5
2.4 Neglecting chemical potentials for relativistic species 5
2.5 Thermal equilibrium in the expanding Universe 6

3 The fate of thermal relics **7**
3.1 Formalizing these thoughts - Boltzmann's transport equation 7
3.1.1 Boltzmann's equation without the RHS interaction terms 7
3.1.2 Boltzmann's equation with interactions 9
3.1.3 Hot versus cold relics 11
3.2 Hot relics 12
3.2.1 Neutrinos as 'canonical' hot relics 12
3.2.2 Intermezzo: Can neutrinos be the DM? 12
3.3 Cold relics 13
3.3.1 General formulae for pair annihilation 14
3.3.2 The WIMP miracle 15
3.3.3 Popular CDM particle-physics candidates 16
3.4 Relics with stronger than weak-interaction couplings to the bath 17
3.4.1 Relics with electromagnetic interactions 18
3.4.2 Relics with strong interactions - baryons 19

4 Thermal history of the early Universe **20**
4.1 Modes of survival 20
4.2 Time scales 21

5 Nucleosynthesis **23**
5.1 The deuterium bottleneck 24
5.2 The pre-deuteron nucleon chemistry 26
5.2.1 A remarkable coincidence 27
5.3 The notorious 25% of primordial Helium 28
5.3.1 The ${}^{2}_{1}\mathrm{D}$ creation efficiency at the onset of nucleosynthesis 30
5.4 Relic abundances of other (light) elements and measurement of η_B 31

6 Recombination **33**
6.1 The fraction of free electrons and Saha equation 34
6.1.1 Solving Saha equation 34
6.1.2 Recombination temperature 36
6.2 Photon decoupling 36
6.2.1 Did photon decoupling occur before, at or after recombination? . . . 36
6.2.2 Photon decoupling temperature 37

7 Some very sketchy conclusions **38**

1 The starting point

With David, you have ended up with a conclusion there should be a dark matter component in the Universe with

$$\Omega_{DM}h^2 \approx 0.12 \tag{1}$$

where

$$\Omega_{DM} \equiv \frac{\rho_{DM}}{\rho_{crit.}} \tag{2}$$

provided

$$\rho_{crit.} = \frac{3H^2}{8\pi G} = \frac{3h^2 \times (100\,\mathrm{km.s^{-1}Mpc^{-1}})^2}{8\pi G} \tag{3}$$

which yields about

$$\Omega_{DM}h^2 \approx \rho_{DM} \times 10^{-4}\,\mathrm{cm}^3/\mathrm{eV} \tag{4}$$

so the observation above tells us something about ρ_{DM}. But what can we really learn about DM from this number? Not much until we assume something:

- In what follows, we shall assume DM is particle-like
- We shall also assume that it is a thermal relic

2 The hot early Universe

By this we mean populations of particles that throughout the history have been at least for a while described by a thermal distribution, i.e., they were in TE with something. Asserting this, we should answer two questions:

1. What was that something the DM was in TE with?
2. Is the thermality assumption plausible, i.e., does it make sense in the FLRW context?

2.1 The early Universe thermal bath

As for 1. above, the role of the bath was (and still is) played by photons - they are very abundant, massless = always relativistic, always Planck distributed (even after their decoupling from matter at some 3×10^5 years after the BB!) - this is seen from the CMB which is (up to 10^{-5} relative fluctuations) perfectly thermal with $T_\gamma \sim 2.7$K.
NB. For massless relativistic species the phase-space distribution remains Planckian even after decoupling because the redshift effects corresponding to $p \to xp$ in

$$f(p) = \frac{1}{e^{cp/T} - 1} \tag{5}$$

are easily compensated by $T \to Tx$ and the shape of the distribution does not change.[1]

[1] Note that this is not the case for the "massive" versions of Fermi or Bose-Einstein where the presence of mass (let alone chemical potentials) spoils the argument:

$$f(p) = \frac{1}{e^{(\sqrt{p^2+m^2}-\mu)/T} \pm 1} \tag{6}$$

and the only way to maintain equilibrium-like distribution is changing the chemical potential μ.

In conclusion, photons for sure provide a good bath so we shall use T_γ as a fundamental parameter in distributions of all species that are in contact with photons. NB there is at least one type of particles that is not (and, remarkably, that has its own temperature - neutrinos; they are actually colder (i.e., they do have a temperature T_ν) than photons because they decoupled way earlier - we shall see that - and were not "reheated" ever after).

2.2 Equilibrium thermodynamics & thermal distributions' recap

Assuming homogeneity and isotropy of the system the equilibrium state is described by Fermi-Dirac or Bose-Einstein phase-space distributions

$$f(\vec{p})^{FD}_{BE} = \frac{1}{e^{(E-\mu)/k_BT} \pm 1} \tag{7}$$

which correspond to the density of states around $\vec{p}$ with $E = \sqrt{\vec{p}^2 + m^2c^4}$ at temperature T. Given this (and employing "natural units"[2]) with $k_B = \hbar = c = 1$, the spatial number density (i.e., the number of states per unit volume) can be calculated as

$$n = \frac{g}{(2\pi)^3}\int f(\vec{p})\mathrm{d}^3p = \frac{g}{2\pi^2}\int_0^\infty p^2 f(p)\mathrm{d}p = \frac{g}{2\pi^2}\int_m^\infty E\sqrt{E^2-m^2}f(E)\mathrm{d}E \tag{8}$$

where the degeneracy factor g counts the number of different degrees of freedom distributed by f (typically, the number of spin states) and the $(2\pi)^3$ factor in the denominator provides the correct normalization when the summation over the quantized values of $\vec{p}$ within a box of length L and, thus, volume $V = L^3$ (divided into elementary cells with size $2\pi\hbar/L$) is replaced by integration over continuum. Similarly, the corresponding energy density is given by

$$\rho = \frac{g}{(2\pi)^3}\int E(\vec{p})f(\vec{p})\mathrm{d}^3p = \frac{g}{2\pi^2}\int_0^\infty p^2 E(p)f(p)\mathrm{d}p = \frac{g}{2\pi^2}\int_m^\infty E^2\sqrt{E^2-m^2}f(E)\mathrm{d}E \tag{9}$$

In the ultrarelativistic ($p \gg m$) and non-relativistic ($p \ll m$) regimes these integrations are relatively simple and the results are (neglecting chemical potentials for the relativistic species, see Sect. 2.4):

Relativistic:

$$\rho = \frac{\pi^2}{30}gT^4 \quad \text{(bosons)}, \qquad \rho = \frac{7}{8}\times\frac{\pi^2}{30}gT^4 \quad \text{(fermions)} \tag{10}$$

$$n = \frac{\zeta(3)}{\pi^2}gT^3 \quad \text{(bosons)}, \qquad n = \frac{3}{4}\times\frac{\zeta(3)}{\pi^2}gT^3 \quad \text{(fermions)} \tag{11}$$

Non-relativistic:

$$\rho = m.n \tag{12}$$

$$n = g\left(\frac{mT}{2\pi}\right)^{3/2} e^{-(m-\mu)/T} \quad \text{(Boltzmann)} \tag{13}$$

Let us emphasize that the formulae for ρ do not tell us much about the DM yet because nowadays DM is no longer thermally distributed and one has to do more.

[2]Sometimes, we may accidentally slip to even adopting "supernatural units" with $k_B = \hbar = c = 1 = -1 = 2 = \pi = i$ for which the author would like to apologise in advance (and as for the assistance should you spot something like that throughout the text).

2.3 Few basic conceptual remarks on chemical potentials

Their definition

$$dU = -pdV + TdS + \mu dN \tag{14}$$

is an equation that tells us about different ways the internal energy of a system can transform, namely, into

1. $-pdV$ = Expansion - we shall not entertain fast expansion (on the interaction rate scales), see below.
2. TdS = Irreversible changes where $dS > 0$
3. μdN = Chemical energy

For isolated systems we have energy conservation, i.e., $dU = 0$.

- If we neglect expansion/contraction then for reversible processes we have $dS = 0$ and, thus $\mu dN = 0$. This means that for a reaction of the type

$$1 + 2 \leftrightarrow 3 + 4 \tag{15}$$

  one has (going, for instance, from left to right[3]) $-(\mu_1 + \mu_2) + (\mu_3 + \mu_4) = 0$ and, thus

$$\mu_1 + \mu_2 = \mu_3 + \mu_4 \tag{16}$$

- For irreversible processes we have $dS > 0$ and, thus $\mu dN < 0$. This means that for the same reaction one has, e.g., preference of going in the $1 + 2 \to 3 + 4$ direction iff $-(\mu_1 + \mu_2) + (\mu_3 + \mu_4) < 0$ and, thus

$$\mu_1 + \mu_2 > \mu_3 + \mu_4 \tag{17}$$

Hence, the name of chemical potential - the bigger the more likely the reaction goes towards its minimization.

2.4 Neglecting chemical potentials for relativistic species

One more remark is worth here: in what follows we shall almost always neglect chemical potentials in the relativistic distributions because (note that due to $\mu_\gamma = 0$ one has $\mu_X = -\mu_{\overline{X}}$)

$$n_X - n_{\overline{X}} = \frac{g}{2\pi^2}\int_0^\infty dp \left(\frac{p^2}{e^{(pc-\mu)/T}+1} - \frac{p^2}{e^{(pc+\mu)/T}+1}\right) = \frac{g}{6}T^3\left(\frac{\mu}{T} + \frac{1}{\pi^2}\frac{\mu^3}{T^3}\right) \tag{18}$$

where the defining formula (8) has been used. Hence,

$$\frac{n_X - n_{\overline{X}}}{n_\gamma} \propto \left(\frac{\mu}{T} + \frac{1}{\pi^2}\frac{\mu^3}{T^3}\right) \tag{19}$$

which makes it possible to conclude that for all species that are not "significantly asymmetric" with respect to photons μ/T must be small!

[3]NB The same is obtained going from right to left.

Note that this is not the case for non-relativistic species where one has

$$n_X = g\left(\frac{m_X T}{2\pi}\right)^{\frac{3}{2}} e^{-(m_X-\mu)/T} \tag{20}$$

(and similarly for antispecies) which yields instead

$$\frac{n_X - n_{\overline{X}}}{n_\gamma} \propto \left(\frac{m_X}{T}\right)^{\frac{3}{2}} e^{-m_X/T} \sinh\left(\frac{\mu}{T}\right) \tag{21}$$

so one can have large μ/T even for negligible asymmetries (due to the further exponential suppression provided by the Boltzmann factor on the RHS).

2.5 Thermal equilibrium in the expanding Universe

Strictly speaking, expansion = cooling breaks the TE picture as in TE the time should not play any role. So, what we mean is a sequence of states very close to the TE, i.e., populations described by equilibrium distributions even though $T(t)$ is slowly changing[4]. Again, this should be clarified:

1. The change in $T(t)$ should be in some sense slow... Slow with respect to what? Well, with respect to the typical interaction rates maintaining the approximate TE - this is a particle physics business; we'll come to that.

2. This picture is questionable at the very beginning, i.e., during inflation (if there were any particles at all), but rather likely later on.

The last point is worth a brief verification: From the second Friedmann's equation one has (for $k = 0$ and $\Lambda = 0$)

$$H^2 = \frac{8\pi G}{3}\rho \tag{22}$$

which in the case of our interest, i.e., RD Universe, corresponds to[5]

$$H^2 = \frac{8\pi G}{3}\rho_{REL} = \frac{8\pi G}{3}\frac{\pi^2}{30}g_* T^4 = \frac{8\pi^3}{90}g_* G T^4 \tag{23}$$

and thus

$$H = \sqrt{\frac{8\pi^3}{90}g_*}\sqrt{G}T^2 \approx 1.66 \times \sqrt{g_*}\frac{T^2}{M_{Pl}} \tag{24}$$

where the Planck scale $M_{Pl} \sim 1.2 \times 10^{19}$GeV is defined as

$$M_{Pl} = \sqrt{\frac{\hbar\, c}{G}} \tag{25}$$

[4]The explicit form of the $T(t)$ dependence in the RD Universe shall be discussed in detail in Sect. 4.2; here we shall be just fine with parametrising the evolution with T.

[5]We use the equilibrium energy density for relativistic species with $\mu = 0$ in the form

$$\rho_{REL} = \frac{\pi^2}{30}g_* T^4$$

where g_* "counts" the species as $g_* = \sum_B g_B + \frac{7}{8}\sum_F g_F$ with $g_{B,F}$ corresponding to degeneracy factors, i.e., spin degrees of freedom etc, see e.g. [1, 2].

so we can see that $H(T) \propto T^2$. Hence, the expansion (Hubble) rate is slowing down in the expanding (and, hence, cooling) FLRW RD Universe as a second power of its temperature.

Second, the interaction probabilities of two-to-two processes that we shall be mostly dealing with come in the form of the relevant scattering cross-sections σ (which, barring weak interactions[6], are often relatively mildly energy-dependent[7]) multiplied by the number densities of projectiles and targets (each $\propto T^3$ for relativistic species) and, thus, scale (typically) like $(T^3)^2$, i.e., feature a very strong temperature dependence.

As we shall see later, the proper quantity H should be really compared to is actually the interaction rate per unit target particle density Γ which, by the same argument, should typically go like T^3 (for relativistic species), i.e., with a higher power of T than the square in $H(T)$. Hence it is rather likely that at some point in time in the early Universe the rates of interactions were large enough to thermalise populations of particles which in one way or another do couple to the thermal bath.

3 The fate of thermal relics

Ok, back to the thermal relics. The $\Gamma \propto T^3$ (or so) vs. $H \propto T^2$ makes it not only clear that thermal fate is quite likely (if interactions are non-negligible) but also that with diminishing T over time the interaction rates may cease being fast enough to maintain TE. At some point (T_{FO}), freeze-out occurs and, since then, the population will practically not interact and will be subject to essentially two remaining processes, namely, dilution due to cosmological expansion and (perhaps) decays.

From this perspective, the number quoted at the beginning should correspond to the TE number density of DM at the T_{FO} diluted by the cosmic evolution. But how to calculate T_{FO} in practice?

3.1 Formalizing these thoughts - Boltzmann's transport equation

What we just said can be written in terms of the time derivative of the number density $n_X(t)$ of species X undergoing expansion and interactions.

3.1.1 Boltzmann's equation without the RHS interaction terms

Remarkably enough, for TE the equilibrium $n_X(t)$'s obey an equation that looks like if there were no interactions at all, just dilution:

$$\frac{d}{dt}n_X + 3Hn_X = 0 \tag{26}$$

Several remarks are needed here:

- First, one should check that the relativistic distributions really obey this: Indeed, from the first Friedmann's equation

$$\dot{\rho} = -4H\rho \tag{27}$$

[6]Note that weak interactions, typically, provide inelastic processes which seldom correspond to those responsible for maintaining kinetic equilibrium; the latter is, instead, typically provided by electromagnetic and strong interactions.

[7]This applies namely to the s-channel processes whilst those passing through the t and/or u channels can have the kinematics rather different; however, what we are mainly interested here is the dominant contribution in the scattering cross-section.

one has (plugging in, e.g., $\rho_B(T) = \frac{\pi^2}{30} g_* T^4$)

$$\frac{\pi^2}{30} g_* 4T^3 \frac{dT}{dt} = -4H \frac{\pi^2}{30} g_* T^4 \tag{28}$$

and, thus,

$$\frac{d}{dt} = \frac{d}{dT}\frac{dT}{dt} = -HT\frac{d}{dT} \tag{29}$$

Combining this with, e.g., $n(T) = \frac{\zeta(3)}{\pi^2} g T^3$ (for bosons) one obtains

$$\frac{d}{dt} n = -HT \frac{\zeta(3)}{\pi^2} g \frac{d}{dT} T^3 = -3Hn \tag{30}$$

which was to be shown.

- Second, one should show that this is indeed equation for dilution. Indeed,

$$0 = \frac{d}{dt} n + 3Hn = \frac{d}{dt} n + 3\left(\frac{\dot{a}}{a}\right) n \tag{31}$$

which is equivalent to

$$\left(\frac{d}{dt} n\right) a^3 + 3\left(\frac{d}{dt} a\right) a^2 n = \frac{d}{dt}\left(n a^3\right) = 0 \tag{32}$$

so one has a constant *comoving volume* number density, i.e., like in the pure dilution case (practically like dust).

Lecture II: Boltzmann's equation with interactions, types of relics

Note that this intuitive picture is perfectly compatible with the transport equation you have encountered before in the statistical physics. There the evolution of the phase space distribution $f(\vec{x}, \vec{p}, t)$ was described by means of Liouville equation of the form

$$L[f] = C[f] \tag{33}$$

where the LHS corresponds to the *total* time change of the number density of particles at $\vec{x}$ with $\vec{p}$ represented by the Liouville operator

$$L[f] = \frac{d}{dt} f(\vec{x}, \vec{p}, t) = \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \vec{x}}\frac{d\vec{x}}{dt} + \frac{\partial}{\partial \vec{p}}\frac{d\vec{p}}{dt}\right) f(\vec{x}, \vec{p}, t)\,. \tag{34}$$

This, in case of no collisions (zero RHS) represents nothing but a continuity equation, i.e., conservation of the total number of particles. Indeed, the particles in the small volume around $(\vec{x}, \vec{p})$ at time t make it after Δt to the volume around point $(\vec{x} + \vec{v}\Delta t, \vec{p} + \vec{F}\Delta t)$ (where $\vec{F} = d\vec{p}/dt$ is an external force) and, hence,

$$f(\vec{x} + \vec{v}\Delta t, \vec{p} + \vec{F}\Delta t, t + \Delta t) - f(\vec{x}, \vec{p}, t) = 0 \tag{35}$$

and the rest (i.e., the differential form of the Liouville operator in (34)) follows from the expansion in Δt. Note also that the covariant[8] transcription of L reads

$$L = \frac{1}{E}\left(p^\alpha \frac{\partial}{\partial x^\alpha} - \Gamma^\mu_{\alpha\beta} p^\alpha p^\beta \frac{\partial}{\partial p^\mu}\right)\,; \tag{36}$$

working this out for the FLRW metrics one ends up with the free form of the Boltzmann's equation (31).

[8] Note that L written this way is covariant up to the overall $\frac{1}{E}$ factor which, however, is expected because it corresponds to a time derivative and, as such, it is not a scalar operator.

3.1.2 Boltzmann's equation with interactions

Turning on the collision term in (33) the comoving number densities of species are no longer necessarily conserved due to the presence of inelastic interactions; roughly speaking, n_X will be subject not only to the cosmological expansion driven by H but also to collisions and decays that can further change it. Let us, for simplicity, assume that X undergoes predominantly one type of inelastic processes, namely

$$XY \leftrightarrow AB \tag{37}$$

where A and B are some other species in the bath. These interactions, among other things, are characterized by their cross-sections $\sigma_{XY\to AB}$ and $\sigma_{AB\to XY}$, respectively. As you should know, the physical meaning of these quantities is the *probability of X entering such an interaction if one Y passes through a unit area around the target.* Suppose that Y's are characterized by a volume number density n_Y and that they fly towards one X in a unit volume with velocity v. Then, the number of Y's which in a unit of time pass through the unit area around X is $v.n_Y$ and the number of interactions they inflict over a unit of time is $\sigma_{XY\to AB}.v.n_Y$. For more than a single X in the unit volume the number of X's undergoing such a change in throughout a unit of time is $\sigma_{XY\to AB}.|v|.n_Y.n_X$ and, hence, one should expect a structure like this playing role of the collision term on the RHS of (33).

This, indeed, is the case with one small easy to understand modification, namely, the *statistical* averaging of $\sigma_{XY\to AB}.|v|$ due to the thermal environment this all occurs in which shall be denoted by $\langle\sigma_{XY\to AB}.|v|\rangle$. Thus, for the process at stakes (37) the Boltzmann's equation with collisions reads

$$\frac{d}{dt}n_X + 3Hn_X = -\langle\sigma_{XY\to AB}.|v|\rangle n_X n_Y + \langle\sigma_{AB\to XY}.|v|\rangle n_A n_B \tag{38}$$

Several comments are worth here:

1. First, note that, in principle, in deriving this equation one does not need to make any strong assumptions about the thermal or non-thermal nature of the species involved - neither the effects of cosmological dilution (the H on the LHS) nor the collision picture really depended on whether they were thermally distributed or not. The only small trace of the need to understand something about that is the averaging procedure $\langle\rangle$; that's why we used the term "statistical" rather than "thermal" averaging above. Nevertheless, in reality, for the sake of T_{FO} calculation we shall assume that none of the species involved ever makes it far from thermal equilibrium – that's indeed how T_{FO} is even defined: it is the onset of significant departure from TE.

2. Note that Eq. (38) is perfectly compatible with both the limiting regimes with zero RHS (TE regime, freeze-out regime) we were talking about before and, in fact, complements them only in the transition period between the two. In this respect, it constitutes a complete description of all three modes of evolution the population of X can go through. Assuming without loss of generality that $m_{X,Y} \gg m_{A,B}$ (and, for simplicity, that[9] $Y = \bar{X}$ and $B = \bar{A}$, i.e., that the processes under consideration are the annihilation and pair creation of $X\bar{X}$) one has:

   - Very busy, perfectly thermal mode with $T \gg m_X$ where destruction of X and its creation are perfectly balanced. In this case, $n_X = n_X^{EQ}$ and we already know that n_X^{EQ} obeys the homogeneous form of Eq.(38).

[9] The bar denotes antiparticles.

- With dropping T the $n_{X,\bar{X}}$ and also $n_{A,\bar{A}}$ go down; for $T \sim m_X$, however, it will become somewhat complicated for the bath to re-create $X\bar{X}$ pairs and the ballance in $X\bar{X} \leftrightarrow A\bar{A}$ will be shifted towards $\rightarrow$, i.e., destruction of $X\bar{X}$ (due to annihilation but in principle also by other processes like decays or other inelastic reactions) prevails.
- At some point, with $n_{X,\bar{X}}$ going even lower even the destruction eventually stops because X will not be able to find a partner $\bar{X}$ to co-annihilate. The RHS of (38) is again zero and, since then, the population is further diluted only due to the H-term on the LHS.

  As an extreme example of these principles one can take the fate of the neutrons right after they popped up in the early Universe thermal bath just after T dropped below about 100 MeV. In such conditions the neutron population was subject to many types of processes among which the prominent role was played by 1) their annihilation into pions (unbalanced by their very seldom recreation by inverse processes because $T \ll m_n$), 2) the "chemical" reactions due to weak interactions in which they were converted into protons and vice versa (these were still fast because electrons and neutrinos were very abundant at these temperatures) and, 3) their three-body decay $n \rightarrow p^+ e^- \bar{\nu}$. The RHS of Eq.(38) would then read

$$\begin{aligned} &-\langle\sigma_{n\bar{n}\rightarrow\pi\pi}.|v|\rangle n_n^2 + \langle\sigma_{\pi\pi\rightarrow n\bar{n}}.|v|\rangle n_\pi^2 \\ &-\langle\sigma_{ne^+\rightarrow p\bar{\nu}}.|v|\rangle n_n n_e + \langle\sigma_{p\bar{\nu}\rightarrow n\bar{n}}.|v|\rangle n_p n_\nu - \langle\sigma_{n\nu\rightarrow pe}.|v|\rangle n_n n_\nu + \langle\sigma_{pe\rightarrow n\bar{\nu}}.|v|\rangle n_p n_e \\ &-\Gamma_{n\rightarrow pe\bar{\nu}}\, n_n \end{aligned} \tag{39}$$

  where the first line corresponds to the "universal" relativistic annihilation and creation all particles can undergo if there is enough energy (note that from the chemical potentials' perspective do not play much of a role here because for pair annihilations and creation they cancel to a large degree on both sides), the second line represents the "chemical processes" which are species-specific and occur even in the non-relativistic regime (and, as we shall see, chemical potentials provide non-trivial information about these) and, finally, the third line is the (slow) decay process.

  A picture of the evolution of n_n with RHS as in (39) would be worth here - later.
- In several cases of our interest (i.e., DM disappearance from the early Universe – note that DM by definition does not entertain much of a chemistry or decay) the relativistic pair processes will (or will be assumed to) dominate throughout the entire thermal epoch of a population which will eventually annihilate into species (say $A\bar{A}$) that will be still strongly coupled to the bath, i.e., they will still be distributed thermally. In such a case one has

$$\frac{d}{dt}n_X + 3Hn_X = -\langle\sigma_{X\bar{X}\rightarrow A\bar{A}}.|v|\rangle n_X n_{\bar{X}} + \langle\sigma_{A\bar{A}\rightarrow X\bar{X}}.|v|\rangle n_A^{EQ} n_{\bar{A}}^{EQ} \tag{40}$$

  and the last term on the RHS can be translated into something way more favourable by noting that

$$\langle\sigma_{A\bar{A}\rightarrow X\bar{X}}.|v|\rangle n_A^{EQ} n_{\bar{A}}^{EQ} = \langle\sigma_{X\bar{X}\rightarrow A\bar{A}}.|v|\rangle n_X^{EQ} n_{\bar{X}}^{EQ} \tag{41}$$

Thus, one can finally rewrite (40) into the form

$$\frac{d}{dt}n_X + 3Hn_X \approx \langle\sigma_{X\bar{X}\to A\bar{A}}.|v|\rangle \left(n_X^{EQ}n_{\bar{X}}^{EQ} - n_X n_{\bar{X}}\right) \tag{42}$$

where we have neglected small variations of $\langle\sigma_{X\bar{X}\to A\bar{A}}.|v|\rangle$ for n's close to n^{EQ}'s.

Now, the physics of (43) is very clear. Swapping the dilution term to the RHS too one arrives at

$$\frac{d}{dt}n_X \approx \langle\sigma_{X\bar{X}\to A\bar{A}}.|v|\rangle \left(n_X^{EQ}n_{\bar{X}}^{EQ} - n_X n_{\bar{X}}\right) - 3Hn_X \tag{43}$$

which, among other things, makes it clear that:

1. In the regime where interactions dominate over the dilution one is essentially left with just the first bracket on the RHS which essentially forces n_X to follow n_X^{EQ}: for any positive upper fluctuations of $n_X n_{\bar{X}}$ over $n_X^{EQ}n_{\bar{X}}^{EQ}$ the RHS is negative and the system is pushed back towards n_X^{EQ}; vice versa for the negative fluctuations.

2. Since interactions are quadratic in n's while dilution is only linear sooner or later the dilution term starts kicking in and will have a tendency to draw n_X away from n_X^{EQ}. The freeze-out moment then corresponds to the situation when interactions cease power to bring not-so-small fluctuations back to zero, i.e., when

$$\langle\sigma_{X\bar{X}\to A\bar{A}}.|v|\rangle \left((n_X^{EQ})^2 - n_X^2\right) \sim \langle\sigma_{X\bar{X}\to A\bar{A}}.|v|\rangle 2n_X^{EQ}(n_X^{EQ} - n_X) \tag{44}$$

become comparable to $3Hn_X^{EQ}$, which means that H becomes comparable to

$$\frac{n_X^{EQ} - n_X}{n_X^{EQ}}\langle\sigma_{X\bar{X}\to A\bar{A}}.|v|\rangle n_X^{EQ} \tag{45}$$

for $O(1)$ fluctuations (the $n_X^{EQ} - n_X/n_X^{EQ}$) term above. Defining

$$\Gamma \equiv \langle\sigma_{X\bar{X}\to A\bar{A}}.|v|\rangle n_X^{EQ} \tag{46}$$

the annihilation freeze-out condition thus reads

$$\boxed{\Gamma \sim H} \tag{47}$$

Once again, if $\Gamma > H$ the system is able to compensate even for (local) $O(1)$ fluctuations and bring n_X back to n_X^{EQ} while for $\Gamma < H$ the interactions are diminishing and can be practically neglected. In practice, however, the transport equation should be solved in full glory in order to get the numbers really right.

3.1.3 Hot versus cold relics

With (47) at hand one can determine T_{FO} by substituting T^2/M_{Pl} to the RHS and by inputting the thermal n_X together with the information on σ from your favourite particle physics model(s) to the LHS and taking that as an equation for T.

In general, there are two basic classes of relics which differ by whether the population was still relativistic upon freeze-out (a hot relic) or not (a cold relic). This seems trivial but it is not because an easy solution for T_{FO} from (47) is available only in the former case.

3.2 Hot relics

3.2.1 Neutrinos as 'canonical' hot relics

As a vanilla example of a hot relic let's consider the freeze-out of neutrinos which, by the way, look as a good SM candidate for the DM that we are after. What is T_{FO} for the neutrino interactions in the early Universe?

The most efficient inelastic neutrino interactions in the early Universe are their weak annihilations and scattering on other leptons (which, for sufficient energies, were as abundant as neutrinos were) with a typical cross-sections of the order of (using for simplicity the old Fermi contact interaction theory[10])

$$\sigma \sim G_F^2 E^2 \tag{48}$$

which essentially follows from the dimensionality arguments. Since the relativistic leptonic target's densities are $\propto T^3$ and $|v| = 1$ in such conditions one has

$$\Gamma \sim G_F^2 T^5 \tag{49}$$

which is comparable to $H \sim T^2/M_{Pl}$ for

$$T_{FO}^{\nu} \sim \sqrt[3]{\frac{1}{G_F^2 M_{Pl}}} \sim 1\text{MeV} \tag{50}$$

So, indeed, the neutrino interactions freeze-out in the regime where ν's are highly relativistic and, thus, they are *hot relics*.

3.2.2 Intermezzo: Can neutrinos be the DM?

With this at hand, we can ask whether neutrinos can play the role of DM that we see, i.e., whether their today's neutrino number density can correspond to the measured $\Omega_{DM}h^2 \approx 0.12$ or not. The answer, however, is negative, for at least two reasons:

1. Cowsik-McClelland bound (PRL 29 (1972))
   At T_{FO}^{ν} the neutrino number density is given by

   $$n_{\nu}^{EQ}(T_{FO}^{\nu}) = \frac{3}{4}\frac{\zeta(3)}{\pi^2} g {T_{FO}^{\nu}}^3 \tag{51}$$

   From that moment on they undergo just the cosmological dilution with constant na^3, and thus their today's number density should be

   $$n_{\nu}^{now} = n_{\nu}^{EQ}(T_{FO}^{\nu}) \left(\frac{a(T_{FO})}{a_0}\right)^3 = n_{\nu}^{EQ}(T_{FO}^{\nu}) \left(\frac{T_{\nu}^{now}}{T_{FO}^{\nu}}\right)^3 = \frac{3}{4}\frac{\zeta(3)}{\pi^2} g (T_{\nu}^{now})^3 \tag{52}$$

   Assuming neutrinos are non-relativistic now (which should indeed be the case for at least two of their mass eigenstates) one can write (for today's energy density stored in relic neutrinos):

   $$\rho_{\nu} = m_{\nu}\frac{3}{4}\frac{\zeta(3)}{\pi^2} g (T_{\nu}^{now})^3 \tag{53}$$

[10] As we shall see this is a good approximation indeed.

Substituting this into formula (4) one obtains

$$\Omega_\nu h_0^2 = 0.95 \times 10^{-4} m_\nu \frac{3}{4} \frac{\zeta(3)}{\pi^2} g (T_\nu^{now})^3 \mathrm{cm}^3/\mathrm{eV} \tag{54}$$

which for $g = 2$ (assuming two-component neutrinos) and $T_\nu^{now} \sim 2$ eV yields

$$\Omega_\nu h_0^2 \approx 10^{-2} \frac{m_\nu}{\mathrm{eV}} \,. \tag{55}$$

Thus, in order for the neutrinos to amount for the measured $\Omega_\nu h_0^2 \approx 0.12$ one would have to have $\sum m_\nu > 10$ eV which is very hard to reconcile with the existing beta decay limits etc. Note that historically this was cast as one of the first upper limits on the neutrino masses (from the requirement that neutrinos do not overclose the Universe) by Cowsik and McClelland some 40 years ago.

2. After the decoupling at $T \sim$1 MeV the hot neutrinos free stream and, as such, they tend to erase all seeds of structure which nowadays corresponds to the supercluster size. The argument goes as follows:

   - Once free streaming, they can not clump and thus (assuming they dominate the DM) they will not seed the potential wells the baryons would start falling into until they get non-relativistic.
   - This happens at $T \sim m_\nu$; thus, the area they would erase is roughly of the size of the horizon R_H at this moment. Since, however, $R_H \sim H^{-1}$ and[11] $H \sim T^2/M_{Pl}$ one gets $R_H \sim M_{Pl}/T^2$ which, at the relativistic to non-relativistic transition is $R_{\mathrm{erased}} \sim M_{Pl}/m_\nu^2$.
   - If one translates it further to the mass within such a radius (dominated again by neutrinos) one has

     $$M_{\mathrm{erased}} \sim R_{\mathrm{erased}}^3 \times m_\nu \times n_\nu (T \sim m_\nu) \sim M_{Pl}^3/m_\nu^2 \sim 10^{16} M_\odot (m_\nu/\mathrm{eV})^{-2} \tag{56}$$

     (with the solar mass $M_\odot \approx 2 \times 10^{30}$ kg), i.e., the supercluster mass.

The last remark touches upon the role T_{FO} plays in the predictions for today's $\Omega_\nu h_0^2$ - remarkably, for hot relics this quantity does not really depend on it because $n^{EQ}(T_{FO}^\nu)$ is cubic in T_{FO}^ν which is entirely undone by the third power of the scaling factor ratio.

Lecture III: Cold relics, WIMPs, other DM candidates, thermal history of the Universe

3.3 Cold relics

Thus, it is time to turn our attention to cold relics, i.e., to species for which the annihilation rates were large enough to dominate over H even when they got non-relativistic, i.e., in the regime when their pair re-creation from the bath essentially stopped. Hence, their number densities experienced a Boltzmann suppression (i.e., a downhill slide) for at least a while before dilution eventually made also their annihilations inefficient.

[11] This is a simplification because they may become non-relativistic below the RD→MD transition; however, as an order of magnitude it works because the RD→MD temperature is comparable to m_ν.

It is rather clear that, unlike for the (stable) hot relics like neutrinos or photons whose comoving number density is practically constant ever since they first got thermal (given just by their today's temperature), the number densities of cold relics (X) will depend rather strongly on the (ratio of their) mass m_X and T_{FO} – the larger it is (i.e., the longer they experience fast annihilations and/or other inelastic processes) the smaller their relic number densities.

However, determination of T_{FO} for cold relics is by far not as straightforward as it was for the hot ones. Indeed, if one tries to simply compare H and Γ there is an issue with the transcendental form of the resulting equation:

$$H \sim \Gamma \Leftrightarrow \frac{T^2}{M_{Pl}} \sim \langle \sigma_{X\bar{X}\to A\bar{A}}.|v| \rangle n_X^{EQ} \tag{57}$$

provided

$$n_X = g \left(\frac{m_X T}{2\pi} \right)^{\frac{3}{2}} e^{-(E-\mu)/T} \tag{58}$$

even if $\langle \sigma_{X\bar{X}\to A\bar{A}}.|v| \rangle$ is a polynomial in T (which, on the other hand is often the case, at least at the leading order).

3.3.1 General formulae for pair annihilation

In general, one has to solve (often numerically) the full system of the Boltzmann's equations. For the simplest case of pair annihilation this solution can be written in a relatively simple form [Kolb,Turner]:

$$x_{FO} \equiv \frac{m_X}{T_{FO}} \approx \log K - (n + \tfrac{1}{2}) \log(\log K) \tag{59}$$

with

$$K \approx 0.038(n+1) \frac{g}{\sqrt{g_*}} M_{Pl} m_X \sigma_0 \tag{60}$$

where n and σ_0 are defined via

$$\langle \sigma_{X\bar{X}\to A\bar{A}}.|v| \rangle \approx \sigma_0 \left(\frac{T}{m_X} \right)^n \tag{61}$$

for a constant σ_0, g is the degeneracy factor of the species X and

$$g_* = \sum_b g_b \left(\frac{T_b}{T_\gamma} \right)^4 + \frac{7}{8} \sum_f g_f \left(\frac{T_f}{T_\gamma} \right)^4 \tag{62}$$

counts the effective number of relativistic degrees of freedom which are in the bath at around T_{FO}, i.e., those that play role in $\rho_{REL} = \frac{\pi^2}{30} g_* T^4$. The relic number density (not comoving!) today is then given by

$$n_{X0} \approx 10^4 \frac{(n+1) x_{FO}^{n+1}}{g_{s*} \sqrt{g_*}} \times \frac{1}{m_X M_{Pl} \sigma_0} \mathrm{cm}^{-3} \tag{63}$$

where g_{s*} is the effective number of relativistic degrees of freedom entering the entropy density at around T_{FO} given by

$$g_{s*} = \sum_b g_b \left(\frac{T_b}{T_\gamma} \right)^3 + \frac{7}{8} \sum_f g_f \left(\frac{T_f}{T_\gamma} \right)^3 . \tag{64}$$

As arcane as these formulae look like they are relatively easy to justify:

1. First, note the relatively weak dependence of x_{FO} on the underlying physics – in most cases one can expect that this number will be somewhat bigger than 1 but not parametrically larger than $\mathcal{O}(1)$ and it will mainly depend on m_X and σ_0. This can also be phrased as a fact that cold relic species underwent a period of exponentially fast number density decay which, however, did not take too long - indeed, sooner or later their diminishing population was so diluted that their annihilations must have stopped.

2. Second, notice the inverse proportionality of n_{X0} to σ_0 (which can be interpreted as a mean value of the annihilation cross-section around the freeze-out point) – indeed, the bigger it is, i.e., the stronger the annihilations are the smaller is the resulting relic number density.

3. Last remark concerns a rather artificial situation in which, by accident, n_{X0} can be estimated analytically even for cold relics because one does no need to solve (57) for T_{FO}: For $n = -1$ one has for $T = T_{FO}$

$$\frac{T_{FO}^2}{M_{Pl}} \sim \langle \sigma_{X\bar{X}\to A\bar{A}} \cdot |v| \rangle n_X^{EQ}(T_{FO}) = \sigma_0 \left(\frac{T_{FO}}{m_X}\right)^{-1} n_X^{EQ}(T_{FO}) \tag{65}$$

and, hence,

$$n_X^{EQ}(T_{FO}) \sim \frac{T_{FO}^3}{\sigma_0 m_X M_{Pl}} \tag{66}$$

and, finally,

$$n_X(T_0) = n_X^{EQ}(T_{FO}) \left(\frac{T_0}{T_{FO}}\right)^3 \sim \frac{T_0^3}{\sigma_0 m_X M_{Pl}} \,. \tag{67}$$

For $T \sim 3\text{K} \sim 3 \times 10^4 \text{eV}$ which yields $T_0^3 \sim 10^3 \text{cm}^{-3}$ and one recovers the $\sim 10^3 \times \frac{1}{m_X M_{Pl} \sigma_0} \text{cm}^{-3}$ behaviour[12] of formula (63).

In what follows, we shall use these relations for the calculation of the basic characteristics of various cold relics including cold DM and also practically the entire spectrum of species in the Standard Model of particle interactions .

3.3.2 The WIMP miracle

Now the question is what can we learn from (63) about the properties of DM if we assume it is indeed a cold relic. In order to saturate the value of the fraction of the critical density in DM today, i.e.,

$$\Omega_{DM} h^2 \approx \rho_{DM} \times 10^5 \, \text{cm}^3/\text{GeV} \approx 0.12 \tag{68}$$

with the order of magnitude estimate (67)

$$\rho_{DM} = m_X n_{X0} \approx \frac{T_0^3}{\sigma_0 M_{Pl}} \tag{69}$$

one needs

$$\sigma_0 \approx \frac{10^5}{\Omega_{DM} h^2} \frac{T_0^3}{M_{Pl}} \text{cm}^3/\text{GeV} \tag{70}$$

[12]Note that, formally, formula (63) vanishes for $n \to -1$; however, this is due to the fact that it has been derived for the physical situations in which n never attains that value. The merit of the exercise with $n = -1$ is different, namely, to demonstrate that even though (63) fails one should still recover the universal behaviour with $\sigma_0 m_X M_{Pl}$ in the denominator of n_{X0}.

which numerically yields

$$\sigma_0 \sim 10^{-10}\text{GeV}^{-2}\,. \tag{71}$$

From the particle physics courses you should remember that this number is very special! Indeed, it is a typical value of the cross-sections due to weak interactions! For instance, the expected size for an annihilation ratio of fermions X with GeV-scale masses m_X due to contact interactions parametrized by the same Fermi constant $G_F \sim 10^{-5}\text{GeV}^{-2}$ as, for instance, in β-decays, the amplitude of such a process shall be $\mathcal{M} \sim G_F$ and, thus, for the cross-section one gets

$$\sigma \sim |\mathcal{M}|^2 m_X^2 \sim G_F^2 m_X^2 \sim 10^{-10}\text{GeV}^{-2} \tag{72}$$

i.e., right into the bulls eye of (71)! Thus, cosmology suggests that, perhaps, we do not need to go much beyond the picture of weak interactions in order to get a viable DM candidate – the only prerequisite is that it should be massive enough, typically with m_X in the GeV ballpark. Such CDM candidates are then called WIMPs (for weakly-interacting massive particles) and most of the past and current experimental searches for CDM were/are focusing on them.

Laboratory based WIMP-like CDM searches
Here come the slides I have got on the indirect and direct detection; slightly outdated, but the salient features/aspects are still there.

3.3.3 Popular CDM particle-physics candidates

Just for illustration, let us discuss two popular incarnations of CDM in particle physics. Since, as we have seen, the only DM particle candidates (i.e., neutral with respect to electromagnetic and strong interactions) available in the Standard Model of particles and their interactions are neutrinos = hot relics, we have to speculate here about particle physics beyond SM.

Low-scale SUSY neutralinos
Low-energy supersymmetry [3] is still one of the most popular extensions of the SM. Besides devising a bosonic candidate to each SM fermion and vice versa (and, thus, establishing a nice symmetry between fermions and bosons) it gained a lot of attention in the last decades because it can address several issues of the SM:

- Unlike in the SM, in the TeV-scale SUSY context the running gauge couplings converge to practically one value at $M_G \sim 10^{16}$ GeV.
- TeV-scale SUSY with R-parity (an extra symmetry that prevents baryons from decaying overly fast) provides a very natural WIMP-type CDM candidate in the form of the neutralino LSP[13] which is typically a weakly interacting stable TeV-scale fermion (a spin-$\frac{1}{2}$ state admixed from the SUSY partners of the photon, Z-boson and the partners of a pair of neutral Higgs zero-charge scalars). As such, it entertains a lot of attention including intense collider searches at the LHC.

[13] If R-parity is exact the lightest supersymmetric particle (LSP), i.e., a state with odd R-parity, is absolutely stable (because it has nothing to decay into) and, thus, a potentially good DM candidate.

RH neutrinos
Another speculative (yet rather popular) candidate for CDM (not WIMP-like though) is the RH neutrino featuring in perhaps the most popular extension of the SM that can account for non-zero neutrino masses, namely, the type-I seesaw model.

- They are not absolutely stable because they are coupled to the rest of the world through Yukawa interactions with H and L (otherwise they could not supply light neutrino masses).
- Hence, one has to deal with their decays kinematically, i.e., close all fast decay channels. The fastest of these is namely the $N \to \ell H$ (open for $m_N > m_H$); in order for this to be slower than 10^{17} seconds the Yukawa's must be so small that the light neutrino masses can not be obtained and, thus, m_N must be smaller than m_H.
- For $m_N < m_H$ the decays are still possible, either by going to higher loops or by exploiting the LH chirality neutrino (i.e., a component of the $SU(2)_L$ doublet) admixture within the N mass eigenstates (again, inevitable due to the need to implement the seesaw).
- The fastest processes of this kind are radiative decays into $\nu\gamma$ and Z-mediated decays into three light leptons (e.g., $e^+e^-\nu$). Especially the last one is strongly constraining and requires $M_N \lesssim 10$ keV or so.
- However, for so light N's its couplings to the rest of the world are so small that such RH neutrinos are decoupling too soon so they behave more like a warm relic rather than the cold one.

Let me note that none of the candidates above is fully satisfactory - no SUSY effects have been observed at the LHC up to several TeV energy scales; the sub-MeV RH neutrinos are not satisfactory either as they go against the merits of the original seesaw idea and, among other things, they invalidate the natural leptogenesis picture a la Davidson and Ibarra (see the BSM course) that makes it possible to overcome another serious SM issue, namely, its incapacity to account for the relic baryon to photon number density ratio. This, as we shall see, will be a crucial parameter for understanding the yields of another super-important process that took place about 3 minutes after the big bang, namely, the nucleosynthesis.

I'll finish this later.

3.4 Relics with stronger than weak-interaction couplings to the bath

So far, we have been discussing two examples of thermal relics, namely, neutrinos and WIMPs, both from the point of view of their viability as DM candidates. Notice that in both these cases their interactions with the rest of the world were rather weak - besides gravity, neutrinos are subject only to the weak nuclear force; similarly, the "optimal" annihilation cross-sections for GeV-scale CDM WIMPs was also of the size of a typical weak interaction.

Recall also that, in spite of this similarity, these species have otherwise very different properties - neutrinos represent a hot relic while WIMPs are cold. Of course, both this has to do with their masses - the almost massless neutrinos are not coupled to the bath at all when it would eventually loose its kinematical capacity to pair produce them; on the other hand, WIMPs are so massive that the bath can no longer pair produce them way before their annihilation effectively stops.

Let us now turn our attention to the species that are more strongly coupled to the bath, namely, those that entertain electromagnetic and strong interactions.

3.4.1 Relics with electromagnetic interactions

There is only one type of primordial thermal relics of this kind in the Universe nowadays, namely, the electrons[14] On one hand, their electromagnetic interaction with the bath is way stronger than the weak interaction that drove the fate of neutrinos, on the other hand, electrons are still very light so in view of what we just said above it might not be a-priori clear whether they are hot or cold relics.

OK, let us start with showing that electromagnetically interacting relics can not be hot: If they were, we could neglect their rest masses and, thus, their annihilation cross section must go like[15]

$$\sigma \propto \frac{\alpha^2}{T^2} \tag{73}$$

and since $|v| = 1$ and $n_e \propto T^3$ for relativistic species one has

$$\Gamma_{ee\to\gamma\gamma} \sim \alpha^2 T \tag{74}$$

which is supposed to be compared to $H \sim T^2/M_{Pl}$ in order to determine the relevant freeze-out temperature. However, doing so reveals that the ratio $\Gamma/H \sim \alpha^2 M_{Pl}/T$ even grows(!) with decreasing temperature which would mean that as hot relics electrons would never decouple from the bath which is absurd. Note that this is all due to the specific behaviour of (73) in which $n = -2$ because the relevant coupling (the fine structure constant here) is *dimensionless*[16] (as it is the case for all renormalizable interactions).

Hence, electromagnetic relics must necessarily be cold. Then, the calculation of the freeze-out temperature of their annihilations goes as follows: For non-relativistic electromagnetically interacting species one has the Thomson-like[17] cross-section (barring the $O(1)$ angular part)

$$\sigma \propto \frac{\pi\alpha^2}{m|p|} = \frac{\pi\alpha^2}{m^2|v|} \tag{75}$$

rather than (73) and thus $\Gamma_{ee\to\gamma\gamma} \sim \frac{\pi\alpha^2}{m^2}$ which yields $n = 0$ in formula (61). Hence, $T_{FO} \sim m/K$ with

$$K \approx 0.1 M_{Pl} \frac{\alpha^2}{m} \tag{76}$$

which even for heavy (i.e., GeV-scale) electromagnetically interacting species like protons yields $K > 10^{14}$ and, thus,

$$T_{FO}^{elmg.} \lesssim m/30 \tag{77}$$

[14] There is also a small population of positrons out there but, as we shall see, practically all of them have been created as secondaries way later so we won't count them here.

[15] A more detailed QED calculation reveals that $d\sigma_{ee\to\gamma\gamma}/d\Omega \propto \frac{\alpha^2}{4E^2}$ in the ultrarelativistic regime.

[16] To be compared to the Fermi constant governing effective weak interactions which has dimension of mass^{-2}.

[17] To the terminology: Strictly speaking, the term Thomson cross-section refers to the low-energy form of the general formula for the elastic scattering of photons on free charged particles, with "Compton scattering" representing the high-energy counterpart of the same. None of these are directly related to the annihilation cross-sections at stakes here so the term Thomson-like stems solely from the visual similarity between the Thomson formula $\sigma_T = \frac{8\pi\alpha^2}{3m^2}$ and that of (75).

(and even somewhat less for electrons). Thus, one can formulate a general rule that *electromagnetically interacting species are always cold relics, they are kept in a close contact with the bath for a long time after they become non-relativistic and, thus, their relic density is very strongly Boltzmann suppressed.*

Note also that the result of the calculation does not depend that much on the fact that we have taken a cold relic which entertains electromagnetic interactions. As a matter of fact one would obtain just a bit larger suppression even if we swapped the electromagnetic for the strong interaction cross-sections, see sect. 3.4.2.

The only exception to this rule of thumb is the situation when the population is asymmetric, i.e., if there is an appreciable excess of either X or $\bar{X}$ over the other - in such a case the annihilations can stop earlier due to starvation and the population of the excessive species can survive in appreciable amounts. This, as we shall see, is exactly what had to happen to baryons in the early universe - their measured relic number density is way too high to be a result of a symmetric annihilation.

3.4.2 Relics with strong interactions - baryons

Talking about baryons (that we all are made of) or better about the baryon to photon number density ratio let us look at them as at a canonical example of species whose annihilation rates in the early Universe were even faster than those discussed above as they were driven by strong interactions. Note, however, that due to their relatively small masses their freeze-out temperature (which, as we know is below their rest mass scale) was significantly lower than that of WIMPs so baryons begun disappearing from the bath way later than CDM.

The processes of nucleon annihilation (especially that of neutrons which, unlike protons, do not feel electromagnetic interactions) in the early Universe with T below $\Lambda_{QCD} \sim$ 150 GeV are modelled by the pion exchange which suggests $\sigma_0 \propto \frac{1}{m_\pi^2}$ or so. Plugging this into (60) one obtains (neglecting the O(1) factors around there)

$$K \sim 0.04 \times \frac{M_{Pl} m_n}{m_\pi^2} \sim 4 \times 10^{18} \tag{78}$$

and, thus[18],

$$T_{FO}^{n,\bar{n}} \sim m_n / \log K \sim 20\,\mathrm{MeV}\,. \tag{79}$$

For the relic density one has approximately

$$n_{n,\bar{n}}(T_0) \sim \frac{T_0^3}{\sigma_0 m_n M_{Pl}} \sim \frac{m_\pi^2}{m_n M_{Pl}} n_\gamma(T_0) \sim 0.03 \times K^{-1} n_\gamma(T_0) \tag{80}$$

and, thus, the baryon to photon number density ratio (which is a quantity that can be measured in many different ways) obeys

$$\eta_B \equiv \frac{n_{n,\bar{n}}(T_0)}{n_\gamma(T_0)} \sim 10^{-20} \tag{81}$$

This means that the populations of baryons in the early Universe underwent a very drastic number density reduction due to their strong coupling to the plasma which kept them on the Boltzmann suppression descent for a very long time(!)

[18] As expected, the K factor for strong interactions is even larger than that for the electromagnetic ones, see Sect. 3.4.1.

However, the number we obtained is in a flagrant conflict with the experimentally measured value of $\eta_B^{\rm exp.} \sim 10^{-9}$, see sect. 5 and, as such, it represents on of the most prominent issues of the SM of particle interactions.

4 Thermal history of the early Universe

However, in order to understand where does the experimental value of $\eta_B^{\rm exp.} \sim 10^{-9}$ come from we have to divert a bit for a moment and turn our attention to the time scales at which various processes took place in the early universe. This is important namely for the understanding of when the primordial nucleosynthesis, which is one of the classical keys to the measurement of η_B, did take place.

4.1 Modes of survival

Before doing so let us first briefly recap the different ways in which various relics might have survived until now:

0. *Non-thermal cosmologically stable relics* such as primordial monopoles which, after their creation by non-perturbative processes[19] in the early Universe were no longer produced from the bath which was essentially too cold to do that in any appreciable amounts. This, however, does not mean that one can not look for their residual annihilations. We did not say much about their expected number densities because the very powerful guiding principle of thermalization was not applicable here.

1. *Thermal relics that never experienced Boltzmann's suppression (a.k.a. hot relics)*: These are the species that have remained relativistic until very late (like neutrinos) or even never became non-relativistic (like photons or, perhaps, the lightest neutrino) and are stable on the cosmological scales.

2. *Thermal relics that did experience Boltzmann's suppression (a.k.a. cold relics)*:

   - Weakly interacting cold relics that underwent a decent Boltzmann's suppression but managed to get diluted and cease annihilating early enough (due to their feeble annihilation rates) to survive in appreciable amounts - this is the WIMP story.
   - Electromagnetically or strongly annihilating cosmologically stable relics which under normal circumstances (i.e., if their population was exactly symmetric) would undergo drastic Boltzmann suppression; however, they might have survived in appreciable amounts due to a primordial asymmetry. This is the case of protons and electrons
   - Neutrons (though sharing most of the basic characteristics with protons) however, do not belong to the category above because they are not cosmologically stable. As such, their number densities were not fully fixed even after their annihilations stopped and it would have been just a matter of time when they would eventually disappear if they were not captured (and, thus, stabilized) in nuclei or if the capture happened too late after the became free.

[19] Those are processes for which, roughly speaking, one can not draw a Feynman diagram.

Thus, it is important to study the time scales at which neutron capture really took place as only then one can say how many neutrons might have survived until now (predominantly in the form of the primordial $^4_2$He) and compare the theoretical predictions with observations.

4.2 Time scales

So far we were mainly concerned with the temperature of photons T as the main coordinate changing along the early Universe evolution and governing most of the quantities of our interest (equilibrium number densities, cross-sections, the Hubble rate, freeze-outs etc.) Let us now convert this information into time.

The first Fridmann's equation is the key here - in a radiation dominated universe it connects the time derivative of the energy density in relativistic species to itself, namely

$$\dot{\rho} = -4H\rho \tag{82}$$

with

$$H^2 = \frac{8\pi G}{3}\rho \tag{83}$$

provided

$$\rho = \frac{\pi^2}{30} g_* T^4$$

Combining these as we did in Eq. (28) one can readily derive the time evolution equation for the temperature as a function of time

$$\frac{d}{dt}T = -HT = -\sqrt{\frac{8\pi G}{3}\rho}\, T = -CT^3 \tag{84}$$

where

$$C \equiv \sqrt{\frac{8\pi^3 G g_*}{90}} \tag{85}$$

with g_* counting as before the relativistic species as $g_* = \sum_B g_B + \frac{7}{8}\sum_F g_F$. Note that this factor (and, thus, also C) features a residual T dependence through the θ-like changes of g_* as formerly relativistic species become one after another non-relativistic and, thus, practically irrelevant for the evolution. However, in between these transitions C is constant and the differential equation (84) is separable and as such it can be integrated trivially:

$$t_1 - t_0 = \sqrt{\frac{45}{16\pi^3 G g_*}} \left(\frac{1}{T_1^2} - \frac{1}{T_0^2} \right) \tag{86}$$

As we shall see, g_* will be changing considerably only at several characteristic temperature scales which shall be separated enough for each of the preceding $1/T_{i-1}^2$ to be parametrically smaller than $1/T_i^2$. Thus, also each subsequent period of practically constant g_* will be parametrically longer than the preceding ones so at all times (t_1) we shall be able to neglect the previous "offsets" due to t_0's. Hence, for all practical purposes one can write to a very good precision

$$t \approx \sqrt{\frac{45}{16\pi^3 G g_*(T)}} \frac{1}{T^2} \qquad \Leftrightarrow \qquad t[s] \approx 2.4 \frac{1}{\sqrt{g_*(T)}} \left(\frac{\mathrm{MeV}}{T} \right)^2 \tag{87}$$

In any case we first need to draw a "population chart" of the first seconds of the primordial plasma to map the behaviour of g_*.

The SM epoch: $T \gtrsim 200$ GeV

Active fields	masses ranges	spin	$g_{\rm eff}$
quarks+antiquarks 6 flavors, 3 colors, 2 spins each	$m_t \sim 175$ GeV $m_b \sim 4$ GeV $m_c \sim 1$ GeV $m_{s,d,u} \sim$ massless	$\frac{1}{2}$	$\frac{7}{8}\,(6\times 3\times 2\times 2)$
charged leptons+antileptons 3 flavors, 2 spins each	$m_\tau \sim 1.7$ GeV $m_{e,\mu} \sim$ massless	$\frac{1}{2}$	$\frac{7}{8}\,(3\times 2\times 2)$
neutrinos+antileptons Weyl: 3 flavors, 1 spin each XOR Majorana: 3 flavors, 2 spins, no anti	$m_{\nu_i} \sim$ massless	$\frac{1}{2}$	$\frac{7}{8}\,(3\times 1\times 2)$ XOR $\frac{7}{8}\,(3\times 2\times 1)$
gauge bosons $SU(3)$: 8 gluons $SU(2)$: 3 A-fields $U(1)$: 1 B-field	$m_G = 0$ $m_W \sim 82$ GeV $m_Z \sim 91$ GeV $m_\gamma = 0$	1	8×2 3×3 1×2
Higgs scalar	$m_H \sim 125$ GeV	0	1

Table 1: The list of species present in the thermal bath in significant amounts (and contributing to the total energy density of the Universe) at and before the onset of the electroweak phase transition.

1. **Age of the Universe at the electroweak phase transition @ $T_{\rm EW} \sim 250$ GeV**: The set of "active" degrees of freedom before and at this moment contains essentially all the SM elementary fields whose characteristics are given in Table 1. Thus, in this epoch, $g_* = 106.75$ and

$$t_{\rm EWPT} \approx 2.4 \times 0.1 \times \left(\frac{1}{250\times 10^3}\right)^2 s \sim 4\times 10^{-12} s \tag{88}$$

2. **The QCD phase transition @ $T \sim \Lambda_{\rm QCD} \sim 150$ MeV**:

   The set of degrees of freedom which were relevant for this epoch including the abrupt change in the plasma composition due to the confinement effects are displayed in Table 2. After the QCD phase transition one has $g_* = 17.25$ and, thus,

$$t_{\rm QCD} \lesssim 2.4 \times 0.25 \times \left(\frac{1}{150}\right)^2 s \sim 3\times 10^{-5} s \tag{89}$$

3. **The neutrino decoupling @ $T_\nu \sim 1$ MeV**: This happened "in the middle" of the "electroweak/QED era" in which the relativistic matter degrees of freedom in the bath were dominated by neutrinos and electrons+positrons, as displayed in Table 3.

   Here $g_* = 10.75$ and, thus,

$$t_\nu \approx 2.4 \times 0.3\, s \sim 1 s \tag{90}$$

4. **The disappearance[20] of positrons @ $T_\nu \sim m_e/5 \sim 0.1$ MeV**: The composition of the plasma did not change much from $T \sim 1$ MeV so the d.o.f. counting is practically

[20] Don't forget that the disappearance temperature (i.e., the temperature at which a population has been decimated to such a degree that it no longer plays any role in ρ, say, $T_{dis} \sim m/5$) for cold relics is typically higher than the freeze-out temperature of their annihilation $T_{FO} \sim m_X/30$, c.f. formula (77)!

The QCD epoch: $T \gtrsim \Lambda_{\rm QCD} \sim 150$ MeV

Active fields	masses ranges	spin	$g_{\rm eff}$
Hadrons above Λ_{QCD}:			
quarks+antiquarks 3 flavors, 3 colors, 2 spins each	$m_{s,d,u} \sim$ massless	$\frac{1}{2}$	$\frac{7}{8}\,(3\times 3\times 2\times 2)$
Hadrons below Λ_{QCD}:			
pions, fast decaying 3 singlet scalars	$m_\pi \sim 130$ MeV	0	3×1
Other fields:			
charged leptons+antileptons 2 flavors, 2 spins each	$m_{e,\mu} \sim$ massless	$\frac{1}{2}$	$\frac{7}{8}\,(2\times 2\times 2)$
neutrinos+antileptons Weyl: 3 flavors, 1 spin each XOR Majorana: 3 flavors, 2 spins, no anti	$m_{\nu_i} \sim$ massless	$\frac{1}{2}$	$\frac{7}{8}\,(3\times 1\times 2)$ XOR $\frac{7}{8}\,(3\times 2\times 1)$
gauge bosons: 1 photon	$m_\gamma = 0$	1	1×2

Table 2: The list of species present in the thermal bath in significant amounts (and contributing to the total energy density of the Universe) right before and right after the QCD phase transition which was one of the most abrupt qualitative changes of the properties of the primordial plasma. Note that the confinement effects reduced the number of active (i.e., relativistic) degrees of freedom considerably. The lifetime of pions is very short compared to the age of the Universe here so they disappeared almost immediately after the temperature of the bath dropped below their rest mass.

identical to that in Table 3 until the last moments which witnessed positron extinction. This event occurred few hundreds seconds after the BB, see also Sect. 5.

Note that the last observation will be important for the discussion of the nucleons' chemistry and nucleosynthesis in the next section.

5 Nucleosynthesis

There are three interesting signals which do not fully conform the purely thermal and symmetric picture we have been drawing so far:

1. $\eta_B \sim 6\times 10^{-10}$ is way above the theoretical prediction of $\eta_B \lesssim 10^{-20}$ based on the assumption of full baryon-antibaryon asymmetry in the early universe.

2. There is a clear excess of baryons over antibaryons around there.

3. There is an excess of protons over neutrons which is materialized in the observed 25% of primordial ${}^4_2$He.

The first two are clear and we have discussed them briefly in the preceding section. The third one becomes interesting in the context of the findings of the last lecture, namely, the fact that protons and neutrons (as well as their antiparticles) as strong (and thus cold) relics popped up from the quark-gluon plasma only at $T \sim \Lambda_{QCD}$ which, however, was when they were non-relativistic and, thus, already deep down the Boltzmann's exponential

The QED epoch: $T \gtrsim 0.1$ MeV

Active fields	masses ranges	spin	$g_{\rm eff}$
electrons and positrons 2 spins each	$m_e \sim 0.511$ MeV	$\frac{1}{2}$	$\frac{7}{8}\,(2\times 2)$
neutrinos+antileptons Weyl: 3 flavors, 1 spin each XOR Majorana: 3 flavors, 2 spins, no anti	$m_{\nu_i} \sim$ massless	$\frac{1}{2}$	$\frac{7}{8}\,(3\times 1\times 2)$ XOR $\frac{7}{8}\,(3\times 2\times 1)$
gauge bosons: 1 photon	$m_\gamma = 0$	1	1×2

Table 3: The list of species present in the thermal bath during the QED phase stretching essentially from the moment of QCD phase transition down to the time of positron disappearance/nucleosynthesis at $T \sim 0.1$ MeV. Note that this is the phase where the most important phase of the nucleon chemistry processes took place along with the freeze-out of weak interactions which eventually established the number density ratio of neutrons and protons for deuterium synthesis.

suppression. Even with matter-antimatter asymmetry which eventually saved them from a total extinction it is not entirely clear why such non-relativistic residuals did not start forming heavier nuclei right away.

5.1 The deuterium bottleneck

The answer to this question has to do with the basic features of the residual strong interactions among nucleons:

- Since the typical binding energies of nuclei are in the $T \sim$ few MeV ballpark the plasma at $T \sim \Lambda_{QCD}$ was simply too hot for nucleons to form bound states; we certainly need to wait for longer.

- $T \sim$ few MeV the plasma is too diluted (with density way smaller than that in the solar core) to admit multi-particle scattering. This is relatively easy to see from the baryon number density nowadays encoded in

$$n_B^0 = \frac{\rho_B^0}{m_N} = \Omega_B h_0^2 \times 10^4 \frac{\rm eV}{m_N} {\rm cm}^{-3} \tag{91}$$

  which for $\Omega_B h_0^2 \sim 0.02$ yields about $2\times 10^{-7}{\rm cm}^{-3}$. This, however, is today, i.e., for $T \sim 3$K and one has to rescale this number to say $T \sim 10$ MeV which consists in the multiplication by a factor of

$$\left(\frac{a_0}{a({\rm MeV})}\right)^3 = \left(\frac{10\,{\rm MeV}}{3{\rm K}}\right)^3 \sim \left(\frac{10\,{\rm MeV}}{3\times 10^{-4}{\rm eV}}\right)^3 \sim 4\times 10^{31}\,. \tag{92}$$

  The result is $n_B \sim 10^{25}{\rm cm}^{-3} = 10^{31}{\rm m}^{-3}$ which for $m_N \sim 10^{-27}$ kg yields[21] some 10 t/m^3, i.e., density which is at least factor of 10 less[22] than that in the solar core (some 150 t/m^3).

[21] In this conversion you may appreciate that $1\,{\rm MeV} \approx 11.6\times 10^9$ K or, equivalently, $1\,{\rm K} \approx 0.86\times 10^{-4}$ eV.

[22] Interestingly, the density of the plasma at this moment is less than that in the solar core while its temperature $T \sim 10$ MeV is way higher than the solar core one ($T_\odot \sim 15\times 10^6$ K). This apparent discrepancy (both are thermal systems indeed) is clearly due to the lack of significant pressure gradients in the relatively homogeneous primordial plasma as compared to the stellar environment.

- It is impossible to form anything but deuterons in two-particle interactions (di-protons can not exist due to electromagnetic repulsion and di-neutrons do not exist due to the isospin contribution to the total energy of the system which is negative for isotopic singlet and positive for isotopic triplet – both neutrons have $T_3 = 1/2$ so their compound can not have zero total isospin – even so much that the latter can not be formed at all).

Hence, the first step that had to precede the formation of nuclei was the formation of deuterons (i.e., $p^+ + n \to^2_1 \mathrm{D} + \gamma$). This, as we shall argue now, is very difficult though due to the very low binding energy of $^2_1\mathrm{D}$ ($E_B^D \sim 2.2$ MeV only) and, thus, a very powerful inverse (deuteron photo-dissociation) process $^2_1\mathrm{D} + \gamma \to p^+ + n$ until relatively late (below $T \sim 0.1$ MeV which corresponds to the time of the order of few hundreds of seconds)!

How is this determined? Suppose that protons and neutrons were thermal[23], i.e., distributed by

$$n_p = g_p \left(\frac{m_p T}{2\pi}\right)^{\frac{3}{2}} e^{-(E-\mu_p)/T} \tag{93}$$

$$n_n = g_n \left(\frac{m_n T}{2\pi}\right)^{\frac{3}{2}} e^{-(E-\mu_n)/T} \tag{94}$$

where we had to restore the chemical potentials (first, because they are non-relativistic and, second, because their populations are highly asymmetric with respect to those of $\bar{p}$ and $\bar{n}$). We know little about μ_p and μ_n individually but it does not matter that much because, in equilibrium, their sum is equal to the chemical potential of deuterons (which are charged and, thus, very thermal for a long time) with

$$n_D = g_D \left(\frac{m_D T}{2\pi}\right)^{\frac{3}{2}} e^{-(E-\mu_D)/T} \tag{95}$$

In the equilibrium the creation and destruction processes

$$p^+ + n \leftrightarrow {}^2_1\mathrm{D} + \gamma \tag{96}$$

are balanced, i.e.,

$$\mu_p + \mu_n = \mu_D \tag{97}$$

(recall $\mu_\gamma = 0$). Thus, in the particular combination of

$$\frac{n_D}{n_p n_n} = \frac{3}{4}\left(\frac{2\pi m_D}{m_n m_p T}\right)^{\frac{3}{2}} e^{-(m_D - m_p - m_n)/T} \sim \frac{3}{4}\left(\frac{4\pi}{m_N T}\right)^{\frac{3}{2}} e^{+E_B^D/T} \tag{98}$$

the chemical potentials are gone and one can write

$$\frac{n_D}{n_n} \sim \frac{3}{4}\left(\frac{4\pi}{m_N T}\right)^{\frac{3}{2}} n_p e^{E_B^D/T} \tag{99}$$

and we see that deuterons can appear in appreciable amounts (i.e., their number density ratio to that of neutrons becomes macroscopic) if and only if

$$\frac{n_D}{n_n} \sim \frac{3}{4}\left(\frac{4\pi}{m_N T}\right)^{\frac{3}{2}} \eta_B n_\gamma e^{E_B^D/T} \sim \frac{3}{4}\left(\frac{4\pi}{m_N T}\right)^{\frac{3}{2}} \eta_B g_\gamma \frac{\zeta(3)}{\pi^2} T^3 e^{E_B^D/T} \tag{100}$$

[23]Especially for neutrons this is not so clear; we shall look at this detail later.

$$= \eta_B \left(\frac{T}{m_N}\right)^{\frac{3}{2}} e^{E_B^D/T} \times \frac{3}{4} g_\gamma \frac{\zeta(3)}{\pi^2} (4\pi)^{3/2} \tag{101}$$

was non-negligible. Neglecting the last $\mathcal{O}(1)$ factor one has

$$\frac{n_D}{n_n} \sim \eta_B \left(\frac{T}{m_N}\right)^{\frac{3}{2}} e^{E_B^D/T} \tag{102}$$

which makes it possible to determine the critical T at which, eventually, $\frac{n_D}{n_n} \sim 1$ from

$$\left(\frac{T}{m_N}\right)^{\frac{3}{2}} e^{E_B^D/T} \sim \eta_B^{-1} \sim 10^{10} \tag{103}$$

which yields $T \sim E_B^D/30 \sim 0.06$ MeV quoted above.

Note that the deuterium production temperature/time depends extremely strongly on the baryon to photon number density ratio! Two comments are perhaps worth here:

- It may look weird that a compound with a binding energy of some 2.2 MeV must wait well below 0.1 MeV to be formed; the reason is of course the enormous overabundance of photons in the bath which are Planck distributed and, thus, even for $T \ll E_B^D$ there can be energetic photons in the distribution tail which will still be able to fast dissociate deuterons.

- As we shall see it is not only the deuterium production onset but also the residual number density (after nucleosynthesis goes further, i.e., beyond deuterons, mainly into helium 4) which strongly depends on η_B and, thus, the observations of relic deuterium fraction in the intergalactic hydrogen clouds is a great handle on this very important parameter.

Finally, to set the time scale of the nucleosynthesis onset via formula (87) note that at this epoch $g_* \sim 5$ to 6 (2 from photons and less than 4 from neutrinos due to their smaller than photon temperature) and one has $\sqrt{g_*} \sim 2.5$ and, thus, this temperature corresponds to more than 200 seconds after the BB. These, indeed, are Weinberg's *first three minutes* [4] when most of the crucial particle processes took place.

5.2 The pre-deuteron nucleon chemistry

OK, but how much deuterium was really formed at $T \sim$0.06 MeV? Clearly, the populations of protons and neutrons could not have been symmetric (i.e., $n_p \neq n_n$), if for nothing then clearly for the finite free neutron lifetime which, accidentally also falls in to the higher hundreds of seconds ballpark.

However, during those more than 200 seconds of waiting for deuterium other processes which shaped the neutron number density at nucleosynthesis took place and it is almost a miracle that there were even any neutrons left at that moment!

This is due to the fact that the residual number-density-changing reactions neutrons and protons underwent (due to weak interactions[24]) before the onset of nucleosynthesis

[24]Naively, one might think that the rates of the feeble weak interactions the nucleons can undergo should be totally irrelevant after the freeze-out of the nucleon annihilations (driven by the almighty strong interactions) which occurred at around $T_{FO}^{N,\bar{N}} \sim \frac{m_N}{50} \sim 20$ MeV, cf. Eq. (79). However, unlike for the $N - \bar{N}$ annihilations the projectiles bombarding nucleons in weak processes (i.e., electrons, positrons and neutrinos) are still very abundant even for T well below $T_{FO}^{N,\bar{N}}$.

were (from a certain T onwards) strongly preferring protons over neutrons. Indeed,

$$n + e^+ \leftrightarrow p^+ + \bar{\nu} \tag{104}$$

$$n + \nu \leftrightarrow p^+ + e^- \tag{105}$$

together with the very slow decay $n \to p^+ e^- + \bar{\nu}$ for $T \lesssim 1$ MeV go better from the left to the right than in the opposite direction. The reason is the mass difference $m_n \sim 939.56$ MeV vs. $m_p \sim 938.27$ MeV and the non-zero mass of electrons and positrons $m_e \sim 0.511$ MeV which yield kinematical thresholds of about 1.8 MeV for (104) and some 0.8 MeV for (105) in the right-to-left direction while there is no threshold in the opposite path.

Again, using the distribution (93) and (94) and, in particular, forming their ratio in order to get rid of the chemical potentials (we neglect those of relativistic species) one has

$$\frac{n_n}{n_p} \sim \left(\frac{m_n}{m_p}\right)^{3/2} e^{-(m_n - m_p)/T} \sim e^{-(1.3\mathrm{MeV})/T} \tag{106}$$

and, in principle, with lowering T the neutron to proton number density ratio would undergo an exponential suppression. This is physically clear – for T below about 1 MeV the bath will not have enough energy to produce $n + \nu$ in the $p^+ + e^-$ scatterings and neutrons will undergo just destruction, i.e., the same principle like the Boltzmann's suppression of annihilating cold relics strongly coupled to the bath.

Hence, if the right-going branches of processes (104) and (105) were fast enough there will be almost no neutrons left at the onset of deuterium creation at $T \sim 0.1$ MeV (better to say they would be suppressed by a factor of some e^{-13}, so some 2 ppm). However, we are living in a Universe with about 25% of baryons in helium so there must be something which stops this neutron slaughter very soon after it starts.

The point is that even the neutron destruction becomes inefficient below about 1 MeV in spite of being kinematically perfectly allowed. This has to do with the fact that the nucleon chemistry of (104) and (105) requires weak interactions to be efficient (i.e., faster than the Hubble rate) which, as we already know, is not true below $T \sim T^{\nu}_{FO} \doteq 1$ MeV (this we have already calculated before in Sect. 3.2.1, formula (50)).

5.2.1 A remarkable coincidence

Note again that the key to this claim is that neutron-proton mass difference is comparable to the freeze-out temperature of weak processes! However, from the perspective of the Standard model of particles and their interactions there is no reason whatsoever for this to be the case:

- T^{ν}_{FO} is governed by the Fermi constant (or g – the $SU(2)$ gauge coupling – and m_W in the modern language) and M_{Pl}.
- $m_n - m_p$ is governed by the difference of the up and down quark masses (let me deliberately ignore all the subtleties related to their definitions here) and the effects of electromagnetic interactions – indeed, as members of one isotopic doublet the only difference (besides their constituents' masses) between n and p is the positive electric charge of the latter.

These, however, correspond to practically uncorrelated sectors of the SM; better to say, there is a priori no (known) reason (besides our very existence, i.e., anthropic arguments)

why the critical parameters which drive these quantities could not have acquired different values, with profound implications for the chemistry/composition of the Universe today[25].

5.3 The notorious 25% of primordial Helium

Hence, the leftover neutron fraction is unlikely to be strongly Boltzmann suppressed, i.e., it should be macroscopic at the onset of nucleosynthesis due to the fact that it had to be macroscopic at $T \sim T^{\nu}_{FO}$ and $n \to p^+ e^- + \bar{\nu}$ is slow (even at the scale of hundreds of seconds!).

A detailed account that can be found e.g. in Weinberg's Cosmology [5] yields

$$\frac{n_n}{n_p}(T \sim T^{\nu}_{FO}) \doteq 0.166 \tag{107}$$

which is further reduced to about 0.14 at the nucleosynthesis onset $T \lesssim 0.1$ MeV (i.e., when deuterons begin to form in appreciable amounts). Assuming that practically all neutrons available at that moment eventually make it into ${}^4_2$He nuclei (see below) the amount of primordial helium is readily calculated: if there is 1 neutron per 7 protons at this moment it is like having 2 neutrons per 14 protons. Out of these, there will be 1 piece of ${}^4_2$He created along with 12 pieces of ^{1}H whose mass ratio is 4:12 = 1:3, i.e., in mass fractions, 25% in ${}^4_2$He and 75% in ^{1}H, as simple as that.

There are, however, several questions/caveats of this simple picture to be commented upon:

1. Since it took a long time (more than three minutes) before the conditions (temperature of photons in particular) became favourable for deuterons to be created the bath in fact underwent a very significant dilution; the question may be whether at $T \sim 0.1$ MeV it was still dense enough for neutrons to meet protons with large enough frequency for any appreciable amount of deuterons to be really created.

2. Why do we assume that all neutrons ended up in ${}^4_2$He ? How about other nuclei? In order to understand what is going on it is good to look first at Figure (1) in which the basic information on light nuclei up to $A \sim 13$ is collected (namely, whether a specific ${}^A_Z X$ exists at all and, for those that do, their lifetimes). The first key observation that can be made is that there are no stable (on the nucleosynthesis timescale) isotopes with $A = 5$ or $A = 8$. Second, looking at Figure 2 one can see that ${}^4_2$He is special as it is the nucleus with maximum binding energy per nucleon around (at least until $A = 12$). These elementary observations give rise to the following picture (assuming only two-body reactions running at any appreciable pace, see Section 5.1):

    - The onset of BBN is indeed defined by the moment sufficient amount of ${}^2_1$D appears in the plasma via

$$p^+ + n \;\; \to \;\; {}^2_1\mathrm{D}$$

[25] It is an extremely interesting exercise to think about (for instance) what would change if the weak interactions were somewhat weaker or stronger than what they really are – you'd conclude that even variations as small as a factor of 2 in the value of the Higgs field vacuum expectation value would leave behind Universes that would look totally different from what we observe now! NB Please give it a try; if for nothing then for the fact that I like to ask this question at the exams :-)

Figure 1: Lifetimes (and natural abundances if stable) of the lightest nuclei. Credit: Japan Atomic Energy Agency, Nuclear Data Center, `https://wwwndc.jaea.go.jp/CN14/sp/`

- With p^+, n and ${}^2_1\mathrm{D}$ present in appreciable amounts, the BBN can proceed further via

$$\begin{aligned} p^+ + {}^2_1\mathrm{D} &\rightarrow {}^3_2\mathrm{He} \\ n + {}^2_1\mathrm{D} &\rightarrow {}^3_1\mathrm{T} \\ {}^2_1\mathrm{D} + {}^2_1\mathrm{D} &\rightarrow {}^4_2\mathrm{He} \end{aligned}$$

- With p^+, n, ${}^2_1\mathrm{D}$, ${}^3_2\mathrm{He}$, ${}^3_1\mathrm{T}$ and ${}^4_2\mathrm{He}$ available now, one can imagine a subsequent production of yet heavier nuclei with A from 5 to 8; however, given the instability of the $A = 5$ and $A = 8$ isotopes there are not that many options left in this domain and only ${}^6_3\mathrm{Li}$, ${}^7_3\mathrm{Li}$ and ${}^7_4\mathrm{Be}$ can be produced in any appreciable amounts. However, their production in the reactions on ${}^4_2\mathrm{He}$ is not energetically favourable (the difference of the binding energies of ${}^4_2\mathrm{He}$ and any of these three is simply too large, typically over 2 MeV which is too much given the $T \lesssim 0.1$ MeV available from the bath). Similarly, the rates of their production in other processes (like ${}^3_2\mathrm{He} + {}^3_1\mathrm{T}$) are typically strongly suppressed (${}^3_2\mathrm{He} + {}^3_1\mathrm{T} \rightarrow {}^6_3\mathrm{Li}$ is overwhelmed by ${}^3_2\mathrm{He} + {}^3_1\mathrm{T} \rightarrow {}^4_2\mathrm{He} + p^+ + n$). Hence, the amounts of ${}^6_3\mathrm{Li}$, ${}^7_3\mathrm{Li}$ to be produced are likely to be very small.
- Finally, it is unlikely that any of these (i.e., ${}^6_3\mathrm{Li}$, ${}^7_3\mathrm{Li}$ and ${}^7_4\mathrm{Be}$) undergo a subsequent conversion into anything with $A > 8$ (like ${}^9_4\mathrm{Be}$, ${}^{10}_4\mathrm{Be}$, ${}^{10}_5\mathrm{B}$ or ${}^{11}_5\mathrm{B}$) because the corresponding rates are again strongly suppressed with respect to the dominant branches landing at ${}^4_2\mathrm{He}$+ leftover p^+'s and n's.

All these effects are nicely depicted in Figure 3. To conclude, one may expect that practically all neutrons end up captured within ${}^4_2\mathrm{He}$ with only small amounts of ${}^2_1\mathrm{D}$, ${}^3_2\mathrm{He}$, ${}^6_3\mathrm{Li}$ and ${}^7_3\mathrm{Li}$ left behind, see also Figure 4.

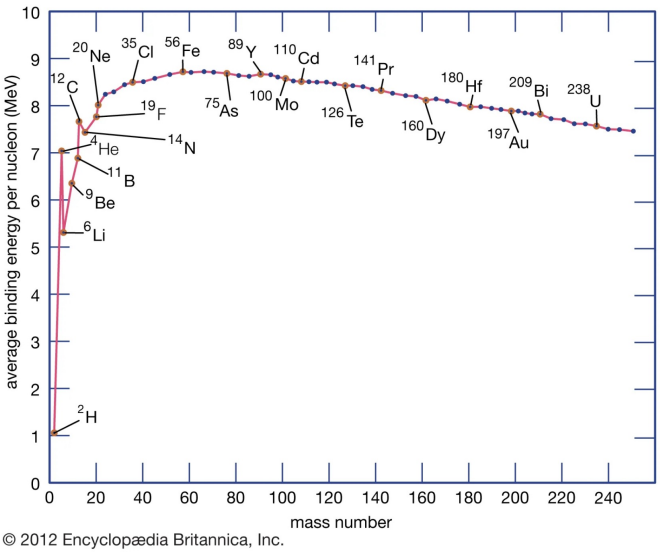

Figure 2: Binding energy (per nucleon) of (some of) the stable elements. One should note in particular the depth of the $^{4}_{2}$He potential well (corresponding to the spike in the Figure) which is behind the strong affinity of many light element's nuclear reactions in the primordial plasma towards this nuclide.

5.3.1 The $^{2}_{1}$D creation efficiency at the onset of nucleosynthesis

The picture sketched above is, however, a bit worrisome because with the BBN onset delayed for so long the plasma has meanwhile undergone a great deal of further dilution and, in principle, the $^{2}_{1}$D production rate may easily become inefficient due to diminishing n_p and n_n (and, as a result, no BBN would occur). Again, this is a question of the relative size of $\Gamma_{np\to D\gamma}$ and H. As for the former, one has

$$\Gamma_{np\to D\gamma} = \langle \sigma_{np\to D\gamma}|v|\rangle n_p \tag{108}$$

where $|v|$ corresponds to the non-relativistic formula

$$|v| = \sqrt{\frac{2T}{m_n}} \sim 0.014 \qquad \text{for} \qquad T \sim 0.1\ \text{MeV} \tag{109}$$

and n_p can be recast in terms of η_B as $n_p = \eta_B n_\gamma$. However, the relevant interaction cross-section $\sigma_{np\to D\gamma}$ is extremely difficult (if not impossible) to calculate from the first principles so here one should better refer to experimental data. These are collected, e.g., in Fig. 5 where one has a product of σ and v in fancy mb×m/ns units. Looking at the value of $\sigma.v$ at around $E_n \sim 0.1$ MeV one reads out $\sigma|v| \sim 4\times 10^{-4}$ mb×m/ns which, to a good precision, we shall also take as the thermal average needed in $\Gamma_{np\to D\gamma}$. A simple unit conversion exercise (1 barn $= 10^{-24}$ cm^2, 1 cm $\approx 10^{-4}$ eV^{-1}, 1 ns ≈ 30 cm) yields

$$\Gamma_{np\to D\gamma} = 4\times 10^{-4}\times 10^{-27}\text{cm}^2 \times 100\,\text{cm}/30\,\text{cm} \times \eta_B n_\gamma \approx 13\times 10^{-31}\text{cm}^2 \times \eta_B T^3 \tag{110}$$

which, substituting $\eta_B \doteq 6\times 10^{-10}$ and $T \sim 10^5$ eV gives

$$\Gamma_{np\to D\gamma} \approx 8\times 10^{-17}\text{eV} \tag{111}$$

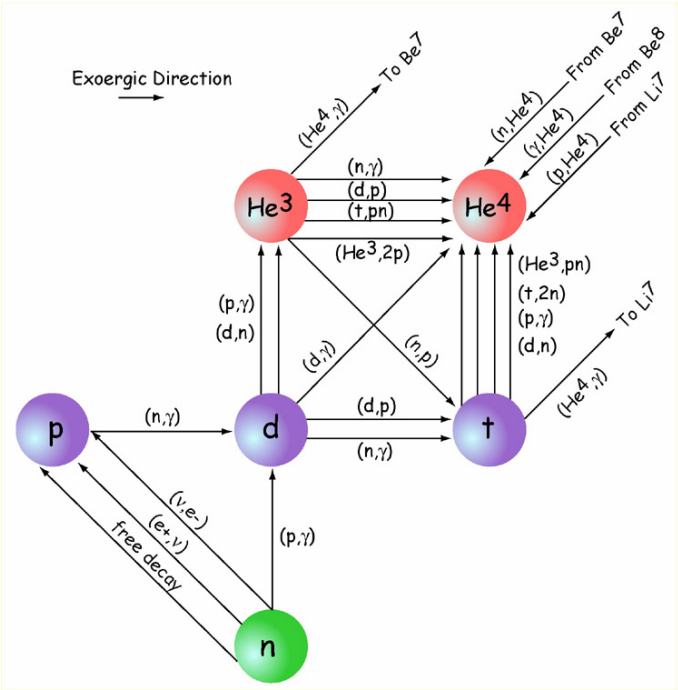

Figure 3: The main nuclear reactions relevant for the primordial (Big-bang) nucleosynthesis. Notice the central role played by ${}^4_2$He as the universal attractor (i.e., all arrows corresponding to the increase in binding energy incoming). Credit: F.X.Timmes, `cococubed.com`

This is to be compared with the Hubble rate at $T \sim 0.1$ MeV which is

$$H(T \sim 0.1\,\mathrm{MeV}) \approx \frac{(10^5\,\mathrm{eV})^2}{10^{28}\,\mathrm{eV}} \approx 10^{-18}\,\mathrm{eV} \tag{112}$$

so we can conclude that $\Gamma_{np\to D\gamma}/H \approx 10^2$ at the onset of nucleosynthesis and, hence, in spite of significant dilution the neutron capture was still efficient and there was no obstacle to deuterium production.

5.4 Relic abundances of other (light) elements and measurement of η_B

As we have already argued, due to the the short lifetimes of the nuclei with $A = 5, 8$ and the affinity of their reactions towards ${}^4_2$He the amount of trans-${}^4_2$He nuclei that were formed in the primordial (Big-bang) nucleosynthesis processes was extremely limited. The only chance for anything else than ${}^4_2$He (along with the dominant leftover hydrogen of course) to survive until today was that not all the long-living progenitors of ${}^4_2$He (i.e., ${}^2_1$D, ${}^3_2$He, ${}^6_3$Li and ${}^7_3$Li) were actually really fished by the ${}^4_2$He-forming reactions; in other words, such processes became inefficient well before this might have happened. Note that the chance that this might be so (i.e., that some of the progenitors might have survived) is indeed indicated by the calculation of the $T \lesssim 0.1$ MeV deuteron creation rate above which, remarkably, was not parametrically larger than the corresponding Hubble rate. Thus, one can indeed expect small amounts of leftover (i.e., relic) ${}^2_1$D, ${}^3_2$He, ${}^6_3$Li and ${}^7_3$Li in the Universe.

Out of these nuclei (predominantly bound in neutral atoms of the intergalactic clouds nowadays) it is namely ${}^2_1$D that deserves a special attention:

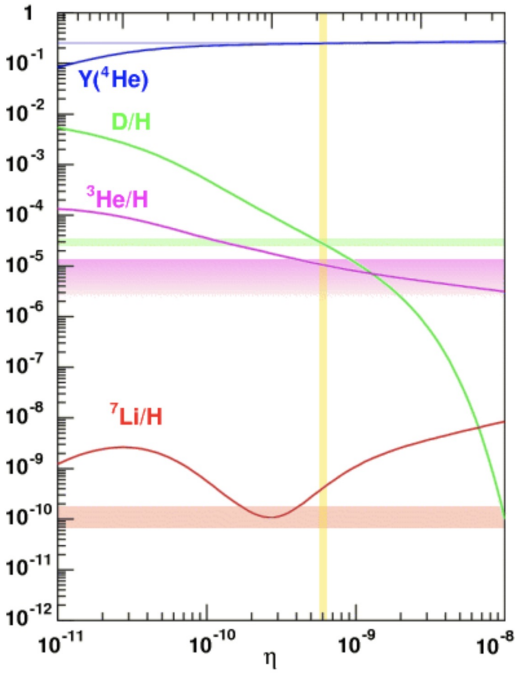

Figure 4: Theoretical relic abundances of light nuclei (at the end of primordial nucleosynthesis) up to ${}^{7}_{3}$Li as functions of η_B. It is interesting to note that the abundance of ${}^{2}_{1}$D is changing relatively fast with η_B (notice its negative slope, for explanation see the text) and, as such, it can be used for independent measurement of this very important parameter. Credit: NASA/WMAP Science Team.

- Unlike for the other nuclei above there is no known astrophysical process that would create ${}^{2}_{1}$D (but there are many which would consume it - nuclear fusion in stellar cores, residual photo-dissociation, interactions with cosmic rays etc.). In this sense, all deuterium that we can observe nowadays is the relic one. This, however, also means that in order to measure its true relic abundance one should look at the most "boring" environments where it is unlikely that ${}^{2}_{1}$D underwent any "reprocessing" in the later stages. That's why the cold intergalactic hydrogen clouds are objects of primary interest here.

- The relic abundance of ${}^{2}_{1}$D should be rather sensitive to the value of the η_B parameter. Indeed, looking at the theoretical curves in Fig. 4 one can see that this is the case and, thus, ${}^{2}_{1}$D is often used as one of the primary sources[26] of information about η_B. It may be surprising that the amount of relic ${}^{2}_{1}$D is decreasing with increasing η_B, i.e., with increasing fraction of baryons – this is due to the fact that with less photons in the bath (i.e., with hypothetically $\eta_B < 6 \times 10^{-10}$) the ${}^{2}_{1}$D production starts earlier (when the plasma is denser) and, thus, the ${}^{4}_{2}$He formation is more efficient; hence, less ${}^{2}_{1}$D remains unprocessed.

[26] Another source of information about baryonic fraction of the critical density (and, thus, η_B) is the CMB power spectrum that you will study with David - the more baryons are there in the primordial plasma the more pronounced the peek corresponding to baryonic acoustic oscillations therein (the first one corresponding to largest angular distances).

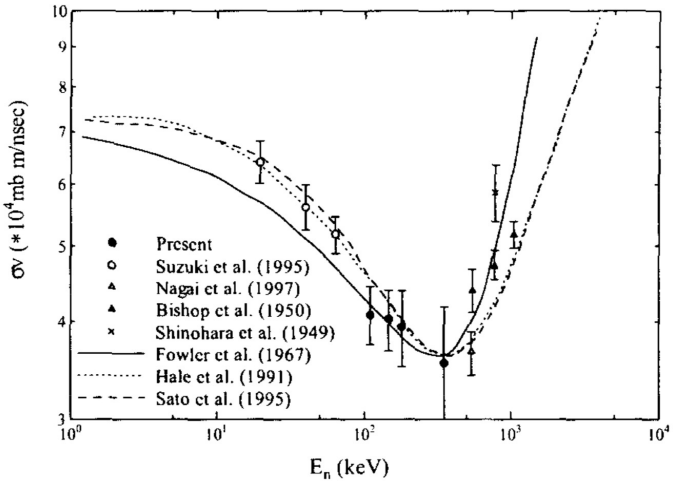

Figure 5: Deuterium production cross section times velocity $\sigma_{n+p\to^2D+\gamma} \times |v|$ as a function of the incident neutron energy E_n in the laboratory frame where protons are practically at rest. Source: A. Tomyo at al., Nucl. Phys. A718 (2003) 401c [6]

6 Recombination

Below $T \sim 0.1$ MeV (which not only corresponds to the onset of nucleosynthesis but also (roughly) to the temperature at which positrons effectively disappeared[27] from the bath), the composition of the thermal primordial plasma became very simple indeed – only photons, atomic nuclei (dominated by hydrogen and ${}^4_2$He) and residual electrons[28] (positrons are just retreating) are still in thermal equilibrium. Dark matter, neutrinos and antineutrinos have decoupled long ago so we shall not consider their effects here.

Given this, the last episode with global effects one can foresee is the moment when electrons and nuclei (re)combine into neutral atoms (mostly hydrogen) by means of the reaction

$$e^- + p^+ \leftrightarrow {}^1H^0 + \gamma \tag{113}$$

which, by the way, is remarkably similar to that of deuterium production, cf. Eq. (96). On the other hand, the binding energy of neutral hydrogen (in the ground state) is about $E_B^H \approx 13.6$ eV, i.e., some five orders of magnitude lower than that of deuterons. Hence, one can foresee the onset of recombination to be in the tens of thousands years' ballpark[29] rather than minutes.

Nevertheless, even in such a case one can to a large degree use the thermal description to capture the salient qualitative features of the recombination process. While this is very clear for the free electrons and protons it may be less obvious for the atomic hydrogen, at least for $T \ll E_B^H$. On the other hand, for T in the eV ballpark (i.e., with large enough

[27] See Sect. 4.2. Note also that this is perfectly consistent with the assumption of the baryon-antibaryon asymmetry that we have employed to save baryons from total extinction – if for nothing then the positive net electric charge of leftover protons had to be compensated by the negative electric charge of the excess of electrons over positrons!

[28] Note that their charge exactly cancels that of the nuclei in the bath.

[29] Here we quote the very naive estimate obtained by just rescaling the time of the nucleosynthesis onset ($\sim$ 3 mins) by the ratio of the two binding energies (squared); many other effects are neglected including different masses of constituents, different values of the relevant cross sections, possible transition from the RD to MD mode etc.

fraction of photons that can still dissociate atomic hydrogen (which, on top of that, will not be all in the ground state)) even the atomic hydrogen fraction in the plasma should still be practically thermal.

This, in turn, admits a mere repetition of what we did for the deuterium equilibrium production (see formulae (93)-(98)) with the result

$$\frac{n_H}{n_p n_e} = \left(\frac{2\pi m_H}{m_e m_p T}\right)^{\frac{3}{2}} e^{-(m_H - m_p - m_e)/T} \sim \left(\frac{2\pi}{m_e T}\right)^{\frac{3}{2}} e^{+E_B^H/T} \tag{114}$$

where the main difference with respect to the former case (98) consists in the value of the prefactor corresponding to the ratio of the g-factors (here it is 1) and the electron rather than nucleon mass in the pre-factor of the exponential term above.

6.1 The fraction of free electrons and Saha equation

Hewever, the quantity on the LHS of Eq. (114) is not very useful for drawing a qualitative picture of what is going on, in particular, how big a fraction of electrons is still free (which, indeed, is what drives the optical thickness of the medium). Hence, in what follows we shall rather work with a quantity

$$R_e \equiv \frac{n_e}{n_B} \tag{115}$$

where n_B is the total number density of baryons (dominated by protons, either free or those already forming neutral hydrogen atoms, i.e., $n_B = n_p + n_H$; note that this quantity is constant in a given comoving volume) with a clear physical meaning of the fraction of electrons which, at the given temperature, are still free. As such, it is an intuitive measure of the degree of ionisaiton ($R_e = 1$ and 0 corresponds to full and no ionisation, respectively).

The point here is that formula (114) can be readily recast for R_e: since $n_p = n_e$ (from the neutrality of the plasma) one has

$$\frac{n_H}{n_p n_e} = \frac{n_B - n_p}{n_e^2} = \frac{n_B - n_e}{n_e^2} = \frac{1 - R_e}{R_e^2} \times \frac{1}{n_B} = \frac{1 - R_e}{R_e^2} \times \frac{1}{\eta_B n_\gamma} \tag{116}$$

and, from (114) one obtains

$$\frac{1 - R_e}{R_e^2} = \eta_B n_\gamma \left(\frac{2\pi}{m_e T}\right)^{\frac{3}{2}} e^{+E_B^H/T} \approx \eta_B \left(\frac{2\pi T}{m_e}\right)^{\frac{3}{2}} e^{+E_B^H/T} \times 2\frac{\zeta(3)}{\pi^2} \tag{117}$$

where the last factor corresponds to the (usually omitted $\mathcal{O}(1)$) factor in the translation of the photon number density into temperature; we retain it at this place for the sake of completeness; however, in what follows, we shall often neglect it. What we have got is the so called *Saha ionization equation* named after the Indian astrophysicist Meghnad Saha who formulated it back in 1920's in his studies of stellar equilibrium.

6.1.1 Solving Saha equation

The next step is to get R_e as a function of T from (117). This is a trivial exercise which yields

$$R_e = \frac{-1 + \sqrt{1 + 4C(T)}}{2C(T)} \tag{118}$$

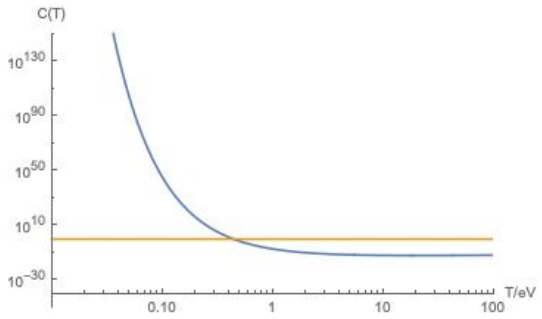

Figure 6: The behaviour of the $C(T)$ defined in Eq. (119) as a function of temperature T. Note that the time axis here goes from the right to the left. The straight horizontal line depicts the $C(T) = 1$ level (corresponding to the ionization fraction $R_e \approx 0.62$, cf. Eq. 118). The recombination temperature (which may be defined by, e.g., $R_e \approx 1\%$, see Sect. 6.1.2) is just a little bit to the left of the intersection of the two lines.

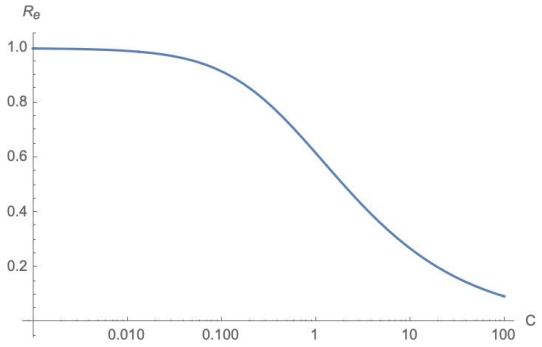

Figure 7: The ionization fraction of the primordial plasma R_e defined in Eq. (115) as a function (118) of the C factor encoding the temperature, see Eq. (119) and Figure 6. Note that here the time goes from the left to the right.

where

$$C(T) \equiv \frac{2\zeta(3)}{\pi^2}\eta_B \left(\frac{2\pi T}{m_e}\right)^{\frac{3}{2}} e^{+E_B^H/T} \tag{119}$$

is a monotonic function which rises with diminishing T (or, equivalently, rises with time), see Figure 6. Two modes of R_e behaviour depending on $C(T)$ are readily recognized, cf. Figure 7:

1. For large temperature the $C(T)$ factor is suppressed (i.e., $C(T) \ll 1$, cf. Figure 6) and $R_e \to 1$ (see Figure 7); hence, the ionization is maximal.

2. Later on, the temperature drops, $C(T)$ enlarges (i.e., $C(T) \gg 1$), R_e starts dropping like $R_e \doteq \frac{1}{\sqrt{C(T)}}$ and the ionization fraction tends to 0.

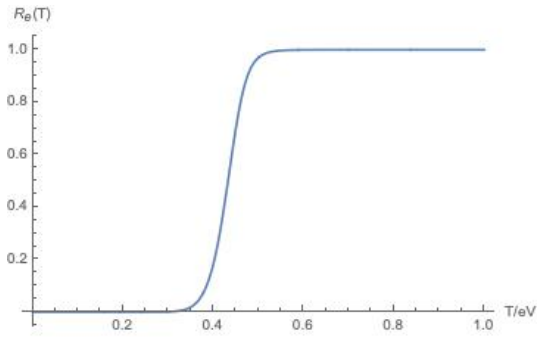

Figure 8: The ionization fraction R_e as a function of T obtained as a composition of the two relations depicted in Figures 6 and 7. Obviously, time here goes from the right to the left.

6.1.2 Recombination temperature

With this at hand, we can finally ask at what temperature $T_{\text{rec.}}$ the recombination occurred. By convention we shall identify it with the situation when there remained the last[30] 1% of electrons free in the plasma, i.e., to the situation with $C(T_{\text{rec.}}) \approx 10^4$. Plugging this back to Eq. (119) one obtains after some numerical tedium $T_{\text{rec.}} \sim 0.3$ eV (see also Figure 8) which corresponds to about 300 000 years after the BB, i.e., not far (on the log scale) from our initial expectation.

6.2 Photon decoupling

However, $T_{\text{rec.}}$ as we defined it is rather arbitrary and it is not even clear how is it related to one of the most valuable sources of information about the early Universe we have (actually, the eldest optical snapshot of it) – the CMB anisotropy map. This, indeed, reflects the state of the primordial plasma at the moment of *photon decoupling* which, in principle, did not need to occur at $T_{\text{rec.}}$ at all.

6.2.1 Did photon decoupling occur before, at or after recombination?

This is not a-priori clear because:

- On one hand, even before recombination matter was already heavily diluted (indeed, we had to wait for very long, i.e., until T dropped below 0.3 eV). So it might have already been very rare for an individual photon to meet a free electron (or a nucleus) even before $T_{\text{rec.}}$. This makes it possible that photons were no longer meeting charged atomic constituents at any appreciable rates even before the onset of recombination.

- On the other hand, before recombination matter was fully ionised and the interaction rate for photons on atomic constituents is parametrically larger than that on neutral atoms and the environment could have easily been optically thick even if the free electron number densities were very low (note that neon lamps have no issues with

[30] This definition of the recombination temperature is of course far from unique and, if you prefer, you can define it in other ways like, e.g., the time when R_e begun deviating from 1 by some amount etc. In this sense, there is nothing really special about $T_{\text{rec.}}$ as we have it besides that it shall correspond to a time which is easy to recall and, thus, can be easily mentioned in the fairy tales you tell to your children.

glowing at pressures as low as few hundreds Pa). This would suggest that photon decoupling rather occurred after recombination.

So the first question to answer is whether photon decoupling occurred before or after recombination. As a matter of fact, if that was the latter case one would have to take into account the Saha suppression of the density of charged particles due to recombination in the calculation of $T^{\gamma}_{\rm dec.}$ from the comparison of $\Gamma_{\gamma e\to\gamma e}$ and H. Note that H here (i.e., for photon decoupling after rec.) would rather be the one for the MD Universe because $T_{\rm equality}\sim 1$ eV is larger than $T_{\rm rec.}\sim 0.3$ eV!

6.2.2 Photon decoupling temperature

To make the puzzle even more interesting let us note that the electron's interaction rate with photons becomes negligible only at around $T\sim 10^{-8}$ eV (see Appendix), i.e., free electrons actually never decoupled from photons! This, however, *does not mean that photons never decoupled from electrons* because their situation is FAR FROM SYMMETRIC: as there are many γ's per relic electron, *individual electrons will always meet some photons*; however, there are very few electrons per photon so *individual photons will almost never meet an electron.*

So, when $\Gamma_{\gamma e\to\gamma e}$, i.e., the rate of photon scattering on electrons (not the opposite!) becomes smaller than H?

1. Assume first that this happened before the onset of recombination, i.e., when the fraction of free electrons that were captured in neutral atoms was still negligible, i.e., there was no or very small Saha suppression. If that was before $T_{\rm eq.}\sim 1$ eV then one can use the RD Hubble rate, if later one should use MD Hubble rate. In the former case one gets (taking $R_e\sim 1$ and using the approximate formula for the Thomson scattering cross-section $\sigma_{\gamma e\to\gamma e}\approx\frac{8\pi\alpha^2}{3m_e^2}$)

$$\Gamma_{\gamma e\to\gamma e}=\sigma_{\gamma e\to\gamma e}|v|n_e\approx\frac{8\pi\alpha^2}{3m_e^2}\eta_B n_\gamma\approx\frac{8\pi\alpha^2}{3m_e^2}\eta_B T^3\stackrel{!}{=}\frac{T^2}{M_{\rm Pl}} \tag{120}$$

   (note that $|v|=1$ here) which yields $T^{\gamma}_{\rm dec.}\approx 0.0001$ eV, while in the latter scenario one has

$$\frac{8\pi\alpha^2}{3m_e^2}\eta_B T^3\stackrel{!}{=}10^{-33}\left(\frac{T}{T_0}\right)^{3/2} \tag{121}$$

   resulting in $T^{\gamma}_{\rm dec.}\approx 0.0036$ eV.

   In either case these results contradict the assumption that $T^{\gamma}_{\rm dec.}>T_{\rm rec.}$

2. Hence, photon decoupling must have occurred *after* recombination and, thus, one has to use the Saha suppression factor in the free electron number density estimate and, at the same time, the MD Hubble rate.

   It should be safe to assume that C in the Saha equation won't be close to 1 (because then, as we saw above, one gets $T^{\gamma}_{\rm dec.}\sim 0.0036$ which is way below $T_{\rm rec.}$) so it is safe to calculate with an approximate value for $R_e\sim 1/\sqrt{C}$, i.e.,

$$R_e=\left(\frac{2\pi T}{m_e}\right)^{-3/4}\eta_B^{-1/2}e^{-E_B^H/2T} \tag{122}$$

Hence, after the onset of recombination the photon interaction rate on free electrons becomes (cf. (120))

$$\Gamma_{\gamma e\to\gamma e} = \sigma_{\gamma e\to\gamma e}|v|n_e \times R_e = \frac{8\pi\alpha^2}{3m_e^2}\left(\frac{2\pi T}{m_e}\right)^{-3/4} e^{-E_B^H/2T}\sqrt{\eta_B}\,T^3 \tag{123}$$

which is again to be compared to the MD Universe Hubble rate. The solution for $T^{\gamma}_{\mathrm{dec.}}$ here is 0.27 eV which can be finally translated to the famous figure of 380 000 years after the BB.

7 Some very sketchy conclusions

In this set of lecture we have learned several interesting things about the Universe and the miracle of our existence:

- There is a lot one can understand about (not only) the arcane first fractions of a second after the Big Bang on the basis of the simple but profound assumption that the main constituents of the primordial plasma were thermal.
- We are rather sure we understand what was going on in the Universe all the way down to the times as early as one picosecond after its beginning. This is particularly fascinating if compared to our uncertainty of what will happen tomorrow.
- What we see nowadays (including ourselves) is actually the ashes of the annihilation inferno of the first moments of the existence of matter.
- It is almost like a miracle that there was something rather than nothing left in the form of baryons. The Standard model of particle interactions married with the Standard ΛCDM model of cosmology have little to say about this surprise.
- The cold DM component of the Universe which accounts for five times more matter than what is stored in baryons may, to some extent, behave quite like other weakly interacting species we know, but it can not be identified with SM neutrinos.
- Neutrons are perhaps the bravest particles in the entire Universe – nobody else that we know of underwent so many phases of extinction as they did.
- It took eternity (i.e., three minutes) before nuclei beyond that of hydrogen could have first popped up in the primordial plasma. The life based on carbon would be impossible without first generation stars which had to produce it.
- We should be very grateful (to whom you prefer) that the weak and electromagnetic interactions conspired in such a way that the early Universe did not finish in one of the two boring states with practically no relic helium or no hydrogen whatsoever.

In any case I hope you enjoyed the course and I'm looking forward to seeing you at the exams.

In Prague, May 25, 2023
MM

Appendix

Hubble rate in the RD Universe
This was already derived in Sect. 2.5 so here we just collect the result obtained there:

$$H_{\rm RD} = \sqrt{\frac{8\pi^3}{90} g_*} \sqrt{G} T^2 \approx 1.66 \times \sqrt{g_*} \frac{T^2}{M_{Pl}} \tag{124}$$

Hubble rate in the MD Universe
If the Universe is dominated by matter the second Friedmann's equation (for $k = 0$ and $\Lambda = 0$) reads

$$H^2 = \frac{8\pi G}{3} \rho_M \tag{125}$$

and one can conveniently relate ρ_M at a given temperature to that of matter (dark+baryonic) nowadays:

$$\rho_M(T) = \rho_M(T_0) \left(\frac{a(T_0)}{a(T)} \right)^3 = \Omega_{M0} \rho_{\rm crit.0} \left(\frac{T}{T_0} \right)^3 = \Omega_{M0} \times \frac{3H_0^2}{8\pi G} \left(\frac{T}{T_0} \right)^3 \tag{126}$$

where we used $\rho_{\rm crit.0} = 3H_0^2/8\pi G$. Combining the two equations above one arrives at

$$H^2 = \Omega_{M0} H_0^2 \left(\frac{T}{T_0} \right)^3 \tag{127}$$

and, thus,

$$H = \sqrt{\Omega_{M0} H_0^2} \left(\frac{T}{T_0} \right)^{3/2} \tag{128}$$

where the factor under the square root is known; indeed,

$$\begin{aligned} \Omega_{M0} H_0^2 &= (\Omega_{B0} + \Omega_{DM0}) h_0^2 \times (100\,{\rm km s^{-1} Mpc^{-1}})^2 \doteq 0.14 \times (3 \times 10^{-18} {\rm s^{-1}})^2 \\ &= 0.14 \times (2.2 \times 10^{-33} {\rm eV})^2 \end{aligned} \tag{129}$$

Finally, one obtains

$$H_{\rm MD} = 0.4 \times 2.2 \times 10^{-33} {\rm eV} \left(\frac{T}{T_0} \right)^{3/2} \sim 10^{-33} {\rm eV} \left(\frac{T}{T_0} \right)^{3/2} \tag{130}$$

Radiation – matter equality temperature $T_{\rm eq.}$
This is just a matter of comparing the two:

$$H_{\rm RD}(T_{\rm eq.}) = H_{\rm MD}(T_{\rm eq.}) \Leftrightarrow 10^{-33} {\rm eV} \left(\frac{T_{\rm eq.}}{T_0} \right)^{3/2} = 1.66 \times \sqrt{g_*} \frac{T_{\rm eq.}^2}{M_{\rm Pl}} \tag{131}$$

which yields

$$T_{\rm eq.} \approx 0.36 \times 10^{-66} {\rm eV}^2 \times \frac{M_{\rm Pl}^2}{g_* {T_0}^3} \tag{132}$$

Given that $M_{\rm Pl} \sim 10^{28}$ eV, $T_0 \sim 2 \times 10^{-4}$ eV and $g_* \approx 8$ (for photons and neutrinos as the only abundant relativistic species left) one gets $T_{\rm eq.}$ in the eV ballpark.

When did free electrons decouple from the bath?

- Assuming first that this happened during the RD epoch one has to compare

$$\Gamma_{\gamma e\to\gamma e} = \sigma_{\rm Thomson}|v|n_\gamma = \frac{8\pi\alpha^2}{3m_e^2}T^3 \tag{133}$$

  with $H_{\rm RD} \approx T^2/M_{\rm Pl}$ which yields $T^e_{\rm dec.} \sim 10^{-13}$ eV which is obviously inconsistent, so one should rather calculate this in the MD Universe.

- With the MD Hubble rate one has instead

$$\frac{8\pi\alpha^2}{3m_e^2}(T^e_{\rm dec.})^3 \stackrel{!}{=} 10^{-33}{\rm eV}\left(\frac{T^e_{\rm dec.}}{T_0}\right)^{3/2} \tag{134}$$

  and, thus,

$$T^e_{\rm dec.} \approx \frac{1}{T_0}\left(\frac{3m_e^2}{8\pi\alpha^2}\times 10^{-33}{\rm eV}\right)^{2/3} \tag{135}$$

  which yields $T^e_{\rm dec.} \sim 10^{-8}$ eV. Hence, electrons (if they would remain free) would never really decouple[31] from photons!

Note that the fact that $T^e_{\rm dec.}$ calculated from the MD Hubble rate is higher than that calculated with $H_{\rm RD}$ makes sense because in the MD Universe the Hubble rate drops slower with T than that in the RD Universe so Γ's (compared to H) drop relatively faster than in the RD universe so the freeze-outs generally occur earlier.
Note that the same calculation for protons yields about 10^{-4} eV.

References

[1] S. Dodelson, *Modern Cosmology* (Academic Press, Amsterdam, 2003).

[2] E. W. Kolb and M. S. Turner *The Early Universe* Vol. 69 (, 1990).

[3] S. Martin, (1997), hep-ph/9709356.

[4] S. Weinberg, *The First Three Minutes. A Modern View of the Origin of the Universe* (, 1977).

[5] S. Weinberg, *Cosmology* (, 2008).

[6] A. Tomyo *et al.*, Nucl. Phys. A **718**, 401 (2003).

[31] Please do not take this as a claim that electrons in the Universe are still thermal – note that practically all of them are nowadays either captured in neutral intergalactic clouds or in stars, i.e., out of direct contact with CMB photons.