Cosmology

Mikhail Shaposhnikov[a]
[a]EPFL, Lausanne, Switzerland

This series of lectures covers the basics of cosmology from a particle physics point of view. The following topics will be partially covered: expanding Universe, cosmological parameters, generic approach to physical processes in the early Universe, cosmic microwave background radiation, nucleosynthesis, baryogenesis, dark matter, cosmological phase transitions and inflation.

This chapter should be cited as: Cosmology, Mikhail Shaposhnikov, DOI: 10.23730/CYRSP-2025-009.65,
in: Proceedings of the 2023 European School of High-Energy Physics,
CERN Yellow Reports: School Proceedings, CERN-2025-009, DOI: 10.23730/CYRSP-2025-009, p.65.

1 Introduction

At first sight, cosmology and particle physics seem to be unrelated branches of physics. Particle physics aims to describe elementary particles and their fundamental interactions at small scales, say, $l < 10^{-14}$ cm. On the contrary, the goal of cosmology is to describe the structure of the Universe at enormous length scales, $l > 10$ kpc $\simeq 10^{22}$ cm. Can we learn anything from cosmology for particle physics? Can we learn anything from particle physics for cosmology?

The evolution of the Universe provides the bridge between particle physics and cosmology. The Universe is expanding, meaning that it was very dense in the past. The physics of the early Universe depends crucially on the properties of elementary particles and the interactions between them. The observations of the present Universe can give us information about the early Universe, and, therefore, about particle physics. The composition of the Universe provides us with the strongest indication that the Standard Model (SM) of elementary particles is not complete, as it cannot explain why we have more matter than antimatter and what is Dark Matter (DM). Cosmology may supply us with constraints on particle physics theories, sometimes superior to those coming from terrestrial experiments. The list of cosmological insights into high energy physics also includes bounds on neutrino masses and the number of light particle species, and constraints on the properties of hypothetical particles and their interactions.

On the other hand, progress in particle physics has led to many advances in cosmology. The non-conservation of baryon number already present in the Standard Model unified with the new sources of CP violation existing in many of its extensions, has led to a qualitative understanding of the absence of antimatter in the Universe; new stable particles, predicted by different Beyond the SM (BSM) models, may play the role of dark matter in the Universe; consideration of phase transitions in particle-physics models has led to the suggestion of a new paradigm in cosmology - inflation. So a “simple” thing as the dynamics of a free quantum scalar field in the expanding Universe proposes a solution to several outstanding problems in cosmology, such as flatness, horizon, homogeneity and structure formation.

The importance of cosmology for particle physics has risen immensely during the last few years. Indeed, the experiments in neutrino physics and robust conclusions of modern observational cosmology call for extensions of the Standard Model of particle physics. However, the situation is very different from what we had in the years preceding the discovery of the Higgs boson. The consistency of the SM, together with experimental results that existed at the time, allowed us to firmly conclude that either the Higgs boson had to be discovered at the LHC, or new physics beyond the SM must show up. Currently, we know for sure that the SM is incomplete and has to be extended. However, we do not have a firm prediction from particle theory of where to search for new particles and what their masses, spin, and interaction type could be.

The famous cosmologist, Zeldovich, used to say: “The Universe is a poor men accelerator - it can produce very heavy particles and very weakly interacting particles. Unfortunately, the experiment happened just once.” In the absence of hints from particle physics experiments for new physics, cosmology

may help to identify the future directions.

The plan of the lectures is as follows. First, we are going to remind the necessary elements of General Relativity. Then we will go through the basic facts about the Universe: its large-scale isotropy and homogeneity, its Hubble expansion, and learn how to describe the Universe's evolution with the Friedman equations. We will consider the content of the Universe and see that its composition presents a puzzle which cannot be explained by the Standard Model of particle interactions. Then we will turn to the foundations of the Big Bang Theory and set up a semi-quantitative way to analyse the processes in the early Universe, based on the notions of thermal equilibrium and non-equilibrium, on freeze-in and freeze-out of particle reactions. After that, we will use these tools to discuss the origins of CMB (Cosmic Microwave Background), photon and neutrino decoupling in the early Universe, nucleosynthesis, and baryogenesis, leading to a matter-antimatter asymmetric Universe. Next, we will review the evidence for Dark Matter and Dark energy and describe one of the possible Dark Matter candidates. The last topic is inflation: we shall discuss the problems of the standard cosmological model and how they can be solved by cosmological inflation.

This writeup is **not self-contained**, it should be used together with the slides shown at the School. In particular, most of the figures are not reproduced here.

Unless otherwise specified, we will use the natural system of units, in which $\hbar = c = 1$ and energy is measured in GeV.

There are several excellent textbooks on cosmology (e.g. [1–6]) that a reader can consult for a thorough study of the subject. Another source of information is https://pdg.lbl.gov, where updated reviews on Astrophysics and Cosmology can be found. No references to the original works will be given here, see the resources listed above and reviews cited in the text.

2 Some elements of General Relativity

In three-dimensional Euclidean space, the infinitesimal distance dl in Cartesian coordinates x, y and z is written as

$$dl^2 = dx^2 + dy^2 + dz^2 \,. \tag{1}$$

This expression is invariant under symmetries of the Euclidean space, namely rotations and translations. The object which is used for the construction of a relativistic theory is the interval ds,

$$ds^2 = dt^2 - dl^2 = \eta_{\mu\nu} dx^\mu dx^\nu \,, \tag{2}$$

where $t = x^0$ is the time coordinate, and the Minkowski metric corresponding to the flat space-time is given by the diagonal matrix $\eta_{\mu\nu} = diag(1, -1, -1, -1)$. In addition to translations and rotations, the interval is invariant under Lorentz transformations, corresponding to transition between different inertial frames.

The generalisation to curved space-time is associated with the interval expressed through the general metric $g_{\mu\nu}$ which is some function of arbitrary coordinates x^μ,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \,, \tag{3}$$

The metric $g_{\mu\nu}$ is now a dynamical variable, describing the non-trivial geometry of the space-time which is determined by matter and the gravitational field itself. Einstein's theory of gravity is constructed in such a way that its field-theory action is invariant under general coordinate transformation,

$$x^\mu \to x'^\mu = f^\mu(x)\,, \tag{4}$$

where $f^\mu(x)$ are arbitrary functions of coordinates. This ensures, in particular, the validity of the equivalence principle (inertial mass is the same as the gravitational mass).

The action of the 'theory of everything" can be symbolically written as

$$S = S_{\rm gravity} + S_{\rm matter}\,, \tag{5}$$

where $S_{\rm matter}$ includes the Lagrangian of the Standard Model and yet other pieces remain to be determined, whereas the lowest order gravity action is that of Einstein-Hilbert,

$$S_{\rm gravity} = -\frac{1}{8\pi G}\int d^4x\sqrt{-g}\left(\frac{R}{2}+\lambda\right)\,, \tag{6}$$

where g is the determinant of the metric, G is the Newton constant of gravity, λ is the cosmological constant, and R is the scalar curvature constructed out of the Riemann curvature tensor as $R = g^{\nu\sigma}R^\mu_{\nu\mu\sigma}$. The equations of motion, as usual, are derived by the variation of the action with respect to the metric. These are the famous Einstein equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}\,, \tag{7}$$

where $R_{\mu\nu}$ is the Ricci tensor, and $T_{\mu\nu}$ is the matter stress-energy tensor. These equations are capable of describing all gravitational phenomena observed till now. We will use them for the analysis of the expanding Universe. It is customary to absorb the cosmological constant into the matter tensor as

$$\tilde{T}_{\mu\nu} = T_{\mu\nu} + \frac{\lambda}{8\pi G}g_{\mu\nu}\,. \tag{8}$$

This operation makes the Einstein equations look like

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}\,, \tag{9}$$

where we removed the tilde to simplify the notations.

3 Structure of the Universe at large distances and the FRW metric

To give a sense of scales for various objects studied in cosmology and astrophysics, we list here the sizes of some of the most recognisable structures in the Universe:

- Earth radius $\approx 6.4 \times 10^8$ cm.
- Solar radius $\approx 7.0 \times 10^{10}$ cm.
- Earth-Sun distance $\equiv 1$AU $\approx 1.5 \times 10^{13}$ cm.

To go beyond this point, it is useful to adopt a new unit of distance called parsec

$$1 \text{ pc} = \frac{1 \text{ AU}}{1 \text{ arcsec}} = 3.26 \text{ lightyear} = 3 \times 10^{18} \text{ cm}. \tag{10}$$

- A galaxy can be thought of as a disk of radius ~ 30 kpc and thickness ~ 1 kpc.
- The galaxies in our Universe are clumped together in clusters of typical size ~ 10 Mpc. The Virgo galaxy cluster, for example, consists of $\sim 10^3 - 10^4$ galaxies.
- Finally, the size of our observable Universe is $\sim 5 \times 10^3$ Mpc.

The most important observational fact about the Universe is that it is isotropic and homogeneous on large scales, i.e. much larger than the galaxy cluster scale. The isotropy of the Universe can be verified by counting and comparing the number of galaxies in different directions or by observing the cosmic microwave background. To prove the homogeneity of the Universe one should reconstruct its three-dimensional picture by measuring distances between galaxies.

The mathematical description of the homogeneous and isotropic Universe is based on the metric which is known as the Friedmann-Lemaitre-Robertson-Walker metric or simply the FLRW metric or even FRW metric:

$$ds^2 = dt^2 - a^2(t) d\bar{l}^2\,, \;\; d\bar{l}^2 = \left(\frac{d\bar{r}^2}{1 - \kappa \bar{r}^2} + \bar{r}^2 d\Omega^2 \right)\,, \tag{11}$$

where $\bar{r}$ is the radial coordinate, $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the solid angle, and the parameter κ describes the global characteristic of the Universe:

$$\kappa = \begin{cases} +1, & \text{closed, positive curvature} \\ 0, & \text{flat, zero curvature} \\ -1, & \text{open, negative curvature}\,. \end{cases} \tag{12}$$

The function $a(t)$ is called the scale factor. Increasing with time $a(t)$ leads to an expanding Universe, whereas if $a(t)$ decreases the Universe would be collapsing.

The structure of the FRW metric allows us to establish the Hubble law for the expanding Universe. If the galaxies are not too far from each other, the physical distance l between them can be written as

$$l = a(t) d\bar{l}\,, \tag{13}$$

leading to

$$\dot{l} = \dot{a}(t) d\bar{l}\,, \tag{14}$$

where the dot is the time derivative (note that the coordinate distance $d\bar{l}$ is time-independent). Excluding $d\bar{l}$, we arrive to the Hubble law

$$\dot{l} = H l\,, \;\; \text{or} \;\; \vec{v} = H\vec{r}, \tag{15}$$

saying that the running out velocity $\vec{v}$ is proportional to the distance $\vec{r}$. Here $H = \dot{a}/a$ is the Hubble constant.

The running-out velocity and the distance can be traded off to the more convenient quantities, which can be experimentally measured. One of them is the "redshift" z and another is the "luminosity distance" d.

There are several types of objects in the Universe that can be considered as "standard candles". This means that we know their total luminosity L and the spectrum of emitted light *at the position they find themselves*. Examples of such objects include supernovae of the type Ia, first-ranked E galaxies in nearby groups and clusters, first-ranked cluster galaxies in rich clusters, etc. With this knowledge, we can find for these objects the redshifts z, defined as:

$$z = \frac{\lambda_{\rm rec} - \lambda_{\rm emit}}{\lambda_{\rm emit}} \,, \tag{16}$$

where $\lambda_{\rm emit}$ and $\lambda_{\rm rec}$ are the wavelengths of the emitted and received light respectively.

Simultaneously, one can find the luminosity distance r to the corresponding object, by measuring the energy flux f (apparent brightness) from the star or galaxy, through the equation

$$f = \frac{L}{4\pi d^2} \,, \tag{17}$$

For distances much smaller than the size of the visible Universe the luminosity distance is the same as the physical distance, whereas the redshift can be associated with the Doppler effect since the relation between the frequency of emitted and received light is given by

$$z = \sqrt{\frac{1+v}{1-v}} - 1 \simeq v \,, \tag{18}$$

for $v \ll 1$, where v is the relative velocity of the emitter to the receiver. This means that the relation between z and d, in this case, reads $z = Hd$. It has a universal character and does not depend on the matter content of the Universe, on the type of the object, on the frequency of the emitted light or on the direction in the sky. It has been verified in numerous observations, confirming the Hubble law.

If the luminosity distance is large, the Hubble law starts to depend on the composition of the Universe. The speed of light is finite, observation of the sources at large distances means that we observe them in the past. Another way to write the redshift is

$$z = \frac{a_{\rm now}}{a_{\rm emit}} - 1 \,, \tag{19}$$

where $a_{\rm now}$ and $a_{\rm emit}$ are the scale factors of the Universe at present and at the time the light was emitted. This follows from the fact that the frequency ν of light changes in an expanding Universe in such a way that $\nu a =$ const, which is easy to understand because the product νa just shows the number of wavelengths in a box of the size a and this number does not change if the size of the box changes.

A somewhat more involved consideration allows one to relate the luminosity distance d to the

coordinate distance r in the FRW metric. It reads

$$d = \bar{r}\frac{a_{\text{now}}^2}{a_{\text{emit}}}\,. \tag{20}$$

The generalised Hubble law, relating d and z can be found from Equations (19,20), it depends on the way the Universe evolved from the time the light was emitted and absorbed. This, in turn, is a function of the content of the Universe via Einstein's equations (9), which for the homogeneous and isotropic Universe are simplified to the Friedmann equations, discussed below.

4 The Friedmann equations

Since the Universe is homogeneous and isotropic, the stress-energy tensor in (9) must obey the same properties. It is a good approximation to assume that the Universe is a perfect fluid. Its energy-momentum tensor is then

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu - p g_{\mu\nu}\,, \tag{21}$$

with ρ, p, and u_μ being the energy density, pressure, and 4-velocity of the fluid respectively. The Universe is not boost invariant and has a preferred cosmic rest frame, namely the rest frame of the fluid, in which $u_\mu = (1, 0, 0, 0)$. In this coordinate system, the non-zero components of the stress-energy tensor are

$$T_{00} = \rho\,, \tag{22}$$

$$T_{ij} = -p g_{ij}\,. \tag{23}$$

Using this form of the stress tensor we arrive at two equations, being the 00^{th} and ij^{th} components of Equation (9):

$$\frac{\dot{a}^2}{a^2} + \frac{\kappa}{a^2} - \frac{\Lambda}{3} = \frac{8\pi G}{3}\rho\,, \tag{24}$$

$$2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{K}{a^2} - \Lambda = -8\pi G p\,. \tag{25}$$

All the other components of the Einstein equation are identically zero. Note that as a consequence of the homogeneity and isotropy of the Universe, the ten Einstein equations have been reduced to only two. Equations (24) and (25) are known as the *Friedmann equations*.

The two Friedmann equations have three unknowns: ρ, p, and a, i.e. they are under-determined. For them to be solvable, one more equation is needed. The extra equation complementing them is usually taken to be the equation of the state of the Universe: an equation of the form $\rho = wp$, where w depends on the properties of the energy component in question. Below are three types of energy components that are of special importance in cosmology together with their equations of states:

- Non-relativistic matter/dust: $p = 0$.
- Relativistic matter/radiation: $p = \rho/3$.
- Cosmological constant/dark energy: $p = -\rho$ (following from $T_{\mu\nu} = \frac{\lambda}{8\pi G} g_{\mu\nu}$).

If several species are present, the equation state is given by the sum of each component

$$p = \sum_i w_i \rho_i \,. \tag{26}$$

Let us take $\kappa = 0$ for simplicity and consider several important cases.

(i) The radiation-dominated Universe, $p = \epsilon/3$. Here

$$\epsilon = \frac{3}{32\pi G}\frac{1}{t^2}\,,\; a = a_0\left(\frac{t}{t_0}\right)^{\frac{1}{2}}\,,\; H = \frac{1}{2t}\,. \tag{27}$$

(ii) The matter-dominated Universe, $p = 0$. Here

$$\epsilon = \frac{1}{6\pi G}\frac{1}{t^2}\,,\; a = a_0\left(\frac{t}{t_0}\right)^{\frac{2}{3}}\,,\; H = \frac{2}{3t}\,. \tag{28}$$

(ii) The vacuum-energy-dominated Universe, $\epsilon = -p,\ \epsilon > 0$. Here

$$\epsilon = \text{ const}\,,\; a = a_0 \exp\left(Ht\right)\,,\; H = \text{ const} = \sqrt{\frac{8\pi G\epsilon}{3}}\,. \tag{29}$$

The last equation may look counter-intuitive since, despite the expansion of the Universe, the energy density does not change. This is related to the fact that the vacuum pressure is negative and it performs negative work which keeps the energy density exactly constant.

A more general case is a mixture of radiation, non-relativistic matter and the vacuum-energy density. Let us introduce different densities, specific for each type of matter,

$$\Omega_M = \frac{\epsilon_M}{\rho_c},\;\; \Omega_r = \frac{\epsilon_r}{\rho_c},\;\; \Omega_\Lambda = \frac{\epsilon_\Lambda}{\rho_c}\,, \tag{30}$$

where the indices $M,\ r$ and Λ refer to the contributions of matter, radiation and vacuum energy respectively. The parameter ρ_c called the critical density, is given by

$$\rho_c = \frac{3H_0^2}{8\pi G}\,. \tag{31}$$

where index 0 refers to the present time. The Hubble constant H_0 is usually parameterized as $H_0 = 100\ h\ \frac{\text{km}}{\text{s}\cdot\text{Mpc}}$ where h is taken from observations, $h \approx 0.67$. Numerically, the critical density is $\rho_c = 1.88h^2 \cdot 10^{-29}$ g/cm^{-3}.

As the Universe expands, different components of the energy scale in the following way:

$$\epsilon_M \sim a^{-3}\,,\;\; \epsilon_r \sim a^{-4}\,,\;\; \epsilon_\Lambda \sim \text{ const}\,, \tag{32}$$

which follows from Equations (27,28,29) and is easy to understand. The equation for matter tells us that the total energy of non-relativistic matter is conserved, the equation for radiation shows that the total number of photons and other light particles is conserved, while their energy is red-shifted. Thus,

Equation (24) can be written as

$$H^2 = H_0^2 \left(\Omega_r \frac{a_0^4}{a^4} + \Omega_M \frac{a_0^3}{a^3} + \Omega_\Lambda + \Omega_k \frac{a_0^2}{a^2} \right) , \tag{33}$$

where the curvature contribution, $\Omega_k \equiv \frac{k}{a_0^2 H_0^2}$ has been introduced for uniformity of notation. As before, the index 0 refers to the present moment of the Universe expansion and we have

$$\Omega_r + \Omega_M + \Omega_\Lambda + \Omega_k = 1 \, . \tag{34}$$

The dominant component of energy density in the early Universe is that related to radiation. Later on, matter dominates. The curvature contribution, potentially important for the evolution of the Universe at a later time, happens to be numerically unimportant, and the Λ term dominates. The schematic dependence of the scale factor on time is represented in Figure 1. The moment when the matter energy density starts to dominate the radiation, $\Omega_M = \Omega_r$ is important for structure formation. This happens at redshift $z_{\rm eq}$ approximately equal to $z_{\rm eq} = 3.1 \cdot 10^4 \Omega_M h^2 \simeq 3500$ and corresponds to the age of the Universe $t_{\rm eq} = 5.2 \cdot 10^4$ years.

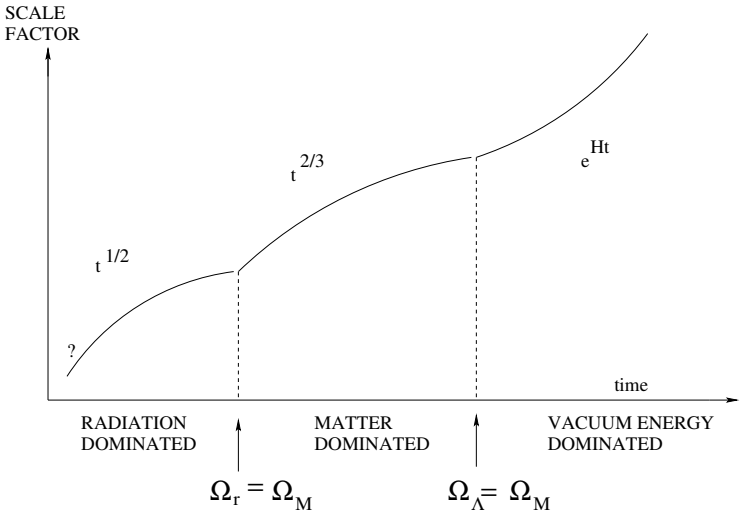

Fig. 1: Dependence of the scale factor on time.

The analysis of the Hubble law at large redshifts based on supernovae, the CMB, the baryon acoustic oscillations, and the dynamics of clusters allows us to pin down the cosmological abundances, giving

$$\Omega_\Lambda \approx 0.68 \, , \;\; \Omega_M \approx 0.32 \, , \;\; \Omega_{\rm DM} \approx 0.27 \, , \;\; \Omega_{\rm baryon} \approx 0.05 \, , \tag{35}$$

where we introduced $\Omega_{\rm DM}$ and $\Omega_{\rm baryon}$ for the DM and baryon abundances respectively (they are discussed in Sections 11,12).

With this set of parameters, the Universe accelerates and will expand forever. The total energy density in the Universe is close to the critical one. This could be considered as an indication of the validity of the inflationary-Universe scenario, discussed below. The age of the Universe, for $\Omega_{\rm tot} = 1$ is $\simeq 13.8$ Gyr, and the value of the Hubble constant is $H_0 \simeq 67$ km s^{-1}Mpc^{-1}.

All these numbers (35) represent a puzzle. If the Standard Model was right, one would expect to have $\Omega_{\rm DM} \ll 1$ and $\Omega_{\rm baryon} \ll 1$ (see Sections 11,12)).

5 Big Bang theory

The common feature of the matter and radiation-dominated regimes of the Universe expansion is the mathematical singularity of the solutions (27,28) at $t \to 0$: the scale factor of the Universe goes to zero, $a \to 0$ and the energy density goes to infinity, $\rho \to \infty$. Physically, all the formulas in this limit do not make any sense: the analysis carried out so far was classical and is not valid when the quantum gravity effects become essential, namely at $\rho \sim M_{Pl}^4$. Here M_{Pl} is the Planck mass, related to Newton's gravitational constant as $M_{Pl} = G^{-\frac{1}{2}} \simeq 1.2 \cdot 10^{19}$ GeV. We do not know what happened then, but the solutions show that the energy density of the Universe in the past was much higher than today. The Universe looked like a plasma of different elementary particles which were close to each other. Hence, reactions between particles were rapid enough, which means that the system was driven to a state of thermal equilibrium and had maximal entropy. This is the essence of the Gamow's Big Bang Theory.

This picture leads immediately to several predictions. One of them is the existence of the CMB - cosmic microwave background radiation which has the Planck thermal spectrum. The other is Big Bang nucleosynthesis, leading to the prediction of abundances of light elements, such as ^4He. The CMB has a thermal equilibrium Planck spectrum with temperature $T = 2.73$ K,

$$dI_\nu \sim \frac{\nu^3 d\nu}{\exp\left(\frac{\nu}{T}\right) - 1} \,, \tag{36}$$

where I_ν is the energy density. The CMB we observe today is simply the ensemble of the relic photons, red-shifted to the present time.

The CMB is highly isotropic. Its dipole anisotropy tells us that the Earth is moving through the CMB with velocity $\simeq 370$ km/s; its multipole anisotropies are at the level of 10^{-5} signal the presence of primordial density perturbations, important for the structure formation (see Section 14.4).

6 Thermal equilibrium

The main assumption of the Big Bang theory about thermal equilibrium is very powerful. Indeed, to describe a generic state of a homogeneous system one would need to know the distribution functions of every particle species, giving the number of particles with a specific momentum. If the system is in thermal equilibrium, it is sufficient to tell what are its temperature T and chemical potentials, corresponding to conserved quantum numbers.

The equilibrium non-interacting particle number distributions are

$$n(p) = \frac{1}{e^{\frac{E-\mu}{T}} \pm 1} \,, \tag{37}$$

where the plus sign refers to fermions and the minus sign to bosons, E is the energy of the particle and μ is the chemical potential. The values of the chemical potentials, at least in the radiation-dominated epoch, are rather small and may be omitted for most purposes.

The number density N and energy density ρ of free particles in thermal equilibrium are given by

$$N = \frac{g}{(2\pi)^3} \int d^3\mathbf{p} n(p) \,, \tag{38}$$

$$\rho = \frac{g}{(2\pi)^3} \int d^3\mathbf{p} E(p) n(p) \,, \tag{39}$$

where g is the number of spin states of a particle, $E(p)$ is the energy of a particle with momentum p. In the relativistic limit $T \gg m$, the integrals in (38) and (39) can be calculated explicitly. The results are

$$\rho = \begin{cases} \frac{\pi^2}{30} g T^4 & \text{boson} \\ \frac{7}{8}\frac{\pi^2}{30} g T^4 & \text{fermion} \end{cases} \tag{40}$$

$$N = \begin{cases} \frac{\zeta(3)}{\pi^2} g T^3 & \text{boson} \\ \frac{3}{4}\frac{\zeta(3)}{\pi^2} g T^3 & \text{fermion} \,, \end{cases} \tag{41}$$

where $\zeta(x)$ is the Riemann zeta function, $\zeta(3) \approx 1.2$. If both bosons and fermions are present, the total energy density can be written compactly as

$$\rho = \frac{\pi^2}{30} g_* T^4 \,, \tag{42}$$

where

$$g_* \equiv \sum_{\text{bosons}} + \frac{7}{8} \sum_{\text{fermions}} \,, \tag{43}$$

is the effective number of relativistic degrees of freedom. The sums above are taken over species that are relativistic at temperature T. For example, for the Standard Model $g^* = 106.75$, if the temperature is well above the Fermi scale. This equation allows one to write a relation between the temperature and the expansion time, combining Equations (27) and (40):

$$t = \frac{1}{2H} = 0.301 \frac{M_{Pl}}{\sqrt{g_*} T^2} \equiv \frac{M_0}{T^2} \,. \tag{44}$$

To appreciate the orders of magnitude, here is an equation to remember:

$$t[\text{s}] = 1/\text{T}[\text{MeV}]^2 \,. \tag{45}$$

In the non-relativistic limit, $T \ll m$, the particle-number densities are

$$n = g \left(\frac{mT}{2\pi} \right)^{\frac{3}{2}} e^{-\frac{m-\mu}{T}} \,. \tag{46}$$

This leads to the energy density $\epsilon = mn$ and to the pressure $p \sim nT \ll \epsilon$. Yet another important thermodynamic quantity is the entropy density, given by

$$s = \frac{2\pi^2}{45} g_* T^3 \tag{47}$$

in the relativistic limit.

7 Freeze in and freeze out

The evolution of the number density N of particles in an expanding Universe is driven by two effects: the diluting effect of the expansion, whose strength is characterized by the Hubble parameter H, and the thermalising effect of particle collisions, whose strength is characterized by the reaction rate Γ

$$\Gamma = \frac{1}{\langle \sigma N v \rangle}\,, \tag{48}$$

where σ is a cross-section of the reaction, N is the relevant particle density, and v is the relative velocity of colliding particles.

The temperature T_* at which

$$\Gamma(T_*) = H(T_*) \tag{49}$$

marks the transition between two qualitatively different behaviours of the number density N. In one extreme $\Gamma(T) \gg H(T)$, the expansion of the Universe is negligible and thermal equilibrium can be achieved without hurdle. In the other extreme $\Gamma(T) \ll H(T)$, interactions happen so rarely that the number of particles in a comoving volume essentially freezes and hence the number density simply scales as $N \propto 1/a^3$. These behaviours are summarized below

$$N \approx \begin{cases} N_{\rm eq}(T), & \Gamma(T) \gg H(T) \\ N_{\rm eq}(T_*) \left(\dfrac{a(T_*)}{a(T)}\right)^3, & \Gamma(T) \ll H(T)\,. \end{cases} \tag{50}$$

In the expression for $\Gamma(T) \ll H(T)$, we have picked the point $T = T_*$ as the scaling reference for N. Since at that point $\Gamma(T_*) = H(T_*)$, and not $\Gamma(T_*) \ll H(T_*)$, the expression is only correct approximately, up to an O(1) factor. The temperature T at which $H = \Gamma$ is called *freeze-in* temperature or *freeze-out* temperature, depending on how the ratio Γ/H changes with time. Freeze-in happens when the system starts in the out-of-equilibrium regime, $\Gamma(T)/H(T) \ll 1$, and moves towards the $\Gamma(T)/H(T) \gg 1$ regime, getting thermalised in the process. Freeze out is the opposite process occurring when an initially thermalized system in the $\Gamma(T)/H(T) \gg 1$ regime moves towards the $\Gamma(T)/H(T) \ll 1$ regime, getting driven out of equilibrium in the process. We will see these processes more explicitly when we study the example below. Before going any further, we would like to point out two caveats of (49) and (50). First, $\Gamma(T)$ here refers to only reactions/collisions that change the number of particles under consideration. Second, (50) is not applicable for unstable, decaying particles.

To get a grasp of the concept of freeze in and freeze out, let us take a look at a gas of electron and positron, interacting mainly via the process $e^+e^- \leftrightarrow \gamma\gamma$[1], in the radiation domination epoch. The

[1]There are of course other processes such as $e^-\gamma \leftrightarrow e^-\gamma$, $e^-e^- \leftrightarrow e^+e^+$, $e^-e^+ \leftrightarrow e^-e^+$, etc, but those processes merely interchange the momenta of the particles involved without changing their particle numbers and therefore irrelevant as far as freeze-in or freeze-out processes are concerned.

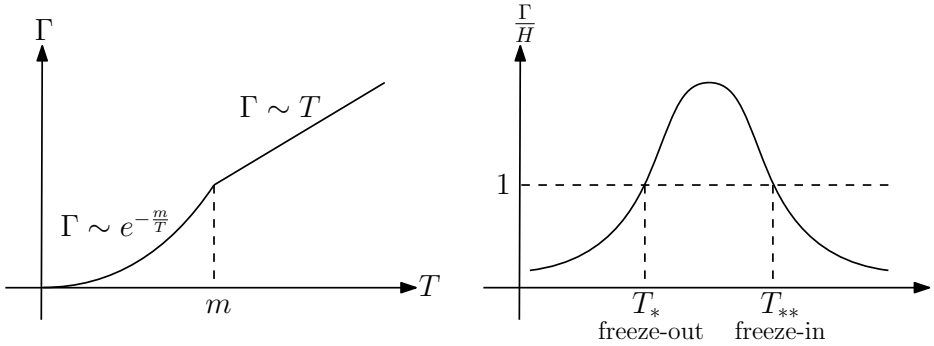

Fig. 2: Temperature dependence of the rate of interaction Γ (left) and its relative size compared to the Hubble parameter Γ/H (right) for the electron-positron-photon system in the radiation domination epoch.

cross-section of the process can be calculated in the centre-of-mass frame to be

$$\sigma \approx \begin{cases} \dfrac{1}{2v}\pi r_e^2, & v \ll 1 \\ \dfrac{m_e^2}{E^2}\pi r_e^2 \left(\log\dfrac{4E^2}{m_e^2} - 1\right), & v \approx 1\,, \end{cases} \tag{51}$$

where $r_e = \alpha/m_e$ with $\alpha = 1/137$ being the fine-structure constant and E is the total energy of the colliding particles in the centre-of-mass frame. The rate of reaction Γ for the cross-section (51) can be computed as

$$\Gamma = \langle \sigma N v \rangle \approx \begin{cases} \pi r_e^2 \left(\dfrac{m_e T}{2\pi}\right)^{3/2} e^{-m_e/T}, & T \lesssim m_e \\ \pi\alpha^2 \dfrac{1}{6}\dfrac{\zeta(3)}{\pi^2} T, & T \gtrsim m_e\,, \end{cases} \tag{52}$$

where we have used (46) in the $T \lesssim m_e$ case, and (41) and $E \sim 3T$ in the $T \gtrsim m_e$ case. The qualitative behaviour of $\Gamma(T)$ is depicted in Figure 2; it is large at high T and small at low T, which follows our intuition: higher temperature implies higher thermal equilibrium density, which, in turn, implies that the particles interact more frequently, hence the higher rate Γ. In the radiation epoch, the Hubble parameter $H(T)$ is given by (44), so

$$\frac{\Gamma(T)}{H(T)} = \begin{cases} \pi r_e^2 \left(\dfrac{m_e T}{2\pi}\right)^{3/2} e^{-m_e/T}\dfrac{M_0}{T^2}, & T \lesssim m_e \\ \pi\alpha^2 \dfrac{1}{6}\dfrac{\zeta(3)}{\pi^2} T\dfrac{M_0}{T^2}, & T \gtrsim m_e\,. \end{cases} \tag{53}$$

As can be seen in Figure 2, $\Gamma(T)/H(T)$ hits one twice, at $T = T_{**}$ (freeze in) and $T = T_*$ (freeze out). These crossing points can be found by equating the above expression to 1. The freeze-in temperature T_{**}, which lies in the $T \gtrsim m_e$ regime, is easier to calculate

$$T_{**} = \alpha^2 \frac{\zeta(3)}{6\pi} M_0 \sim 10^{14}\ \text{GeV}\,. \tag{54}$$

Calculating the freeze-out temperature T_*, which lies in the $T \lesssim m_e$ regime, is somewhat more involved.

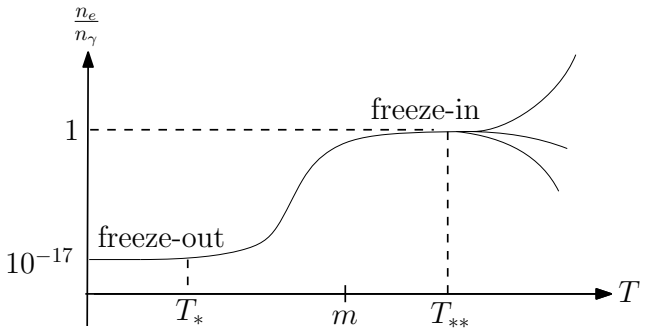

Fig. 3: The behaviour of N_e/N_γ during the freeze-in and freeze-out process. The three different curves on the right part of the graph illustrate the fact that the final, freeze-out value of N_e/N_γ does not depend on the initial value of N_e/N_γ before freeze-in.

It gives

$$\frac{m_e}{T_*} \approx 43\,. \tag{55}$$

After the freeze out, the comoving number density freezes at around its value at $T = T_*$.

In this example, the density of (massless) photons scales as $N_\gamma \propto T^3$ all the time. This means that we can use the ratio N_e/N_γ as a representative of the number density of the electrons. The behaviour of N_e/N_γ during the freeze in and freeze out is shown in Figure. 3. After freeze out ($T < T_*$) N_e/N_γ freezes at its value at $T = T_*$

$$\left.\frac{N_e}{N_\gamma}\right|_{T\lesssim T_*} \sim \left.\frac{N_e}{N_\gamma}\right|_{T=T_*} \approx \left(\frac{43}{2\pi}\right)^{3/2} e^{-43}\frac{\pi^2}{\zeta(3)} \sim 10^{-17}\,. \tag{56}$$

Note that the process of freeze-in, which is nothing but thermalisation, erases any information about the initial conditions of electrons in the Universe. Consequently, the freeze-out density of N_e/N_γ does not depend on the initial state of the electrons at $T > T_*$.

There are plenty of phenomena that can be associated with the freeze-in or freeze-out of different interactions. Below we are going to discuss the decoupling of photons, and neutrinos in the early Universe, nucleosynthesis and baryogenesis as examples.

8 Decoupling of photons

If the temperature of the Universe is larger than the binding energy of electrons in atoms, the cosmic plasma is ionized and the mean free path of photons is rather small so that photons are in thermal equilibrium. When the temperature drops, plasma neutralises and the photons no longer interact with matter but propagate freely. The cosmic microwave radiation, which is observed today, is a snapshot of the Universe at the moment of decoupling. Thus, by the study of CMB today we may find the temperature and matter-density fluctuations, existing at redshifts associated with the photon decoupling.

To estimate the temperature of decoupling one notes that the main reactions to be taken into account are the scattering of photons on electrons, $e\gamma \leftrightarrow e\gamma$ (the cross-section of the γp reaction is much smaller) and the reaction of hydrogen (H) dissociation, $ep \leftrightarrow H\gamma$, that controls the concentration of free electrons. When the second reaction is in thermal equilibrium, concentrations of electrons (n_e), protons

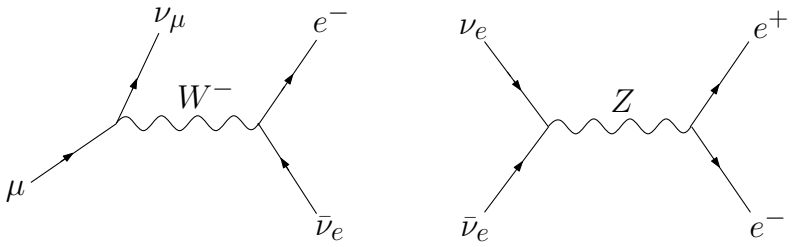

Fig. 4: Charged-current (left) and neutral-current (right) interactions involving neutrinos.

(n_p) and hydrogen atoms (n_H) are related by the Saha formula

$$\frac{n_e n_p}{n_H} = \left(\frac{m_e T}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\frac{I}{T}\right), \tag{57}$$

where $I = 13.6\text{ eV} = 1.58 \cdot 10^5\text{ K}$ is the ionization energy. The decoupling moment is determined by the solution of the equation $\sigma_{\gamma e} n_e \simeq H$, where the Compton scattering cross-section is $\sigma_{\gamma e} = \frac{8\pi\alpha^2}{3m_e^2}$. The system of equations is closed by adding the condition of plasma neutrality ($n_e = n_p$) and introducing as an input parameter the ratio of baryon number (n_B) to the number of photons (n_γ),

$$\eta = \frac{n_B}{n_\gamma} = \frac{n_p + n_H}{n_\gamma} = 6.15 \cdot 10^{-10}, \tag{58}$$

found from different observations.

In addition, the relation between time and temperature can be taken[2] from eq. (44). Numerically, $T^* = 0.25\text{ eV} = 3000\text{ K}$, $z = 1100$, which corresponds to the age of the Universe $t_{\rm dec} \simeq 3.8 \cdot 10^5$ years.

Above T^* the photons are in thermal equilibrium and thus have the thermal Planck spectrum. Below T^* phonons are decoupled and their moment are red shifted. Since the photons are massless, this redshift is equivalent to the change of temperature, explaining why the CMB observed today has the black-body spectrum.

9 Freeze-out and present concentration of relic neutrinos

Apart from the CMB photons, the present Universe is also filled with relic neutrinos that were once in thermal equilibrium but decoupled at some point. Let us estimate the present concentration of these relic neutrinos. For the current purpose, it is safe to assume that the neutrinos are massless.

Before their decoupling, the neutrinos were held in thermal equilibrium mainly by the charged-current and neutral-current interactions, both of which are depicted in Figure 4. The cross-section for these interactions is roughly

$$\langle \sigma_W v \rangle \sim G_F^2 E^2, \tag{59}$$

where $G_F = g^2/M_W^2 \sim 10^{-5}\,\text{GeV}^{-2}$ is the Fermi constant. For $E \sim 1\,\text{MeV}$, we have $\sigma_W \sim 10^{-16}\,\text{GeV}^{-2}$, which is tiny in comparison to typical cross-section for electromagnetic interactions at

[2]Strictly speaking, the decoupling of photons occurs in the epoch of matter dominance, shortly after the moment at which $\Omega_M = \Omega_r$. Accounting for this fact has only a slight influence on the estimate.

the same energy $\sigma_{\rm EM} \sim \alpha^2/E^2 \sim 10^2$ GeV^{-2}. Neutrino decoupling happens when the rate of reaction

$$\Gamma_W = \langle \sigma_W n v \rangle \sim G_F^2 T^5 \tag{60}$$

is equal to the Hubble parameter (44), namely at

$$T_* \sim \left(G_F^2 M_0\right)^{-1/3} \sim 2 \text{ MeV} \,. \tag{61}$$

Interestingly, T_* is larger than the electron mass but much smaller than the masses of all particles other than photons and neutrinos. It means that at $T \approx T_*$ the Universe is populated by a thermal mixture of e^+, e^-, γ, ν_e, ν_μ, and ν_τ and their antiparticles.

Not long after the neutrino decoupling, the temperature of the Universe goes below electron mass due to the expansion, and e^+'s and e^-'s annihilating into photons. By the time $T \ll m_e$, the remaining particles are γ, ν_e, ν_μ, and ν_τ. The e^+-e^- annihilation process increases the temperature of photons but not that of neutrinos, thus creating a difference between the two temperatures. Let us work out explicitly how much they differ. A time long before the annihilations, when the scale factor is $a_{\rm in}$, both the photons and neutrinos have a temperature $T_{\rm in}$. And, long after the annihilations, when the scale factor is $a_{\rm out}$, the temperature of photons and neutrinos are T_γ and T_ν respectively. Since the Universe expands adiabatically, $H \ll \Gamma_W$, we can make use of the conservation of entropy to determine the change in the photon temperature. During the e^+-e^- annihilation, the entropy of e^+'s and e^-'s, which possess $7/8 \times (2+2)$ degrees of freedom, is transferred to photons, which possess 2 degrees of freedom. Therefore, the entropy conservation for e^-, e^+, and γ reads

$$\begin{aligned} \left[s_{e^+} + s_{e^-} + s_\gamma\right]\big|_{a=a_{\rm in}} a_{\rm in}^3 &= s_\gamma\big|_{a=a_{\rm out}} a_{\rm out}^3 \,, \\ \left[2 + \frac{7}{8}\,(2+2)\right] T_{\rm in}^3 a_{\rm in}^3 &= 2 T_\gamma^3 a_{\rm out}^3 \,. \end{aligned} \tag{62}$$

Neutrinos, on the other hand, are unaffected by the annihilations, so their temperature simply rescales as

$$T_{\nu,{\rm in}} a_{\rm in} = T_{\nu,{\rm out}} a_{\rm out} \,. \tag{63}$$

Combining (62) and (63), we find the present temperature of relic neutrinos to be

$$T_{\nu,0} = \left(\frac{4}{11}\right)^{1/3} T_{\gamma,0} = 2 \text{ K} \,, \tag{64}$$

where $T_{\gamma,0} = 2.73$ K.

Independently of whether the neutrinos are relativistic or non-relativistic today, we can estimate their present number density. The electron mass, and therefore the temperature at which the electrons and positrons annihilated, is much larger than the upper bounds on the neutrino masses. So, the neutrinos can be assumed to be relativistic close to the time of electron-positron annihilations. We can thus relate

the number density of neutrinos with that of photons at a time not long after the annihilations as follows

$$n_\nu = \frac{3}{4}\frac{g_\nu T_\nu^3}{g_\gamma T_\gamma^3} n_\gamma = \frac{3/4}{2}\left(\frac{T_\nu}{T_\gamma}\right)^3 n_\gamma = \frac{3}{22} n_\gamma \,. \tag{65}$$

Since both the number density of neutrino and photon scale in the same way as $n \propto a^{-3}$, the above relation is preserved until today and, for one neutrino degree of freedom,

$$n_{\nu,0} = \frac{3}{22} n_{\gamma,0} \simeq 56 \, \frac{1}{\mathrm{cm}^3} \,. \tag{66}$$

This result can be used for getting an upper bound on neutrino masses, by requiring that the total energy density of the neutrinos is smaller than the observed energy density of dark matter. The energy density of dark matter is

$$\rho_{\mathrm{DM}} = \rho_c \Omega_{\mathrm{DM}} \,, \tag{67}$$

where Ω_{DM} is the dark matter abundance and ρ_c is the critical density define in (31). The total energy density of neutrinos today is given by

$$\rho_{\nu,0} = \sum m_\nu n_{\nu,0} = \sum m_\nu \times \left(\frac{3}{22} n_{\gamma,0}\right) \approx \sum m_\nu \times \left(56 \ \mathrm{cm}^{-3}\right) \,. \tag{68}$$

In the last step, (64) and the observed number density of CMB photons $n_\gamma = 411 \, \mathrm{cm}^{-3}$ was used. Requiring that $2\rho_{\nu,0} < \rho_{\mathrm{DM}}$ (the factor of 2 accounts for antineutrinos), we obtain a very robust bound on the sum of the neutrino masses

$$\sum m_\nu < 100 h^2 \Omega_{\mathrm{DM}} \ \mathrm{eV} \approx 10 \ \mathrm{eV} \,, \tag{69}$$

where we have used $h = 0.67$ and $\Omega_{\mathrm{DM}} \approx 0.27$. This bound is more stringent than the direct constraints from particle physics on muon and tau neutrinos. As for experimental detection of the cosmic relic neutrinos, this is very challenging, because of the small interaction cross-section.

10 Big Bang Nucleosynthesis

Roughly speaking, the baryonic matter sector of the Universe consists of 74% of hydrogen and 24% of helium in terms of mass, with a small contribution of other elements. Thermodynamic cooking in stars is not sufficient to explain the observed amount of helium and other light elements in the Universe, suggesting that a cosmological explanation is needed. Big Bang nucleosynthesis (BBN) is our best cosmological explanation for these abundances.

At temperatures above a few hundred MeV, the best description of the Universe is provided by the quark-gluon plasma. Below this temperature, but above a few MeV (the binding energy of protons and neutrons in nuclei) the primordial plasma consists of nucleons rather than nuclei. At smaller temperatures, it is energetically more favourable to hide protons and neutrons in nuclei. The question arises whether all the chemical content of the Universe can be explained by the nuclear reactions occurring at $T \sim 1$ MeV. If not, which elements can be created?

Deviations from thermal equilibrium coming from the expansion of the Universe play an important role in nucleosynthesis. Indeed, in thermal equilibrium, all baryon number would reside in nuclei with the maximal binding energy per nucleon, which is ^{56}Fe. Thus, the dynamics of decoupling of different nuclear reactions must be taken into account. Nuclear abundances are obtained from the solution of a system of kinetic equations incorporating different processes in the expanding Universe. There are various computer codes written for this purpose, which use experimental data for cross-sections of nuclear reactions, supplemented by necessary theoretical information. We shall not discuss this in any detail.

Instead, we will estimate He^4 abundance, which can be done without complicated computations. The first step is to determine the freezing concentration of neutrons. The equilibrium ratio of neutron to proton concentration is simply

$$\frac{n_n}{n_p} = \exp\left(-\frac{m_n - m_p}{T}\right). \tag{70}$$

It is smaller than unity because neutrons are heavier than protons. The fastest reaction that keeps neutron concentration in equilibrium is $p + e \leftrightarrow \nu + n$. A computation similar to that discussed in Section 9 shows that it goes out of equilibrium at $T \simeq 0.8$ MeV. Therefore, $\frac{n_n}{n_p} \simeq \frac{1}{5}$ for temperature $T \simeq T^*$. This ratio becomes somewhat smaller by the time the BBN starts because of neutron decays, see below.

We can define the starting point of BBN as the point when the first process in the chain reaction $\mathrm{p} + \mathrm{n} \leftrightarrow D + \gamma$ "hides" almost all protons inside the deuterium, i.e. when $n_\mathrm{p} \sim n_\mathrm{D}$. The relevant Saha's equation in the present case is

$$\frac{n_\mathrm{p} n_\mathrm{n}}{n_\mathrm{D}} = \left(\frac{m_\mathrm{p} T}{2\pi}\right)^{3/2} e^{-\Delta_\mathrm{D}/T} \tag{71}$$

with $\Delta_\mathrm{D} = m_\mathrm{p} + m_\mathrm{n} - m_\mathrm{D} = 2.23$ MeV, where index "D" is short for deuterium. When $n_\mathrm{p} \sim n_\mathrm{D}$, the Saha's equation reduces to

$$n_\mathrm{n} \sim \left(\frac{m_\mathrm{p} T}{2\pi}\right)^{3/2} e^{-\Delta_\mathrm{D}/T}. \tag{72}$$

Comparing it with (58) and using $n_\gamma \sim T^3$, we find that it occurs when the temperature is

$$T_\mathrm{BBN} = 70 \text{ keV} \tag{73}$$

corresponding to the time $t_\mathrm{BBN} \approx 4.5$min, which is comparable to the lifetime of neutron. Thus, some portion of neutrons must have decayed by then. By the time the deuterium production becomes effective at T_BBN, the ratio of neutron and proton density has reduced from 1/5 to

$$\frac{n_\mathrm{n}}{n_\mathrm{p}} = \exp\left(-\frac{\Delta_\mathrm{D}}{T_*}\right) \exp\left(-\frac{t_\mathrm{BBN}}{\tau_\mathrm{n}}\right) \approx \frac{1}{7}. \tag{74}$$

Now, if one looks at the binding energies of light elements (say, with an atomic number smaller than 8, the cross-sections for the creation of even heavier elements are exponentially suppressed because of the Coulomb barrier) one finds that it is the highest in He^4. Thus, the abundance Y of He^4 is given simply by the number of available free neutrons in the plasma,

$$Y = \frac{\text{number of baryons in He}^4}{\text{total number of baryons}} = \frac{2n_n}{n_n + n_p} \simeq 0.25. \tag{75}$$

Abundances of other light elements (He^3, D and Li) can be found from kinetic equations and theoretical predictions can be compared with cosmological observations, see Figure on page 64 of the slides of the Lecture 1. It is plotted as a function of parameter $\eta = \frac{n_B}{n_\gamma} = \eta_{10} \cdot 10^{-10}$, showing the ratio of baryon to photon density for the case of three neutrino species[3]. Amazingly, all light-element abundances are following observations if η is in the interval $\eta = 6.14(19) \times 10^{-10}$, which may be considered as a most important confirmation of the Big Bang theory up to temperatures of the order of 1 MeV. Other elements, present in the Universe, with atomic numbers greater than 12 are believed to be created in massive stars, while lighter elements, such as B, ^{6}Be, and Li are created by a cosmic ray spallation process.

11 Baryogenesis

11.1 Baryon asymmetry in the early Universe

As we discussed in the previous sections, the parameter $\eta = \frac{n_B}{n_\gamma}$, plays an important role in cosmology. It determines the moment of matter-radiation equality and influences primordial abundances of light elements and structure formation. As we shall see, it is related to the fact that there is no antimatter in the Universe (at least, not in amounts comparable to matter).

Let us understand the parameter η better. Once the temperature goes below the electroweak scale, ~ 130 GeV, the baryon number is known to be conserved to a very good approximation (see below). In an expanding Universe, this translates to

$$(n_B - n_{\bar{B}})a^3 = \text{const}\,, \tag{76}$$

where $n_{\bar{B}}$ is the concentration of antibaryons. Furthermore, as long as the Universe expands adiabatically, the entropy is conserved

$$sa^3 = \text{const}\,. \tag{77}$$

If we take the ratio between (76) and (77), the scale factors would cancel and we get

$$\frac{n_B - n_{\bar{B}}}{s} = \text{const}\,, \tag{78}$$

Taking $n_{\bar{B}} = 0$ and $s \simeq 7n_\gamma$ as the values of these quantities today, one gets

$$\frac{n_B - n_{\bar{B}}}{s} \simeq \frac{1}{7}\eta \simeq 9 \times 10^{-11}\,. \tag{79}$$

Consider now high temperatures, say $T \simeq 130$ GeV. The number of quark degrees of freedom at this time is $\simeq 72$, giving $s = 21.4(n_B + n_{\bar{B}})$. This allows us to estimate the baryon asymmetry defined as

$$\Delta \equiv \frac{n_B - n_{\bar{B}}}{n_B + n_{\bar{B}}} \simeq 3\eta \simeq 2 \times 10^{-10} \tag{80}$$

at this temperature. In other words, the baryon-to-photon ratio today gives us an estimate of the (tiny) baryon asymmetry in the early Universe.

[3] Changing the number of neutrino species changes the rate of the Universal expansion and thus predictions of Big Bang nucleosynthesis. One cannot admit more than four types of massless neutrinos in order not to spoil successful predictions of BBN.

When the Universe cools down from this state, the symmetric part of the baryon-antibaryon background annihilates into photons and neutrinos, but the nucleons that do not find a pair survive, see Fig. 5. These give rise to baryonic matter in the Universe, and eventually galaxies, stars and planets.

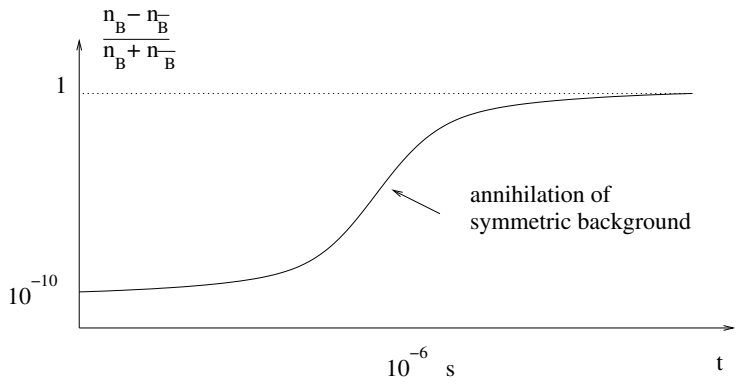

Fig. 5: Dependence of baryon asymmetry on time.

11.2 Sakharov Conditions

Rather than assume that the Universe contained more baryons than antibaryons from the very beginning, it is more compelling to think that the Universe started in a baryon-symmetric state and somehow produced baryon asymmetry as it evolved. The initial baryon symmetric state may be a consequence of cosmological inflation, as will be discussed below. The process of creating baryon asymmetry is dubbed baryogenesis.

From the viewpoint of particle physics, to produce a net baryon asymmetry in the Universe the three so-called Sakharov conditions must be met:

1. Existence of baryon number violating processes.
2. Violation of discreet symmetries C and CP.
3. Departure from thermal equilibrium.

In this case, the Universe could start its expansion from a truly symmetric state, containing an equal number of particles and antiparticles. Then, in the course of the expansion, the particle physics reactions with B, C and CP non-conservation would produce an excess of particles over antiparticles. The third condition is required because in thermal equilibrium the baryon number of the system must be zero: the total rate of the processes which increase the baryon number is exactly compensated by the rate of the processes that decrease it, as a consequence of the CPT theorem.

Depending on the way the three Sakharov conditions are implemented, one can consider different types of baryogenesis mechanisms. We will briefly review two of the most popular ideas, called leptogenesis and domain wall electroweak baryogenesis (for detailed reviews see, e.g. [7–10]), leaving others (such as Grand Unified baryogenesis, Affleck-Dine baryogenesis, baryogenesis from black holes, spontaneous baryogenesis, etc) for individual study (for reviews see, e.g. [11–13]).

11.2.1 Baryon number violation in the Standard Model

Qualitatively speaking, the Standard Model has all the necessary ingredients to create baryon asymmetry: the Cabibbo-Kobayashi-Maskawa (CKM) matrix that describes the mixing among the different quark flavours contains a CP-violating phase, a departure from thermal equilibrium comes from the Universe expansion, and baryon-number violation occurs through the so-called "sphaleron" processes.

On a perturbative level, the Standard Model has four conserved fermionic numbers: B - baryon number, and $L_e,\ L_\mu,\ L_\tau$ - leptonic numbers. The non-perturbative effects break these conservation laws and leave intact only three of them, $L_e - B/3,\ L_\mu - B/3$, and $L_\tau - B/3$. In other words, there is anomalous fermion-number non-conservation in the SM. It can not be described by Feynman diagrams and uses for its description several advanced non-perturbative methods of Quantum Field Theory, such as triangular anomalies, instantons, and non-trivial topological structure of the gauge vacua.

The sphaleron process is a non-perturbative high-temperature reaction violating (B+L) that does roughly the following conversion:

$$9q + 3l \to \mathcal{O}\left(\frac{1}{\alpha_W}\right) bosons\,.$$

It requires the presence of a large number of bosons, coming out from a "sphaleron" decay, with sphaleron being a specific static but unstable solution of the classical equations of motion of the bosonic sector of the SM.

The rate of anomalous fermion-number non-conservation at zero and non-zero temperatures is of the order of:

$$\Gamma \sim \begin{cases} \exp(-\frac{4\pi}{\alpha_W}) \sim 10^{-160}, & T = 0 \\ (\alpha_W T)^4 (\frac{M_{sph}}{T})^7 \exp\left(-\frac{M_{sph}}{T}\right), & T < T_c \\ (\alpha_W)^5 T^4, & T > T_c \end{cases} \tag{81}$$

where $M_{sph} \sim M_W/\alpha_W$ is the sphaleron mass and $T_c \simeq 160$ GeV is the temperature of the electroweak (EW) crossover (see below).

The anomalous fermion number violation of the SM looks to be specially designed for baryogenesis: it freezes in at $T \simeq \alpha_W^5 M_0 \sim 10^{12}$ GeV, freezes-out out at $T \simeq 130$ GeV, and is consistent with all constraints on the matter stability at zero temperature, such as the proton decay.

11.3 Baryon asymmetry and the Standard Model

Still, the Standard Model cannot explain the observed baryon asymmetry. The reason is twofold. First, the SM CP violation has a peculiar structure: it vanishes if any pair of up or down-type quarks degenerate in mass. This allows us to estimate the CP-violating effects without a complicated computation. The relevant energy scale for baryogenesis is the temperature at which sphaleron transitions are active, $T > T_{\rm sph}$. This temperature is larger than the quark masses, which can be considered as a small perturbation.

Now, a polynomial constructed from the quark masses, that obeys the above property is

$$P = (m_t^2 - m_c^2)(m_t^2 - m_u^2)(m_c^2 - m_u^2)(m_b^2 - m_s^2)(m_b^2 - m_d^2)(m_s^2 - m_d^2)\,. \tag{82}$$

Note that the squares of the quark mass should be used, as no physical quantity can depend on the first power of the quark mass. To get a measure of CP violation, this combination should be multiplied by $\sin(\theta_{12})\sin(\theta_{23})\sin(\theta_{13})\sin\delta_{CP}$, where δ_{CP} is the KM phase and the Particle Data Group parametrization of the CKM matrix is used. Indeed, if one or more of the mixing angles in the CKM matrix vanishes, then with a physically unobservable change of phases of the quark fields, the CKM matrix can be made purely real and there is no CP violation. Finally, to get a dimensionless quantity, the combination (82) should be divided by the 12th power of the relevant temperature. Numerically, the effects of the Kobayashi-Maskawa CP violation are of the order

$$\delta_{KM}^{CP} \sim \frac{P}{T^{12}} \sim 10^{-20}\,, \tag{83}$$

i.e., some ten orders of magnitude smaller than the baryon asymmetry of the Universe.

Secondly, the rates of strong, weak and electromagnetic interactions at the sphaleron freeze-out temperature 130 GeV are much faster than the Hubble rate. They keep the distributions of quarks and leptons in equilibrium, and, because of this, the production of any sizable baryon asymmetry is greatly suppressed, as follows from the third Sakharov condition. So, the presence of the baryon asymmetry of the Universe is a strong argument in favour of new physics, which should bring new sources of CP-violation and new mechanisms for the departure from thermal equilibrium.

11.3.1 Leptogenesis

The Standard Model (SM) of elementary particles, defined as a renormalizable field theory, based on the SU(3)×SU(2)×U(1) gauge group, and containing three fermionic families with left-handed particles being the SU(2) doublets, the right-handed ones being the SU(2) singlets and one Higgs doublet, has been used to successfully predict several particles and their properties. It has a strange asymmetry between the quark and lepton sectors. Namely, every left-handed quark or charged lepton has its right-handed counterpart. The neutral leptons – neutrinos are in the SM in different ways – they do not have right-handed partners. This is exactly the reason why the neutrinos are exactly massless in the SM and why there are three distinct conserved leptonic numbers. Nowadays, it is well established experimentally that neutrinos change their flavours and thus have nonzero masses. Therefore, it is natural to upgrade the SM by adding to it $\mathcal{N} = 3$ right-handed neutral fermion states. They can be called singlet or neutral leptons, sterile neutrinos, or alike. The PDG name for them is "heavy neutral lepton", HNL for short. The number "3" looks natural as this is the number of fermionic generations in the SM. It also happens to be a minimal number, which enables us to explain simultaneously neutrino masses and oscillations, and the presence of baryon asymmetry and Dark Matter (DM) in the Universe.

The most general renormalisable Lagrangian including the SM and several right-handed singlet

fermions N_I reads

$$L = L_{SM} + \bar{N}_I i\partial_\mu \gamma^\mu N_I - F_{\alpha I}\, \bar{L}_\alpha N_I \tilde{\phi} - \frac{M_I}{2}\, \bar{N}_I^c N_I + h.c., \tag{84}$$

where ϕ is the Higgs field and $\tilde{\phi}_i = \epsilon_{ij}\phi_j^*$. This Lagrangian can be used for the explanation of the small values of neutrino masses via the see-saw mechanism, see the lectures on neutrino physics at this School by Gabriela Barenboim. If the masses of HNLs are much larger than the electroweak scale, say of the order of 10^{10}, this Lagrangian is called the "type I see-saw model". If the masses of HNLs are below the Fermi scale, this is dubbed as "νMSM" for the Neutrino Minimal Standard Model, stressing that new fermions are similar to other fermions of the SM (for a review see, e.g. [14]).

A new qualitative feature of this theory is lepton number non-conservation. Indeed, since N_I transform as singlets to the SM gauge group, the Majorana mass terms for them are allowed by the gauge symmetries. The Yukawa interactions of HNLs with the leptons contain CP-violating phases, whereas the mere presence of new particles creates new sources for departure from thermal equilibrium during the freeze in or freeze out of HNLs. HNLs can decay or scatter, producing lepton asymmetry. If this happens above the sphaleron freeze-out, this asymmetry is converted into baryon asymmetry by sphalerons (Section 11.2.1). The resulting baryon asymmetry is just a numerical factor of order one smaller than the generated lepton asymmetry. If the HNL masses are smaller than the mass of the Z-boson, the see-saw mechanism of neutrino mass generation and baryogenesis can be verified experimentally at new experiments at CERN such as SHiP, and later at FCC-ee, see Figure on page 16 of the Lecture 2 slides.

11.3.2 Phase transitions and baryogenesis

Yet another idea of how to produce baryon asymmetry in the Universe is associated with first-order phase transitions. The same substance, depending on temperature T and pressure P, can be in a different phase state. For instance, the water can be in solid, liquid and gaseous states, depending on P and T. Similarly, a hot dense plasma of elementary particles can be in different phases. For instance, there could be a first-order phase transition between the symmetric and Higgs phases of the Standard Model when the Universe was cooling down.

The electroweak phase transition (EWPT), if it is strongly first order, would be quite a violent event. The first-order phase transitions go through the bubble nucleation, qualitatively in the same way the water boils or the vapour condenses. EWPT would start at $T_c \sim 100$ GeV. At this time the size of the event horizon is $\simeq 1$ cm. The critical bubbles are of microscopical size, roughly $R \sim (\alpha_W T)^{-1}$. Bubbles nucleate in different places, expand, and eventually collide, filling the Universe with a new phase. The typical size of bubbles at the time they collide is macroscopic, $\sim 10^{-6}$ cm. Inside the bubble the system is in the Higgs phase and the vacuum expectation value of the Higgs field is non-zero. Outside the bubble, the system is in the symmetric phase. Inside the bubble, the rate of B-violation must be small to evade the destruction of the baryon number due to sphalerons. This requirement puts a constraint on the strength of the phase transition. In contrast, outside the bubble the rate is large. The masses of particles are different inside and outside the bubbles (particles inside the bubble are generically heavier since they get their masses from interactions with a Higgs field), leading to their interaction with the bubble walls.

The non-equilibrium motion of the bubble walls, combined with the effects of CP-violation of

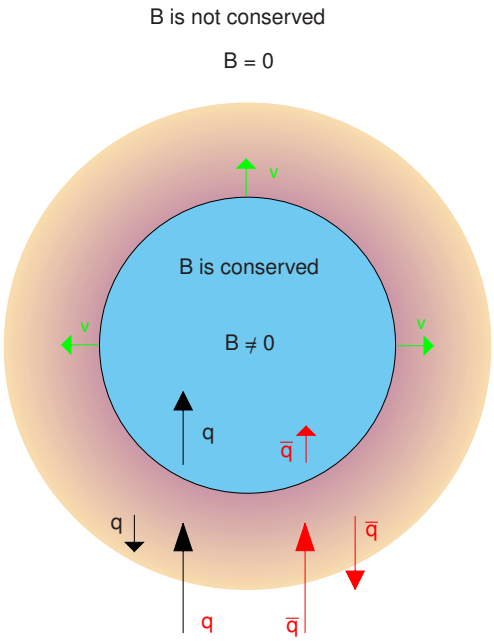

Fig. 6: The expanding bubble. The dark region inside the small circle corresponds to the broken (Higgs) phase where quarks and leptons are massive and baryon number violating processes are strongly suppressed. The region between the two circles corresponds to the plasma in the symmetric phase which is disturbed by the motion of the bubble wall. It is this region which is responsible for the generation of baryon excess. The rate of baryon number non-conservation is high here, while CP-non-invariant interaction of fermions with the domain wall spatially separates fermions from antifermions in a way that fermions are going inside the bubble. The region outside the large circle corresponds to the undisturbed symmetric phase.

fermion scattering on the walls, leads to the separation of baryon number: fermions go inside the bubbles and antifermions outside. The outside excess of antifermions is destroyed in equilibrium sphaleron reactions, whereas the excess of fermions in the Higgs phase remains intact. As the region of the broken phase increases, at the time when different bubbles collide, the whole Universe is baryon-asymmetric.

The scenario described above does not work for the SM. The phase diagram for the electroweak theory is shown in Fig. 7. The vertical axis is the temperature, while the horizontal axis is the Higgs mass in the SM. The phase diagram in many extensions of the standard model looks qualitatively the same, but the horizontal axis for them is a combination of the different parameters rather than the Higgs mass itself.

This diagram is similar to that of a liquid–vapour system. There is an end-point of a line of the first-order phase transitions. If the Higgs mass is equal to the critical value, the phase transition in the system is of second order. For smaller Higgs masses the EW phase transition is of the first kind, while at larger Higgs masses there is no phase transition at all.

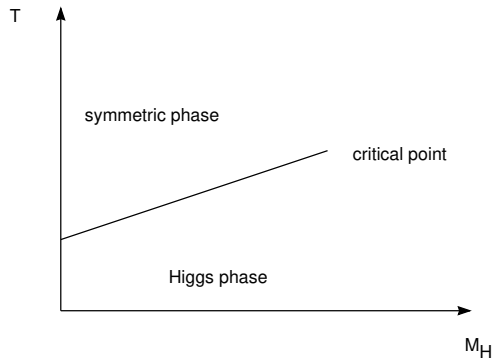

Fig. 7: The phase diagram of the electroweak theory. For small Higgs masses, to the left from the critical point, the electroweak phase transition is of the first kind. At the critical point, it is of the second kind, and the critical properties of the electroweak theory near this point are very similar to those of the liquid-vapour system. At large Higgs masses, to the right from the critical point, there is no phase transition at all.

For the SM, the endpoint of the first order phase transition line can be computed unambiguously and corresponds to $M_H \simeq 72$ GeV. With the experimental value of the Higgs mass 125 GeV, there is no phase transition at all. So, the SM can be excluded as a theory for domain wall electroweak baryogenesis. Extensions of the SM may be viable from the point of view of the strength of the phase transition. For instance, extra scalars in the mass range ~ 100 GeV may lead to the necessary strength of the phase transition.

Assume now that we have a model with a necessary first-order phase transition. What is the *magnitude* of the baryon asymmetry created? It depends on the particle content, the strength of CP-violation, the rate of B-non-conservation in the symmetric phase, the bubble wall velocity, and many other details of interactions. Computations of baryon excess have been carried out in different models, with the result that the asymmetry 10^{-10} may be derived, provided the parameters of a model are chosen in some particular way. In addition to constraints on the particle spectrum, following the requirement of a first-order phase transition, there appear extra constraints to the models, since a sufficient amount of CP violation is required.

12 Dark Matter

12.1 Evidence for Dark Matter

The existence of non-luminous and non-absorbing substance, dubbed dark matter, making up about 27% of the mass of the Universe is now an established fact. Our confidence in its existence has been built on many independent observations pointing to the same conclusion. Let us discuss some of them.

Rotational curves of spiral galaxies. Compelling evidence for dark matter is coming from the orbital velocities of stars located on the disk of a spiral galaxy. The mass distribution of luminous matter in such a galaxy can be inferred by measuring its luminosity as a function of radius. Observations show that the luminosity follows a rough radial dependence of the form

$$I(r) = I_0 \exp\left(-\frac{r}{r_0}\right) . \tag{85}$$

From here we can infer that most of the mass in a spiral galaxy is concentrated near its centre. If we assume that only luminous matter gravitates, the velocities of the stars on the disk at the outer side of the galaxy would fall off according to Kepler's law ($v \propto 1/\sqrt{r}$). Indeed, since most of the mass is clumped near the centre of the galaxy, stars located at different radii sufficiently far from the centre of the galaxy feel essentially the same amount of mass pulling them towards the centre. Mathematically,

$$mv^2 \approx G\frac{Mm}{r} \Longrightarrow v \approx \sqrt{\frac{GM}{r}} . \tag{86}$$

In reality, what we see is a flat profile

$$v \approx \text{constant} \tag{87}$$

suggesting that the M in (86) is not a constant but varies as

$$M(r) \propto r \tag{88}$$

which could be explained by the presence of a dark matter halo with mass distribution at large r

$$\rho_{\text{dark}}(r) \propto \frac{1}{r^2} . \tag{89}$$

Gravitational lensing. Gravitational lensing is a phenomenon where massive objects located between us and distant light sources act as a lens, bending the space around them and consequently bending any light passing nearby. When the effect only causes slight shear deformations in the images of distant luminous objects, it is called weak gravitational lensing. By measuring the amount of such shear deformations on distant galaxies, and combining it with appropriate statistical analysis, we can infer the total mass of the intervening galaxy clusters causing the deformations. This led to the same conclusion that the galaxy clusters are more massive than the total mass of the visible objects belonging to them.

Bullet Cluster. The Bullet Cluster is an aftermath of a collision between two galaxy clusters. The mass distribution of the colliding clusters can be inferred by gravitational lensing. It was found that the non-luminous matter simply passes through one another, showing that the dark matter has essentially no pressure. Furthermore, it was observed that the location of the centre of mass inferred from gravitational lensing is significantly displaced from that of the (luminous) baryonic matter, a fact that can be explained easily by the presence of dark matter.

Big Bang nucleosynthesis. The BBN predicts the abundance of light elements in the Universe as a function of the baryon-to-photon ratio η. Requiring the predictions to match with the observed values allows us to determine the baryon abundance $\Omega_B \simeq 0.05$. Numerically Ω_B is smaller than the

observed abundance Ω_M of non-relativistic matter in the Universe, which also suggests the existence of dark matter.

Combined data from SNe, BAO, and CMB. The combined data from supernovae (SNe), baryon acoustic oscillation (BAO), and the cosmic microwave background (CMB) constraint the abundances of non-relativistic matter Ω_M, non-relativistic matter in the form of baryons Ω_B, and the cosmological constant Ω_Λ to have the values in Equation (35). The fact that $\Omega_B < \Omega_M$ by a significant amount suggest the presence of non-baryonic matter, i.e. dark matter.

12.2 Constraints on dark matter particle

The SM does not provide any candidate for the non-baryonic dark matter and, therefore, cosmological observations point in the direction of physics beyond the standard model.

Lower bound on the mass of fermion DM particle. The Pauli exclusion principle gives an upper limit on how densely fermions can be packed in the phase space $(\mathbf{p}, \mathbf{x})$. Consequently, the number of fermions N_F with typical velocity v contained within an object of size r, e.g. a galaxy halo, is bounded from above

$$N_F \lesssim \frac{1}{(2\pi)^3} \int d^3\mathbf{p} d^3\mathbf{x} n \sim \left(\frac{pr}{2\pi}\right)^3 . \tag{90}$$

In the last step, we have assumed that n has the form of a window function with momentum and coordinate extent of p and r respectively. It follows that the total mass of fermions in a galaxy M is also bounded from above

$$M \lesssim m_\nu N_F \sim m_F^4 (vr/2\pi)^3 . \tag{91}$$

At the same time, Kepler's law states that $v^2 \sim GM/r$, which, when combined with the above bound, gives a lower limit on the masses of fermions in terms of measurable quantities v and r

$$m_F \gtrsim \left(\frac{(2\pi)^3}{Gvr^2}\right)^{1/4} \approx 120 \left(\frac{100\ \text{km/s}}{v}\right)^{1/4} \left(\frac{1\ \text{kpc}}{r}\right)^{1/2} . \tag{92}$$

This constraint is known as the Tremaire-Gunn limit. For example, in our galaxy $r \sim 10\,\text{kpc}$ and $v \sim 220\,\text{km/s}$, giving the fermion mass bound of $m_F \gtrsim 30\,\text{eV}$. A more stringent bound is given by dwarf spheroidal galaxies whose typical masses are around $M \sim 10^6 M_\odot$. Those galaxies give the fermion mass bounds of $m_F \gtrsim 500\,\text{eV}$. This bound, in particular, excludes ordinary neutrinos as dark matter candidates.

Lower bound on the mass of boson DM particle. The above argument does not work for bosons, as the number of bosons per one quantum state is not bounded. Still, the de Broglier wavelength $2\pi/(m_B v)$ of a DM particle with escape velocity v and mass m_B must be smaller than the size r of the compact DM object, such as a dwarf galaxy. This leads to a lower bound $m_B > 10^{-22}\,\text{eV}$. The DM so light is usually called "fuzzy" dark matter.

Cold, warm and hot dark matter. An important constraint on DM particles comes from the study of the structure formation. The analysis of the evolution of the density perturbations reveals that the structure formation starts at the matter radiation equality, at $T_{\text{eq}} \simeq 0.8\,\text{eV}$, corresponding to the redshift $z \simeq 3500$. The free-streaming length λ_{FS} of the DM particles at this time determines the

minimal mass of the structure that can be formed by the gravitational Jeans instabilities,

$$M \lesssim M_{\rm FS} \simeq \frac{4}{3}\pi\lambda_{\rm FS}^3\rho_{\rm DM}\,, \tag{93}$$

where the DM energy density $\rho_{\rm DM}$ is taken at $T_{\rm eq}$. The DM particle traverses the distance l after its decoupling determined from the equation

$$\frac{dl}{dt} = v(t) + Hl\,, \tag{94}$$

where $v(t)$ is the velocity of DM particle, and H is the Hubble rate. A relatively simple computation leads to

$$M_{\rm FS} \simeq 10^{10} M_\odot \left(\frac{7\ {\rm keV}}{M}\right)^3 \frac{g^*({\rm now})}{g^*({\rm decoupling})} \left(\frac{\langle p\rangle}{p_T}\right)^3\,, \tag{95}$$

where $M_\odot$ is the solar mass, $g^*({\rm now})$ and $g^*({\rm decoupling})$ are the number of effectively massless degrees of freedom now and at the moment of DM particle decoupling, $\langle p\rangle$ is the average momentum of the DM particle at decoupling, and $p_T \simeq 3.15T$ is the average thermal momentum of a fermion at this time. Equation (95) tells, in particular, that the dark matter particles must not be relativistic at the onset of structure formation, to admit the formation of structures such as dwarf galaxies.

The ordinary neutrino would be a hot DM particle, with the free streaming mass exceeding the mass of a cluster of galaxies $> 10^{14} M_\odot$. In this case, the smaller structures could not be formed, excluding once more the ordinary neutrino as a DM candidate. A WIMP (weakly interacting massive particle) with a mass, say, 100 GeV leads to a negligibly small free streaming mass, allowing the formation of structures at small scales. This a cold DM candidate. A 7 keV particle, lying somewhere between these two extremes, would be a "warm" DM matter candidate.

Known and unknown properties of dark matter. To summarise, the dark matter particles must be sufficiently stable with a lifetime that is longer than the age of the Universe, or otherwise, they would have decayed. If the dark matter particles are relatively light, $M \lesssim 1$ TeV, then they must be neutral and very weakly interacting, otherwise, we would have detected them. If the dark matter particles are fermions, their mass must be above 500eV (the Tremaine-Gunn bound we obtained above). If they are bosons, their mass should be above 10^{-22} eV. Also, the free streaming length of DM particles should be small enough in order not to get in conflict with structure formation.

The predictions of dark matter masses from different models vary wildly, they go from as light as 10^{-22} eV (stringy axions) to as heavy as 10^{24} GeV (such as supersymmetric Q-balls). The spin of the dark matter particle is not known. Since the presence of dark matter has been based entirely on its gravitational effects, not much is known about its non-gravitational interactions other than they must be very weak. For the same reason, the production mechanism of dark matter and how they are embedded in the big picture of particle physics are the subjects of speculation.

12.3 Sterile neutrino dark matter

It became fashionable to make vote at different conferences and workshops in favour of this or that DM candidate. At the moment, the first place is usually attributed to WIMP (weakly interacting massive

particle), the second to axion, and the third to sterile neutrino. The motivation for WIMPs stems from low-energy supersymmetry and, more generally, from naturalness and the hierarchy problem. The motivation for axion stems from the strong CP problem. The motivation for sterile neutrino comes from neutrino physics. The first two candidates were discussed in the special lecture on Dark Matter by Mads Frandsen (see also the non-shown slides 44-57 of my Lecture 2). In these lecture notes, we will only cover the third one.

Let us consider the theory described in Section 11.3.1, which is just the SM extended by 3 HNLs. As we have already explained, it can describe neutrino masses and oscillations and produce baryon asymmetry of the Universe. In addition, it also gives a suitable DM candidate. Though the νMSM does not have any extra stable particle in comparison with the SM, the lightest singlet fermion, N_1, may play the role of a dark matter particle as its lifetime can greatly exceed the age of the Universe (for a review see, e.g. [15]).

The following considerations determine a possible range of masses and couplings of the DM sterile neutrino:

- The sterile neutrino N_1 can decay via the mixing with the ordinary neutrino into 3 neutrinos, $N_1 \to \nu\nu\bar{\nu}$, $N_1 \to \bar{\nu}\bar{\nu}\nu$ with the width $\Gamma \propto G_F^2 m_N^2 \theta^2$ where m_N is the mass of the DM sterile neutrino, and $\theta = fv/m_N$ is the mixing angle. Here f is the Yukawa coupling of the sterile neutrino and v is the Higgs vacuum expectation value. The sterile neutrino lifetime must exceed the age of the Universe.
- Cosmological production. DM sterile neutrinos are produced in reactions like $l^+l^- \to N_1\nu$ at temperatures of the order of 100 MeV. If N_1 interact too strongly, these reactions would overproduce them, making the abundance of Dark Matter in the Universe larger than observed. This leads to an upper limit on the strength of interaction of N_1. If N_1 is produced only in the reactions of this type, the requirement to produce enough Dark Matter also results in the lower bound on the mixing angle. This lower bound depends on the conditions in the early Universe during the epoch of N_1 production [14], such as the lepton asymmetry of the Universe. Moreover, the lower bound completely disappears if N_1 can also be produced at very high temperatures by interactions related to gravity or at the end of cosmological inflation.
- X-rays. N_1 decays radiatively with the width $\Gamma_\gamma \propto \alpha G_F^2 m_N^2 \theta^2$, $N_1 \to \gamma\nu$, producing a narrow line that can be detected by X-ray telescopes (such as XMM-Newton or Chandra). These considerations result in an upper limit on the sterile neutrino mixing angle with active neutrinos. While this upper limit depends on the uncertainties in the distribution of dark matter in the Milky Way and the other nearby galaxies and clusters and the modelling of the diffuse X-ray background, it is possible to conservatively marginalise over these uncertainties, obtaining very robust constraints (see [15] and references therein).
- Structure formation. If N_1 is too light, a very large number density of such particles is required to make an observed halo of a small galaxy. As HNLs are fermions, their number density can not exceed that of a completely degenerate Fermi gas, resulting in a very robust lower bound $\simeq 0.5$ keV on the mass of the particle. This bound can be further improved by taking into account that, as discussed above, light-dark matter particles remain relativistic till quite late epochs and therefore

suppress or erase density perturbation on small scales. This would affect the inner structure of the halos of the Milky Way and other galaxies, as well as matter distribution in the intergalactic medium.

The summary of constraints is presented in Figure on slide 43 of Lecture 2. There the solid lines represent the most important constraints that are largely model-independent. The *phase space bound* (solid purple line) is based on Pauli's exclusion principle applied to dark matter in dwarf galaxies. The bounds based on the *non-observation of X-rays* from the decay $N_1 \to \nu\gamma$ are shown by the violet area. The dashed lines represent the estimates of the sensitivity of the future X-ray mission ATHENA. The blue square marks the interpretation of the 3.5 keV excess as decaying sterile neutrino DM. Above the line marked "thermal overproduction", the abundance of sterile neutrinos would exceed the observed DM density. All other constraints depend on the sterile neutrino production mechanism.

The upper limits on the strength of interaction of sterile neutrino allow us to fix the overall scale of active neutrino masses in the νMSM. The DM sterile neutrino effectively decouples from the see-saw formula, telling that the mass of one of the active neutrinos is much smaller than the observed solar and atmospheric mass differences. This fixes the masses of two other active neutrinos to $\simeq 0.009$ eV and $\simeq 0.05$ eV for the normal ordering and to the near degenerate value 0.05 eV for the inverted ordering.

The most promising way to find DM matter neutrino is indirect detection with the use of X-ray telescopes in Space. The new X-ray spectrometer XRISM, which was launched in September 2013, has great potential to detect a signal from Dark Matter decay.

13 Dark Energy

The cosmological observations suggest that the Universe is accelerating now. This is perfectly consistent with a non-zero positive cosmological constant, $\Omega_\Lambda \simeq 0.7$, though maybe a signal of the presence of some exotic substance – Dark Energy – with the equation of state close to that of the cosmological constant, namely $p = \omega\rho$, where experimentally $\omega = -1 \pm 0.1 \pm 0.1$ with the first error being systematic and the second being statistical. The observed value of Ω_Λ corresponds to the vacuum energy density

$$V_0 = \left(\frac{\Lambda}{8\pi G}\right)^{\frac{1}{4}} \sim 2 \times 10^{-3}\ \text{eV} \sim 0.01\ \text{cm}^{-1}\,. \tag{96}$$

The value of the cosmological constant in the SM plus gravity cannot be predicted, in the exact similarity with the other parameters of the SM, such as the mass of electrons or other particles. At the present state of knowledge, all these parameters are simply taken from the experiment. A naive computation of the vacuum energy density in the SM gives a divergent result without physical meaning, representing a sum of zero-point energies of all species of the SM.

A particular value of the cosmological constant poses several questions which remain without the answers. The first comes from a comparison between the magnitude of the scale (96) and other known scales, for example, Λ_{QCD}, M_W, and M_{Pl}. Why is the scale associated with Λ so small compared with the other scales? Another problem arises from the comparison of Ω_M and Ω_Λ. At the present stage of expansion of the Universe, they are of the same order of magnitude. Is this a pure coincidence or there is a deep reason behind that?

14 Inflation

14.1 Problems of standard cosmology

To explain what kind of problems the standard cosmology faced before the invention of inflation, we will introduce the notion of particle horizons. In a static Universe, if two events are separated by distance Δl and time Δt, they are causally independent, provided $\Delta l > \Delta t$. What is the analogue of this statement in an expanding Universe? To answer this, let us write the equation describing the propagation of light, taking into account the fact that the speed of light is one in the natural system of units:

$$\frac{dl}{dt} = 1 + \frac{\dot{a}}{a} l \,. \tag{97}$$

The solution of this equation is

$$l(t) = \int_{t_0}^{t} \frac{a(t)}{a(t')} dt' \,. \tag{98}$$

For both radiation- and matter-dominated Universes, with $a(t) \sim t^{1/2}$ and $a(t) \sim t^{2/3}$ respectively, the integral in eq. (98) converges even if $t_0 = 0$. The distance light travels since t_0 is called the particle horizon, $l_H(t)$. For different epochs, we have:

$$l_H(t) = \begin{cases} 2t, \text{ radiation-dominated epoch,} \\ 3t, \text{ matter-dominated epoch.} \end{cases} \tag{99}$$

If the distance between two points is greater than $l_H(t)$, the points were not in causal contact in the past and thus we should expect that the parameters of the Universe (such as temperature) should be different.

The horizon and homogeneity problem. As we have already discussed, the photons decoupled from the plasma at the redshift $z \simeq 1100$ corresponding to the time t_d. Let us assume that there were only radiation- and matter-dominated epochs in the past. Then, when looking at different points of the sky separated by some angle $\theta > \theta_H$ we would observe CMB emitted from regions that were never in causal contact and so should have different temperatures. To estimate θ_H, one should find the present size L of the region that was the horizon at the decoupling time,

$$L \sim 3t_d \frac{a_{\rm now}}{a_d} \sim 3t_d \left(\frac{t_{\rm now}}{t_d}\right)^{\frac{2}{3}} , \tag{100}$$

where $3t_d$ is the horizon scale at t_d. The angle θ_H is simply the ratio of this scale to the present size of the horizon,

$$\theta_H \simeq \frac{3t_d}{3t_{\rm now}} \left(\frac{t_{\rm now}}{t_d}\right)^{\frac{2}{3}} \simeq \left(\frac{t_d}{t_{\rm now}}\right)^{\frac{1}{3}} , \tag{101}$$

which corresponds to $\theta_H \sim 2^\circ$. This means that the present horizon contains $O\left(\frac{t_{\rm now}}{t_d}\right) \sim 10^4$ domains that were not in causal contact before recombination, see Fig. 8. However, observations show that the cosmic microwave background is isotropic for these angles, with accuracy better than 10^{-4}. This is the essence of the horizon and homogeneity problem.

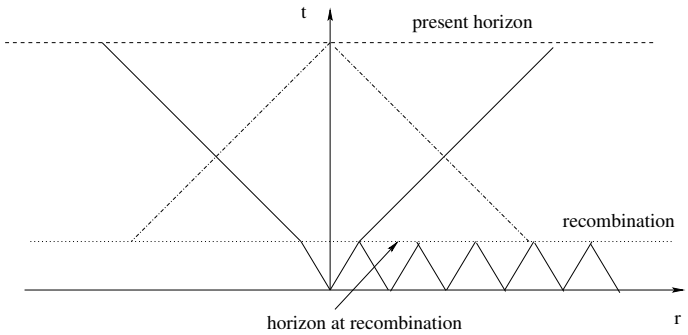

Fig. 8: An observer at point $r = 0$ sees many particle horizons corresponding to photon decoupling.

The flatness problem. Let us consider the relationship (34) in somewhat different form,

$$\Omega - 1 = \frac{k}{a^2 H^2} \,, \tag{102}$$

where Ω is the ratio of the total energy density to the critical density. For both matter- and radiation-dominated epochs, $H \sim \frac{1}{t}$ and $a \sim t^\alpha$, with $\alpha = \frac{1}{2}$ or $\alpha = \frac{2}{3}$. Thus, Ω increases with t as $\Omega - 1 \sim t^{2(1-\alpha)}$. Therefore, to have $\Omega \simeq 1$ at present, as follows from observation, Ω must have been finely tuned to one with huge accuracy in the past. For example, at the nucleosynthesis time, $|\Omega - 1|$ must be of the order of 10^{-15}. It is unclear why the Universe should have been so flat in the past.

To summarise, the observational fact that the present Universe is flat, homogeneous and isotropic is very bizarre. **If** the Universe was dominated by matter or radiation in the past, then the initial conditions for expansion must be highly fine-tuned. The Universe must have been much flatter than it is now, and the causally disconnected regions must have had the same characteristics such as densities, temperatures, etc. The question arises is there any rationale behind these fine-tuned initial conditions? A possible answer, associated with cosmological inflation, is discussed below.

14.2 Inflation as a solution of cosmological problems

The three problems described above are related to each other. The inflationary paradigm provides a simultaneous solution to all of them.

First, let us note that the assumption that the Universe was dominated by radiation or by matter well in the past does not follow from any observation. We can only be sure that the Universe was dominated by radiation at the epoch of BBN, perhaps at the time of baryogenesis at $T \sim 100$ GeV, but what happened before that is not known. Suppose that, for some reason, the dependence of a on t before baryogenesis was such that the integral in (98) is very large and the factor aH increases with time rather than decreases. Then the problem of initial conditions could be solved naturally.

To get a better grasp of the idea, let us see how it works explicitly. Suppose that the Universe was dominated by the vacuum energy at some point before the baryogenesis, BBN, and photon decoupling. As we found earlier, a vacuum-energy-dominated Universe expands as

$$a(t) = a_0 \exp\left[H(t - t_0)\right] \tag{103}$$

with $H = \sqrt{8\pi G\rho_\Lambda/3} = \text{const}$, where ρ_Λ is the vacuum energy density dominating the energy budget of the Universe during the period of inflation. Then, at $t = t_1$ the vacuum-energy dominated epoch ends with a transition to the radiation-domination epoch. After that, the scale factor continues to evolve as

$$a(t) = a_0 \exp\left[H(t_1 - t_0)\right] \left(\frac{t}{t_1}\right)^{1/2} . \tag{104}$$

Hence, using (98) we can calculate the horizon size at photon decoupling as

$$\begin{aligned} \ell_H(t_d) &= \int_{t_0}^{t_1} dt' \exp\left[H(t - t')\right] + \int_{t_1}^{t_d} dt' \left(\frac{t}{t'}\right)^{1/2} \\ &= \frac{1}{H} \exp\left[H(t_1 - t_0)\right] + 2(t_d - t_1) \\ &\sim \frac{1}{H} \exp\left[H(t_1 - t_0)\right] , \end{aligned} \tag{105}$$

where in the last step we have assumed that the period of inflation was sufficiently long that $H(t_1 - t_0) \gg 1$ and $H^{-1} \exp\left[H(t_1 - t_0)\right] \gg 2(t_d - t_1)$. If, for example, $\rho_\Lambda \sim \left(10^{15}\ \text{GeV}\right)^4$ (GUT scale), the horizon problem is solved if the particle horizon at the time of photon decoupling as seen today $\ell_H(t_d) a_0/a_d$ is at least as large as the present Hubble radius H_0^{-1}. The latter requirement gives a lower bound on the number of e-foldings during inflation $H(t_1 - t_0) \gtrsim 65$. The flatness problem is also automatically solved since

$$\left.\frac{\kappa}{a^2 H^2}\right|_{t=t_0} \propto \kappa \exp\left[-2H(t_1 - t_0)\right] \tag{106}$$

becomes exponentially suppressed if $H(t_1 - t_0) \gg 1$. The time-dependence of κ/a^2H^2 is, of course, more complicated than displayed above, but the point is that as long as $H(t_1 - t_0) \gg 1$ the exponential suppression will most likely be the dominant factor.

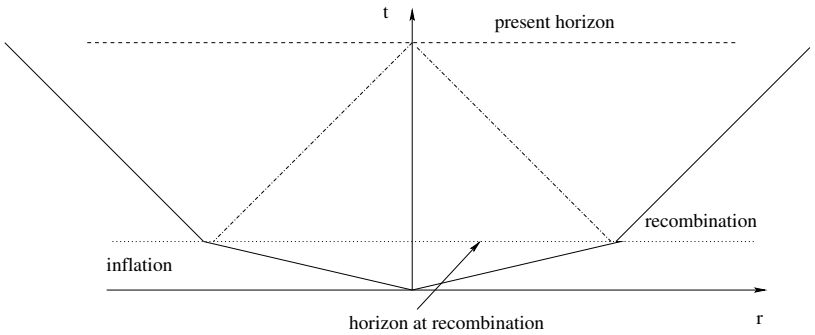

Fig. 9: An observer at point $r = 0$ sees just one particle horizon corresponding to photon decoupling.

14.3 Chaotic inflation

There are many different particle-physics models of inflation. Most of them are associated with the dynamics of single or multiple scalar fields. We refer to a comprehensive review [16] and describe in some detail just one possibility which is called "chaotic inflation".

Consider a theory of a single free scalar field in a curved background with an action

$$S = \int d^4x\sqrt{-g}\left(\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - U(\phi)\right) , \tag{107}$$

where the potential is

$$U(\phi) = \frac{1}{2}m^2\phi^2 . \tag{108}$$

We will assume that $m^2 \ll M_{Pl}^2$.

We do not know how to describe the state of the Universe at the Planck scale, since the classical theory of gravity is not applicable there. Nevertheless, it is natural to assume that at Planck time there were fluctuations in the scalar field with energy density

$$\epsilon \sim \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + U(\phi) \sim M_{Pl}^4 . \tag{109}$$

Suppose that there is a sufficiently large region of space where fluctuations of potential energy dominate, i.e.

$$U(\phi) \sim M_{Pl}^4 \gg (\nabla\phi)^2 \text{ and } \dot{\phi}^2 . \tag{110}$$

In these regions, the value of ϕ is much larger than the Planck scale,

$$\phi \simeq \frac{M_{Pl}^2}{m} \gg M_{Pl} , \tag{111}$$

and the scalar field is nearly homogeneous so that the equation representing its evolution is simply

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dU(\phi)}{d\phi} = 0 , \tag{112}$$

where

$$H^2 = \frac{8\pi G}{3}\left(\frac{\dot{\phi}^2}{2} + U(\phi)\right) . \tag{113}$$

This looks like an equation of motion of a non-relativistic particle with unit mass in the potential $U(\phi)$ with a friction term that depends on the position and velocity of the particle, see Fig. 10. For large values of ϕ the regime is overdamped, with $H \gg \ddot{\phi}/\dot{\phi}$, where

$$H^2 \simeq \frac{4\pi m^2\phi^2}{3M_{Pl}^2} . \tag{114}$$

Thus, eq. (112) has the form

$$\frac{\sqrt{12\pi}m\phi\dot{\phi}}{M_{Pl}} + m^2\phi = 0 \tag{115}$$

and has a solution

$$\phi \simeq \phi_0 - \frac{mM_{Pl}t}{\sqrt{12\pi}} \simeq \phi_0\left(1 - O\left(\frac{m^2t}{M_{Pl}}\right)\right) . \tag{116}$$

The “slow-roll-down" approximation breaks down at $\dot{\phi}^2 \sim V(\phi)$, where $\phi \simeq M_{Pl}$, $t \simeq M_{Pl}/m^2$.

Before this time the Universe expands exponentially and the scale factor changes by

$$\exp\left(Ht\right) \simeq \exp\left(\frac{M_{Pl}^2}{m^2}\right) \gg 1\,. \tag{117}$$

During exponential expansion, the non-homogeneities are red-shifted away and the Universe becomes practically uniform at cosmological distances.

After time $t \simeq M_{Pl}/m^2$ inflaton oscillates near the origin, transferring its energy to other particles. This process is usually called reheating. For $m \ll M_{Pl}$ this simple model of scalar field solves horizon,

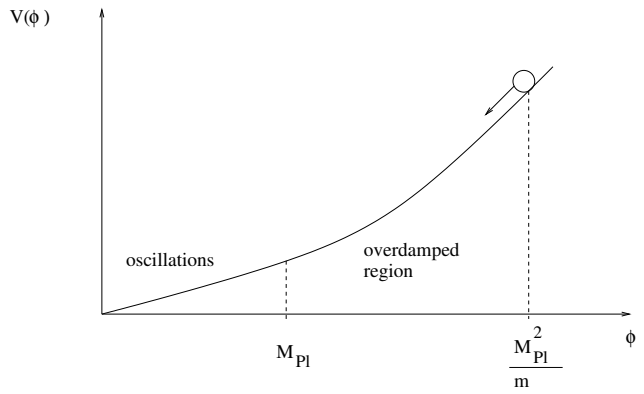

Fig. 10: Inflaton dynamics as the motion of a particle in potential $V(\phi)$ with friction term depending on ϕ.

homogeneity and flatness problems. It does not survive, though, a more elaborated comparison with the data concerning the spectrum of primordial density perturbations, which eventually lead to structure formation in the Universe, see below. However, replacing a free field theory with a more complicated "inflationary" potential allows us to address all the issues.

14.4 Density perturbations from inflation

The paradigm of inflation was originally motivated by solutions to the horizon and flatness problems. It was realised soon after that, the quantum aspect of the same idea could provide the appropriate initial conditions for the primordial density perturbation.

Qualitatively, inflation seeds the density perturbations in the Universe in the following way. Quantum fluctuations $\delta\phi$ in the inflaton field result in adiabatic density perturbation $\delta\rho(\mathbf{x})$ on top of the homogeneous background density ρ. Existing initially at small distances, perturbations are inflated to macroscopic scales due to the exponential expansion of the Universe at the inflationary stage. By definition, the average

$$\left\langle \frac{\delta\rho(\mathbf{x})}{\rho} \right\rangle = 0\,, \tag{118}$$

but

$$\left\langle \frac{\delta\rho(\mathbf{x})}{\rho}\,\frac{\delta\rho(\mathbf{y})}{\rho} \right\rangle = f(|\mathbf{x}\text{-}\mathbf{y}|) \neq 0\,, \tag{119}$$

Note that f depends only on the $|\mathbf{x}-\mathbf{y}|$ as a consequence of isotropy and homogeneity. Due to inflation, for $\mathbf{x}$ and $\mathbf{y}$ located within a causal domain $|\mathbf{x}-\mathbf{y}| \lesssim \ell_H(t_0)$, the correlation function should be close to constant

$$\left\langle \frac{\delta\rho(\mathbf{x})}{\rho} \frac{\delta\rho(\mathbf{y})}{\rho} \right\rangle \simeq \text{const}\,. \tag{120}$$

As it is impossible to have a perfectly constant Hubble parameter during inflation, there is a slight tilt in the correlation function (the Fourier transform of it defines the so-called power spectrum of primordial fluctuations)

$$f(\mathbf{x}) \propto |\mathbf{x}|^{1-n_s}\,, \tag{121}$$

where $n_s \neq 0$ is the so-called spectral index of the scalar perturbations. A scale-invariant spectrum with $n_s = 0$ is called the Harrison-Zeldovich power spectrum.

In addition to scalar perturbations $\delta\rho/\rho$, inflation also predicts tensor perturbations. It is common to measure their amplitude relative to the size of the scalar perturbation in terms of the tensor-to-scalar ratio

$$r = \frac{\rho_2}{\rho_0}\,, \tag{122}$$

where ρ_2 is the energy of spin-2 fluctuations (associated with the gravitational waves created during inflation) and the energy in the scalar fluctuations. These two parameters, n_s and r depend on the model of inflation.

To explain the CMB temperature and polarisation fluctuations, inflation gives as input the form of the power spectrum characterised by n_s and tensor to scalar ratio r. The initial density perturbations get modified by various "ordinary physics" effects, e.g. photon scattering, gravitational lensing, plasma waves, etc that we know how to incorporate. Once these effects are accounted for, one can obtain, for instance, a prediction for the present form of temperature fluctuation $\left\langle \frac{\delta T(\mathbf{x})}{T} \frac{\delta T(\mathbf{y})}{T} \right\rangle$ which we can compare with CMB measurements. The predicted form depends on many parameters, e.g. n_s, r, Ω_B, Ω_ν, etc, enabling themselves to be determined by measuring the CMB temperature and polarisation anisotropies. See Figure at slide 18 of Lecture 3.

The whole picture, called ΛCDM (for Lambda-Cold Dark Matter cosmology) works remarkably well, with only a few fitting parameters describing a huge amount of cosmological data (but several tensions exist, for instance, the Hubble rate measurements at small and large redshifts give different results).

14.5 Particle physics models for inflation

What kind of particle inflated the Universe? Is this a new particle or a part of the Standard Model? The answer is: “We do not know”. There is a huge variety of inflationary models based on different principles and ideas, for a review see [16]. The mass of the inflaton may vary from hundreds of MeV to, say, 10^{10} GeV. The minimal possibility is that the Higgs field of the Standard Model inflated the Universe is in perfect agreement with the data on the spectral tilt n_s and tensor to scalar ratio r (see discussion on slides 19-22 of Lecture 3). To single out the inflationary model we need a theory input in the form of the complete theory of gravity and the SM, together with an even more precise determination of the cosmological parameters allowing us to discriminate the different models.

15 Conclusions

The remarkable progress in particle physics and cosmology resulted in the creation of two standard models – the Standard Model in particle physics and ΛCDM in cosmology, both consistent with the majority of observations in respective domains of physics. There is a clash between these two theories, though. Indeed, only a small part $\sim 5\%$ of the energy density of the Universe can be described by particles of the Standard Model - baryons. Even this number, associated with the baryon asymmetry of the Universe, cannot be explained by the physics of the SM. Another $\simeq 70\%$ can be attributed to the cosmological constant or "Dark Energy" with the equation of state close to that of the cosmological constant. The remaining $\simeq 25\%$ represents dark matter, which cannot be the SM particles. The SM, where neutrinos are massless and neutrino flavours are conserved, is also challenged by neutrino physics.

There are many theoretical ideas in particle physics and cosmology which address the drawbacks of the SM. The structure of the Universe at large scales and density perturbations are neatly explained by inflation. The problems of neutrino masses, baryon asymmetry of the Universe and Dark Matter can be solved in different extensions of the SM. Hopefully, new observations in cosmology and new experiments in particle physics will narrow the search for the theory which provides a better description of Nature.

References

[1] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley and Sons, New York, 1972.

[2] Y. B. Zeldovich and I. D. Novikov, *Relativistic Astrophysics. vol. 2. The Structure and Evolution of the Universe*. 1983.

[3] E. W. Kolb and M. S. Turner, *The Early Universe*, vol. 69. 1990.

[4] V. A. Rubakov and D. S. Gorbunov, *Introduction to the Theory of the Early Universe: Hot big bang theory*. World Scientific, Singapore, 2017.

[5] D. S. Gorbunov and V. A. Rubakov, *Introduction to the theory of the early universe: Cosmological perturbations and inflationary theory*. 2011.

[6] V. Mukhanov, *Physical Foundations of Cosmology*. Cambridge University Press, Oxford, 2005.

[7] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, "Progress in electroweak baryogenesis," *Ann. Rev. Nucl. Part. Sci.* **43** (1993) 27–70, arXiv:hep-ph/9302210.

[8] S. Davidson, E. Nardi, and Y. Nir, "Leptogenesis," *Phys. Rept.* **466** (2008) 105–177, arXiv:0802.2962 [hep-ph].

[9] D. E. Morrissey and M. J. Ramsey-Musolf, "Electroweak baryogenesis," *New J. Phys.* **14** (2012) 125003, arXiv:1206.2942 [hep-ph].

[10] L. Canetti, M. Drewes, and M. Shaposhnikov, "Matter and Antimatter in the Universe," *New J. Phys.* **14** (2012) 095012, arXiv:1204.4186 [hep-ph].

[11] A. D. Dolgov, "NonGUT baryogenesis," *Phys. Rept.* **222** (1992) 309–386.

[12] A. Riotto and M. Trodden, "Recent progress in baryogenesis," *Ann. Rev. Nucl. Part. Sci.* **49** (1999) 35–75, arXiv:hep-ph/9901362.

[13] R. Allahverdi and A. Mazumdar, "A mini review on Affleck-Dine baryogenesis," *New J. Phys.* **14** (2012) 125013.

[14] A. Boyarsky, O. Ruchayskiy, and M. Shaposhnikov, "The Role of sterile neutrinos in cosmology and astrophysics," *Ann. Rev. Nucl. Part. Sci.* **59** (2009) 191–214, arXiv:0901.0011 [hep-ph].

[15] A. Boyarsky, M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy, "Sterile neutrino Dark Matter," *Prog. Part. Nucl. Phys.* **104** (2019) 1–45, arXiv:1807.07938 [hep-ph].

[16] J. Martin, C. Ringeval, and V. Vennin, "Encyclopædia Inflationaris," *Phys. Dark Univ.* **5-6** (2014) 75–235, arXiv:1303.3787 [astro-ph.CO].