

INTRODUCTION TO COSMOLOGY

A.D. Dolgov

ITEP, 117218, Moscow, Russia
INFN, Ferrara 40100, Italy
University of Ferrara, Ferrara 40100, Italy

ITEP Winter School
Moscow

February 9-14, 2009

Abstract

An introductory lectures on cosmology at ITEP Winter School for students specializing in particle physics are presented. Many important subjects are not covered because of lack of time and space but hopefully the lectures may serve as a starting point for further studies.

1 Introduction

Modern cosmology is vast interdisciplinary science and it is impossible to cover it in any considerable detail in five hours allocated to me at this School. The task is even more difficult because of different background and level of the participants. Planning these lectures, I have prepared the following short list of subjects, which is surely will be made much shorter at this lecture course, but hopefully it may be useful for the students who would like to continue studying this field. So the idealistic content could be the following:

1. A little about general relativity and its role in cosmology.
2. Four basic cosmological equations and expansion regimes.
3. Universe today and in the past.
4. Kinetics in hot expanding world and freezing of species.
5. Inflation: kinematics, models, universe heating, and generation of density perturbations and gravitational waves.

6. Big bang nucleosynthesis.
7. Field theory at non-zero temperature and cosmological phase transitions.
8. Baryogenesis and cosmological antimatter.
9. Neutrino in cosmology (bounds on mass, oscillations, magnetic moment, and anomalous interactions).
10. Dark matter and large scale structure (LSS).
11. Vacuum and dark energies.
12. Cosmic microwave radiation (CMB) and cosmological parameters.

In reality about a half of this plan was fulfilled. At least this lectures could be helpful for a first acquaintance with cosmology and as starting point for deeper studies.

We will start from some non-technical introduction to General Relativity and relations between the latter and cosmology, sec. 2. Next we will derive the basic cosmological equations in a rather naive way studying motion of non-relativistic test body in spherically symmetric gravitational field, sec. 3. There we also talk about realistic regimes of the universe expansion and basic cosmological parameters. In the next section, 4 the universe history is very briefly presented. Section 5 is dedicated to thermodynamics and kinetics in the early universe. Section 6 is dedicated to freezing of species and cosmological limit on neutrino mass. Big bang nucleosynthesis is presented in sec. 7. In section 8 the role of neutrinos in BBN is described. Neutrino oscillations in the early universe are considered in sec. 9. In section 10 inflationary cosmology is discussed, and the last section 11 is dedicated to cosmological baryogenesis.

2 Gravity and cosmology

Two simple observations that the sky is dark at night and that there are shining stars lead to the conclusion that the universe is finite in space and time. The first one is the well known Olbers' paradox, based on the estimate of the sky luminosity, which in infinite homogeneous static universe must be infinitely high. Shining stars should exhaust their fuel in finite time and thus cannot exist in the infinitely old universe – thermal death of the universe. General relativity (GR) successfully hit both targets leading to the notion of expanding universe of finite age, but created instead its own very interesting problems which we discuss in what follows.

Newtonian theory of gravity has an evident shortcoming that it has action-at-a-distance property. In other words, gravitation acts instantaneously, at any distance. On the other hand, in the spirit of contemporary wisdom interactions are always mediated by some bosonic fields and are relativistically invariant. If we wished today to generalise Newtonian theory of gravity to relativistic theory we could take, *a priori* as a mediator of interactions scalar, vector, or tensor intermediate bosons, confining ourselves to lower spins.

Since we know that gravity operates at astronomically large distances, the mass of the intermediate boson should be zero or very small. Indeed, massless bosons create static Coulomb type potential, $U \sim 1/r$, while massive bosons lead to exponentially cut-off Yukawa potential, $U \sim \exp(-mr)/r$.

Interactions mediated by vector field are odd with respect to charge parity transformation, C-transformation, and as one can see from the vector boson propagator, such interactions induce matter-antimatter attraction and matter-matter repulsion, recall electromagnetic interactions. Hence vector field cannot mediate attractive gravitational force.

Scalar and tensor mediators lead to attraction of matter-matter and matter-antimatter and both are *a priori* allowed. According to non-relativistic Newtonian theory the source of gravity is mass. Possible relativistic generalisation for scalars should be a scalar quantity coinciding in non-relativistic limit with mass. The only known such source is the trace of the energy-momentum tensor of matter, T^μ_μ . The relativistic equation of motion for scalar gravity should have the form:

$$\partial^2 \Phi = 8\pi G_N T^\mu_\mu, \quad (1)$$

where G_N is the Newtonian gravitational coupling constant. Such theory is rejected by the observed light bending in gravitational field, since for photons: $T^\mu_\mu = 0$. A small admixture of scalar gravity to tensor one, i.e. Brans-Dicke theory [1], is allowed.

There remains massless tensor theory with the source which may be only the energy-momentum tensor of matter, $T_{\mu\nu}$. In first approximation the equation of motion takes the form:

$$\partial^2 h_{\mu\nu} = 8\pi G_N T_{\mu\nu}. \quad (2)$$

This equation is valid in the weak field approximation because the energy-momentum of $h_{\mu\nu}$ itself should be included to ensure conservation of the *total* energy-momentum.

Massless particles, as e.g. gravitons, must interact with a conserved source. Otherwise theory becomes infrared pathological. The energy-momentum tensor of matter is conserved only if the energy transfer to gravitational field is neglected. Taking into account energy leak into gravity leads to non-linear equations of motion and allows to reconstruct GR order by order. For a discussion of this approach see papers [2].

Historically Einstein did not start from field theoretical approach but formulated general relativity in an elegant and economical way as geometrical theory postulating that matter makes space-time curved and that the motion of matter in gravitational field is simply free fall along geodesics of this curved manifold. This construction is heavily based on the universality of gravitational action on all types of matter – the famous equivalence principle, probably first formulated by Galileo Galilei. The least action principle for GR was formulated by Hilbert with the action given by

$$A = \frac{1}{16\pi G_N} \int d^4x \sqrt{-g} R + A_m, \quad (3)$$

where R is the curvature scalar of four dimensional space-time and A_m is the matter action, written in arbitrary curved coordinates. Gravitational field is identified with the

metric tensor, $g_{\mu\nu}$, of the curved space-time. The curvature is created by matter through equations of motion:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G_N T_{\mu\nu}, \quad (4)$$

where $R_{\mu\nu}$ is the Ricci tensor. There is no space here to stop on technicalities of Riemann geometry. A good introduction can be found e.g. in book [3] where one can find definition and properties of the Christoffel symbols, $\Gamma_{\mu\nu}^\alpha$, Riemann tensor, $R_{\mu\alpha\nu\beta}$, Ricci tensor, $R_{\mu\nu} = g^{\alpha\beta}R_{\mu\alpha\nu\beta}$, scalar curvature, $R = g^{\mu\nu}R_{\mu\nu}$, covariant derivatives in curved space-time, D_μ , etc.

The source of gravity is the energy-momentum tensor of matter taken in this curved space-time:

$$T_{\mu\nu} = 2\delta A_m / \delta g^{\mu\nu}. \quad (5)$$

The impact of gravity on matter is included into $T_{\mu\nu}$ due to its dependence on metric and in some more complicated cases on the curvature tensors. Let us repeat that the motion of matter in the gravitational field is simply the free fall, i.e. motion along geodesics.

Classical tensor theory of gravity agrees with all available data and is a self-consistent, very beautiful and economic theory. It is essentially based on one principle of general covariance, which is a generalisation of Galilei principle of relativity to arbitrary coordinate frames. Invariance with respect to general coordinate transformation (which is called general covariance) is a natural framework which ensures vanishing of the graviton mass, m_g . Even if the underlying classical theory is postulated to be massless, quantum corrections should generally induce non-zero mass if they are not prevented from that by some symmetry principle. This is another advantage of tensor gravity with respect to scalar one for which no principle which forbids non-zero mass is known. Though quantum gravity is not yet understood, it is natural to expect that quantum corrections should induce $m_g \neq 0$ in absence of general covariance.

An important property of equations of motion (4) is that their right hand side is covariantly conserved:

$$D_\mu \left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \right) \equiv 0. \quad (6)$$

Accordingly the energy-momentum tensor must be conserved too:

$$D_\mu T_\nu^\mu = 0. \quad (7)$$

Here, D_μ is covariant derivative, as we have already mentioned. To those not familiar with Riemann geometry it may be instructive to mention that covariant derivative appears when one differentiates in curved coordinate system, e.g. in spherical one, even in flat space-time. From another point of view, covariant derivative in curved space-time, which, e.g. is acting on vector field, looks as

$$D_\mu V_\nu = \partial_\mu V_\nu - \Gamma_{\mu\nu}^\alpha V_\alpha. \quad (8)$$

It is similar to covariant derivative in gauge theories, because the latter includes gauge field, A_μ analogous to $\Gamma_{\mu\nu}^\alpha$.

According to the Noether theorem, the conservation of $T_{\mu\nu}$ follows from the least action principle if the matter action is invariant with respect to general coordinate transformation. So the gravitational (Hilbert) part of the action and the matter part lead to self-consistent equations of motion only if general covariance is maintained.

There is a deep analogy between the Einstein gravity and Maxwell electrodynamics. The Maxwell equations have the form:

$$\partial_\mu F^{\mu\nu} = 4\pi J^\nu \quad (9)$$

Owing to anti-symmetry of $F^{\mu\nu}$, the l.h.s. is automatically conserved:

$$\partial_\mu \partial_\nu F^{\mu\nu} \equiv 0, \quad (10)$$

so the current must be conserved too:

$$\partial_\mu J^\mu = 0. \quad (11)$$

These two conditions are consistent due to gauge invariance of the total electromagnetic action with matter included.

Einstein was the first who decided to apply GR equations (4) to cosmology in 1918 and was very much disappointed to find that the equations do not have static solutions. So an advantage of GR was erroneously taken as a shortcoming. Only after the Friedman solution in 1922 which predicted the cosmological expansion [4] and the Hubble discovery of the latter in 1929 [5], the idea that our world is not stationary and may have a finite life-time was established.

The distribution of matter in the universe is assumed to be homogeneous and isotropic, at least in the early stage, as indicated by isotropy of cosmic microwave background radiation (CMB), and even now at large scales. Correspondingly the metric can be taken as homogeneous and isotropic one (FRW metric [6]):

$$ds^2 = dt^2 - a^2(t) [f(r)dr^2 + r^2 d\Omega], \quad (12)$$

where the function $f(r)$ describes 3D space of constant curvature, $f(r) = 1/(1 - kr^2)$.

The evolution of the scale factor $a(t)$, i.e. the expansion law, is determined by the Friedman equations, which follow from the general GR ones for the FRW ansatz. The derivation is straightforward but quite tedious. We will derive them in the next section in very simple but not rigorous way. The derivation may be taken as a mnemonic rule to recall the equations in one-two minutes.

3 Cosmological expansion

3.1 Basic cosmological equations

Here we will present an oversimplified derivation of the Friedman equations. Though the arguments are subject to criticism, the final results are correct. Let us consider a

test body on the surface of homogeneous sphere with radius $a(t)$ and the energy density ρ . The energy conservation condition for the non-relativistic test particle reads $v^2/2 = G_N M/a + \text{const}$ where $M = 4\pi a^3 \rho/3$. It can be rewritten as:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi \rho G_N}{3} - \frac{k}{a^2} \quad (13)$$

This is one of the main cosmological equations. Here is the famous Hubble expansion law, that the object situated at distance d runs away from us with velocity proportional to the distance, $v = Hd$. Notice that for $d > 1/H$ it runs away with superluminal velocity. Superluminal velocities of distant objects are allowed by GR but locally velocities must be always smaller or equal to the speed of light.

Another equation follows from the energy balance of the medium inside the sphere: $dE = -P dV$ where $E = \rho V$ and $dE = V d\rho + 3(da/a)V\rho$. Hence:

$$\dot{\rho} + 3H(\rho + P) = 0. \quad (14)$$

This equation is simply the law of covariant energy-momentum conservation (7) in metric (12).

Problem 1. Derive from eqs. (13) and (14) the law for the acceleration of the test body:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_N}{3}(\rho + 3P) \quad (15)$$

A striking feature of equation (15) is that not only energy but also pressure gravitates. It is always assumed in canonical theory that ρ is positive, though pressure may be negative. Thus if $\rho + 3P < 0$, the cosmological expansion would proceed with acceleration, i.e. antigravity may operate in cosmological scales. In other words, negative pressure is the source of the cosmological expansion. Life is possible only because of that. We believe that the universe was in such anti-gravitating state at the very beginning, during the so called inflationary stage (see below). Surprisingly it was established during the last decade that at the present time the expansion is also accelerating. There existed a simple analogy between the universe expansion and the motion of a stone thrown up from the Earth with some initial velocity v_0 . The speed of the stone drops down with time and it either will return back to the Earth, if v_0 is smaller than a certain value, v_1 . If $v_0 > v_1$ the stone will never come back. In the last case the stone will either come to infinity with non-zero speed or with the vanishing one. All three such regimes could exist in cosmology and they were believed to be realised, depending upon the initial expansion velocity. The first regime corresponds to the closed universe with $\rho > \rho_c$, where ρ_c is the critical or closure energy density, see below. The expansion in this case will ultimately turn into contraction. The other two regimes correspond to open universe which was expected to expand forever. The third one with zero velocity at infinity corresponds to spatially flat universe, with $k = 0$ in eq. (13). Now the picture is very much different. Imagine that you have thrown a stone from the Earth and first the stone moves with normal negative acceleration and after a while starts to move faster and faster as if it has a rocket engine. This is exactly what we see in the sky now. It means, in particular, that the spatially

closed universe may expand forever. This accelerated cosmological expansion, induced by antigravity at cosmological scale is prescribed to existence of mysterious dark energy. It is one the greatest unsolved problems in modern fundamental physics.

Problem 2. We seemingly started from the Newtonian theory but came to the conclusion that pressure gravitates which is not the true in Newtonian case. Where is the deviation from Newton?

Problem 3. Prove that for positive definite energy density, $\rho > 0$, any object of finite size creates an attractive gravitational force, even if inside such an object pressure may be arbitrary negative.

Problem 4. Prove that any finite object with positive energy density gravitates, so anti-gravitational action of pressure can manifest itself only in infinitely large objects.

Except for equations which determine the law of the cosmological expansion, we need an equation which governs particle propagation in FRW metric, i.e. geodesic equation. The latter can be written as:

$$\frac{dV^\alpha}{ds} = -\Gamma_{\mu\nu}^\alpha V^\mu V^\nu + \text{curvature term}, \quad (16)$$

where $V^\alpha = dx^\alpha/ds$ and the curvature term is absent in spatially flat universe, when $k = 0$. In what follows we will consider only this case, moreover, the effects of curvature are typically small.

To solve this equation one needs first to calculate the Christoffel symbols for metric (12). In 3D flat space they have very simple form:

$$\Gamma_{jt}^i = H\delta_j^i, \quad \Gamma_{ij}^t = Ha^2\delta_{ij}. \quad (17)$$

All other are zero. After that the geodesic equation takes a very simple form:

$$\dot{p} = -Hp \quad (18)$$

with an evident solution $p \sim 1/a(t)$ which describes red-shifting of momentum of a free particle moving in FRW background. In derivation of this result one has to pay attention that physical momentum is defined with respect to physical length $dl = a(t)dx$.

We can simply derived the same result taking into account the Doppler red-shift of the momentum of a free particle induced by the cosmological expansion. Let us take two points A and B separated by distance dl . The relative velocity of these two points due to expansion is $U = Hdl$. The Doppler shift of the momentum of the particle moving from A to B with velocity $v = dl/dt$ is

$$dp = -UE = -HEdl. \quad (19)$$

Thus

$$\dot{p} = -HEdl/dt = -Hp. \quad (20)$$

Let introduce at this stage the notion of the cosmological red-shift:

$$z = a(t_U)/a(t) - 1, \quad (21)$$

where t_u is the universe age and $a(t_U)$ is the value of the scale factor today. So the momentum of a free particles drops down in the course of expansion as $p \sim 1/a \sim 1/(z+1)$.

Equations (13-15) are the basic cosmological equations for three unknowns, a , ρ , and P . However, there are only two independent equations. So it is necessary to have one more equation describing properties of matter i.e. the equation of state (e.o.s.): $P = P(\rho)$. Usually this equation is parametrized in the simple linear form:

$$P = w\rho. \quad (22)$$

The parameter w determines matter properties. For non-relativistic matter pressure is negligibly small in comparison with ρ and in a good approximation we can take $w = 0$. For relativistic matter $w = 1/3$. There is also one more type of matter (or vacuum) known to exist in the universe, for which $w = -1$.

However, sometimes the equation of state does not exist but the necessary additional relation (not e.o.s.) can be derived from the equations of motion. E.g. for a scalar field:

$$D^2\phi + U'(\phi) = 0 \quad (23)$$

and one can calculate $T_{\mu\nu}$ and find ρ and P but $P \neq P(\rho)$.

Problem 5. Calculate $T_{\mu\nu}(\phi)$, ρ , and P for homogeneous field $\phi(t)$.

3.2 Expansion regimes

Here we will present solutions of the cosmological equations for several special cases which were/are realised in the universe at different stages of her evolution. We always assume that the three dimensional space is flat, i.e. $k = 0$. As we see below it is true during practically all life-time of the universe.

Let us first consider non-relativistic matter with equation of state $P = 0$. According to eq. (14) the evolution of the energy density is given by:

$$\dot{\rho} = -3H\rho \quad (24)$$

and thus $\rho \sim 1/a^3$. The result is evident, it is simply dilution of the number density of massive particle at rest.

The time dependence of the cosmological scale factor, is determined by eq. (13):

$$\dot{a}/a \sim \sqrt{\rho} \quad (25)$$

and thus in non-relativistic regime $a \sim t^{2/3}$ and $H = 2/3t$.

For relativistic matter the equation of state is $P = \rho/3$ and correspondingly

$$\dot{\rho} = -4H\rho. \quad (26)$$

Thus ρ drops as $\rho \sim 1/a^4$ and the scale factor rises as $a(t) \sim t^{1/2}$, which means that $H = 1/2t$.

The energy density of relativistic particles drops one power of a faster than that of non-relativistic ones due to dilution of their number density as volume, $1/a^3$, and red-shift

of the particle momentum. That's why relativistic matter dominated in the early universe, while at a later stage non-relativistic matter took over. Until last years of the XX century it was believed that the universe today is dominated by non-relativistic matter but then it was established that the dominant matter is the so called dark energy with equation of state close to the vacuum one.

In the vacuum(-like) regime the energy-momentum tensor is proportional to the metric tensor which is the only invariant tensor:

$$T_{\mu\nu} = \rho_{vac} g_{\mu\nu}. \quad (27)$$

Hence $P_{vac} = -\rho_{vac}$ and vacuum energy density remains constant in the course of the cosmological expansion: $\dot{\rho} = -3H(\rho + P) = 0$. The scale factor in this case rises exponentially $a \sim \exp(Ht)$.

According to our understanding, all visible universe originated from microscopically small volume with negligible amount of matter by exponential expansion with practically constant ρ , see below.

It is interesting to calculate the causality distance as a function of time, which is equal to the light path propagating from an initial moment $t_1 = 0$ to final moment t_2 . This distance can be found from the light geodesic equation: $dt^2 - a^2(t)dr^2 = 0$. It can be easily integrated to give:

$$l_\gamma = a(t) \int_0^t \frac{dt'}{a(t')}. \quad (28)$$

Thus for relativistic regime, $l_\gamma = 2t$, for non-relativistic one it is $l_\gamma = 3t$, and for exponential De Sitter (inflationary stage): $l_\gamma = H^{-1}[\exp(Ht) - 1]$.

As we see, cosmological equations do not have stationary solutions, at least for the examples taken. In the case of positive space curvature, i.e. $k > 0$, and normal matter with $\rho \sim 1/a^n$, $n=3,4$, the expansion will ultimately change into contraction. If however, $\rho > k/a^2$, e.g. if ρ is vacuum energy, the expansion may last forever for any k .

Note, that if the cosmological energy density is dominated by the normal matter the Hubble parameter drops down as $H \sim 1/t$, where t is the universe age. If the universe is dominated by vacuum(-like) energy, the Hubble parameter remains constant,

Problem 6. Find $\rho(a)$ and $a(t)$ for general linear equation of state, $P = w\rho$ with arbitrary w . Study the case of $w \leq -1$.

3.3 Cosmological parameters

Before proceeding further let us say a few words about the natural system of units which is used throughout all these lectures. We take speed of light, reduced Planck constant, and Boltzmann constant all equal to unity. $c = \hbar/2\pi = k = 1$. All dimensional quantities have dimension of length, or time, or (inverse) mass or energy – all the same. For example the Newtonian gravitational constant has dimension of inverse mass, $G_N \equiv 1/M_{Pl}^2$; $M_{Pl} = 1.221 \cdot 10^{19} \text{ GeV} = 2.176 \cdot 10^{-5} \text{ g}$; $m_p = 938 \text{ MeV} = 1.67 \cdot 10^{-24} \text{ g}$; $1 \text{ GeV}^{-1} = 1.97 \cdot 10^{-14} \text{ cm} = 0.66 \cdot 10^{-24} \text{ s}$; $1 \text{ eV} = 1.16 \cdot 10^4 \text{ K}$.

Homogeneous cosmology is described in terms of the Hubble parameter $H = \dot{a}/a$, which characterises the universe expansion rate, by the critical or closure energy density

$$\rho_c = 3H^2 M_{Pl}^2/8\pi \quad (29)$$

and by the dimensionless parameter $\Omega_j = \rho_j/\rho_c$, which measures the relative contribution of the energy density of species of type j into the total energy density of the universe. Clearly for spatially flat universe the total energy density is equal to the critical one: $\Omega_{tot} = 1$, if $k = 0$. It remains constant in the course of the universe expansion.

If $k \neq 0$, then from eq. (13) follows that Ω evolves with time as

$$\Omega(a) = \left[1 - \left(1 - \frac{1}{\Omega_0} \right) \frac{\rho_0 a_0^2}{\rho a^2} \right]^{-1}, \quad (30)$$

where the index sub-0 denotes the present day values of the corresponding quantities.

For normal matter $\rho a^2 \rightarrow 0$ if $a \rightarrow \infty$ and Ω runs away from unity: $\Omega(a) \rightarrow 0$ if $\Omega_0 < 1$ and $\Omega(a) \rightarrow \infty$ if $\Omega_0 > 1$. On the other hand, $\Omega(a) \rightarrow 1$ when $\rho a^2 \rightarrow \infty$, e.g. for vacuum energy at expansion for arbitrary initial Ω .

The present day value of the H characterises by dimensionless parameter h as

$$H = 100 h \text{ km/sec/Mps}, \quad (31)$$

where $h = 0.73 \pm 0.05$. The inverse quantity $H^{-1} = 9.8 \text{ Gyr}/h \approx 13.4 \text{ Gyr}$ is approximately equal to the universe age.

An exact expression for the universe age through the present day values of the Hubble parameter and relative energy densities of different forms of matter can be obtained by integration of the equation

$$\dot{a} = [8\pi \rho G_N a^2/3 - k]^{1/2} \quad (32)$$

After simple algebra one finds:

$$t_U = \frac{1}{H} \int_0^1 \frac{dx}{\sqrt{1 - \Omega_t + \frac{\Omega_m}{x} + \frac{\Omega_r}{x^2} + x^2 \Omega_v}}, \quad (33)$$

where Ω_t is the total Ω and $\Omega_{r,m,v}$ are respectively contributions from relativistic and non-relativistic matter and from vacuum energy. All the quantities in this equations are the present day ones; we skipped the sub-index 0 here.

Problem 7. Derive eq. (33). Find t_U for $\Omega_t = \Omega_m = 0, 0.3, 1$. Find t_U for $\Omega_t = 1, \Omega_m = 0.3$, and $\Omega_v = 0.7$.

The “measured” value of the universe age lies in the interval

$$t_U = 12 - 15 \text{ Gyr}, \quad (34)$$

found from the ages of old stellar clusters and nuclear chronology.

The present day value of the critical energy density is :

$$\rho_c = \frac{3H^2 m_{Pl}^2}{8\pi} = 1.88 \cdot 10^{-29} h^2 \frac{\text{g}}{\text{cm}^3} = 10.5 h^2 \frac{\text{keV}}{\text{cm}^3} \approx 10^{-47} h^2 \text{ GeV}^4 \quad (35)$$

It corresponds approximately to 10 protons per m^3 , but the dominant matter is not the baryonic one and in reality there are about 0.5 protons per cubic meter.

3.4 Matter inventory

The relative contributions of different forms of matter into the total energy density are obtained from different independent astronomical observations. Here we only present their numerical values. For discussion of their measurements in more detail a few extra lectures are necessary.

The total cosmological energy density is very close to the critical one: $\Omega_{tot} = 1 \pm 0.02$ as found from the position of the first peak of the angular spectrum of CMBR and the large scale structure (LSS) of the universe.

The usual baryonic matter makes quite small contribution: $\Omega_B = 0.044 \pm 0.004$ as found from the heights of the peaks in angular fluctuations of CMB, from production of light elements at BBN, and from the onset of structure formation with small $\delta T/T$.

Approximately five times more than baryons is brought by the so called dark matter. It is invisible matter with presumably normal gravitational interactions. As is found from galactic rotation curves, gravitational lensing, equilibrium of hot gas in rich galactic clusters, cluster evolution, and LSS: $\Omega_{DM} \approx 0.22 \pm 0.04$.

The rest, $\Omega_{DE} \approx 0.76$, is carried by some mysterious substance, which is uniformly distributed in the universe induces accelerated cosmological expansion. Its equation of state is close to the vacuum one, i.e. $w \approx -1$. The existence and the properties of dark energy was deduced from dimming of high- z supernovae, LSS, CMB spectrum, and the universe age.

I would like to stress that the different pieces of data and their interpretation are independent. It minimises the probability of a possible interpretation error. The numerical values obtained in different type measurements are pretty close to each other.

4 Brief cosmological history

Universe history can be separated into several epochs, some of them are described by established well known physics verified by experiment, some are based on hypothetical physics beyond the standard model, and some (little?) are absolutely dark.

1. Beginning, unknown. Quantum gravity, quantum space-time? Maybe time did not exist? It is so called pre-inflationary cosmology.
2. Inflation, i.e. epoch of exponential expansion of the universe. It is practically “experimental” fact..
3. End of inflation, particle production. At that period dark expanding “emptiness” filled by scalar (or some other) field, inflaton, exploded, creating light and other elementary particles.
4. Baryogenesis. At that time an excess of matter and antimatter in the universe (or vice versa) was created.
5. Thermally equilibrium universe, adiabatically cooled down. Presumably during this epoch several phase transitions took place leading to breaking of grand unified symmetry (GUT), electroweak (EW) symmetry, supersymmetry (SUSY), phase transi-

tion from free quark-gluon phase to confinement phase in quantum chromodynamics (QCD), *etc.* with possible formation of topological defects and non-topological solitons. At the phase transitions adiabaticity of expansion could be broken.

6. Decoupling of neutrinos from electromagnetic part of the cosmological plasma. It took place when the universe was about 1 second old at $T \sim 1$ MeV.
7. Big bang nucleosynthesis (BBN), which proceeded in the time interval from 1 s to ~ 200 s, and $T = 1 - 0.07$ MeV. At that time light elements, 2H , 3He , 4He , and 7Li were formed. Theory is in a good agreement with observations. A different mechanisms for creation of light elements is unknown. It makes BBN one of the cornerstones of the standard cosmological model (SCM).
8. Onset of structure formation which started when the dominating cosmological matter turned from relativistic into non-relativistic one. It took place at the red-shift $z_{eq} \approx 10^4$, $T \sim$ eV.
9. Hydrogen recombination, at $z \approx 10^3$ or $T \sim 0.2$ eV. At that time cosmic microwave radiation (CMB) decoupled from matter and after that it propagated practically freely in the universe. After decoupling of matter and radiation baryons begun to fall into already evolved seeds of structures created by dark matter (DM).
10. Formation of first stars and reionization of the universe.
11. Present time, $t_U = 12 - 15$ Gyr.

5 Hot equilibrium epoch

Usually a system comes to the state of thermal equilibrium after sufficiently long time. Paradoxically, in cosmology equilibrium is reached in the early universe when time is short but temperature is high and the reaction rates Γ exceed the cosmological expansion rate, $H = \dot{a}/a$:

$$\Gamma = \sigma n \sim \alpha^2 T > H \sim T^2/M_{Pl} \quad (36)$$

This condition is fulfilled at high temperatures but bounded by $T \leq \alpha^n M_{Pl}$, where α is the generic value of the coupling constant and $n = 1, 2$ for decays and reactions respectively. At lower T the equilibrium is broken due to Boltzmann suppression of the participating particles.

In equilibrium particle distribution functions are determined by two parameters only, by the temperature, T and, if the particles do not coincide with antiparticles, by their chemical potential, μ . The equilibrium distributions have the well known form:

$$f_{f,b}^{(eq)}(p) = \frac{1}{\exp[(E - \mu)/T] \pm 1}, \quad (37)$$

where $E = \sqrt{p^2 + m^2}$. Equilibrium with respect to the reaction $a_1 + a_2 + a_3 \dots \leftrightarrow b_1 + b_2 + \dots$ imposes the following condition on the chemical potential of the participating particles:

$$\sum_i \mu_{a_i} = \sum_j \mu_{b_j} \quad (38)$$

Chemical potentials are necessary to introduce in charge asymmetric case to describe inequality between number densities of particles and antiparticles: $n \neq \bar{n}$. It is assumed usually that in cosmological situation chemical potentials are very small, as follows from the observed baryon asymmetry of the universe. However, large lepton asymmetry is not excluded and so it may be interesting to consider non-negligible chemical potentials. One should keep in mind however, that chemical potential of bosons cannot be arbitrarily large. It is bounded by the mass of the boson, $\mu < m$, while for fermions there is no upper limit on μ .

What happens if charge asymmetry in bosonic sector, i.e. $(n - \bar{n})$ is so large that $\mu = m$ is not sufficient to realise that? In this case the equilibrium distribution function acquires an additional term:

$$f = [e^{(E-m)/T} - 1]^{-1} + C\delta^3(p), \quad (39)$$

i.e. Bose condensate forms. Notice that equilibrium distributions are always determined by two parameters: T and μ if $\mu < m$ or T and C if μ is fixed by the maximally allowed value, $\mu = m$.

Problem 8. Check that the distribution (39) is indeed an equilibrium solution of kinetic equation, i.e. $I^{coll} = 0$.

If the reactions of annihilation of particle and antiparticle

$$b + \bar{b} \leftrightarrow 2\gamma, 3\gamma \quad (40)$$

are in equilibrium, then from condition (38) follows that $\mu_\gamma = 0$ and that the chemical potentials of particles and antiparticles are equal by magnitude and have opposite signs:

$$\mu + \bar{\mu} = 0. \quad (41)$$

If the equilibrium with respect to annihilation into two photons is maintained, while the annihilation into larger number of photons is out of equilibrium (such reactions are slower due to an extra power of the fine structure constant α and smaller phase space), non-zero chemical potential of photons can be developed. The observed CMB photons have zero or very small chemical potential $|\mu/T| < 10^{-4}$.

The equilibrium number density of bosons with $\mu = 0$ is:

$$n_b \equiv \sum_s \int \frac{f_b(p)}{(2\pi)^3} d^3p = \begin{cases} \zeta(3)g_s T^3/\pi^2 \approx 0.12g_s T^3, & T > m; \\ (2\pi)^{-3/2}g_s(mT)^{3/2}e^{-m/T}, & T < m, \end{cases} \quad (42)$$

where g_s is the number of spin states.

The number density of photons is equal to:

$$n_\gamma = 0.2404T^3 = 412(T/2.728\text{K})^3 \text{ cm}^{-3}, \quad (43)$$

where 2.728 K is the present day temperature of the cosmic microwave background radiation (CMBR).

The equilibrium number density of non-degenerate (i.e. $\mu = 0$) fermions is:

$$n_f = \begin{cases} 3n_b/4 \approx 0.09 g_s T^3, & T > m; \\ (2\pi)^{-3/2} g_s (mT)^{3/2} e^{-m/T}, & T < m. \end{cases} \quad (44)$$

The equilibrium energy density is given by:

$$\rho = \sum \frac{1}{2\pi^2} \int \frac{dpp^2 E}{\exp[(E - \mu)/T] \pm 1}. \quad (45)$$

The total energy density of all species of relativistic matter with $\mu = 0$ is

$$\rho_{rel} = (\pi^2/30) g_* T^4, \quad (46)$$

where $g_* = \sum [g_b + (7/8)g_f]$ and $g_{b,f}$ is the number of spin states of bosons or fermions. *Problem 9.* Calculate g_* for $T \sim 3$ MeV. Answer: 10.75.

Sometimes the total energy density is described by expression (46) with the temperature depending number of species, $g_*(T)$ which includes contributions of all relativistic as well as non-relativistic species.

The energy density of CMB photons is

$$\rho_\gamma = \frac{\pi^2 T^4}{15} \approx 0.2615 \left(\frac{T}{2.728 \text{ K}} \right)^4 \frac{\text{eV}}{\text{cm}^3} \approx 4.662 \cdot 10^{-34} \left(\frac{T}{2.728 \text{ K}} \right)^4 \frac{\text{g}}{\text{cm}^3}. \quad (47)$$

The relative contribution of CMB into the cosmological energy density is small but non-negligible:

$$\Omega_{CMB} = 4.7 \cdot 10^{-5}. \quad (48)$$

Heavy particles, i.e. those with $m > T$, have exponentially small number and energy densities if they are in equilibrium:

$$\rho_{nr} = g_s m \left(\frac{mT}{2\pi} \right)^{3/2} e^{-m/T} \left(1 + \frac{27T}{8m} + \dots \right) \quad (49)$$

If the annihilation stopped, the density of massive particles could strongly exceed the equilibrium one. If the particles are unstable with a large life-time then at later stage their distribution would return to the equilibrium one.

Approach to equilibrium and deviations from it in homogeneous cosmology are described by the kinetic equation in FRW space-time:

$$\frac{df_i}{dt} = (\partial_t + \dot{p} \partial_{p_i}) f_i = (\partial_t - H p_i \partial_{p_i}) f_i = I_i^{coll}, \quad (50)$$

where $\dot{p} = -Hp$ and $I_i^{(coll)}$ is the collision integral, see below eq. (62).

Problem 10. Why in the distribution function p and t are taken as independent variables, while above we treated momentum as a function of time, $p = p(t)$?

In terms of dimensionless variables:

$$x = m_0 a \text{ and } y_j = p_j a \quad (51)$$

the l.h.s. of kinetic equation takes a very simple form:

$$Hx \frac{\partial f_i}{\partial x} = I_i^{coll}. \quad (52)$$

If the universe is dominated by relativistic matter, the temperature drops as $T \sim 1/a$ and the Hubble parameter is expressed through x as:

$$H = 5.44 \sqrt{\frac{g_*}{10.75}} \frac{m_0^2}{x^2 m_{Pl}}. \quad (53)$$

In thermal equilibrium: $g_* = 2$ for photons, $g_* = 7/2$ for e^\pm -pairs, and $g_* = 7/8$ for one family of left-handed neutrino. Since $H = 1/2t$ the relation between cosmological time and temperature of the primeval plasma has the form: $t/\text{sec} \approx (\text{MeV}/T)^2$.

For non-interacting particles, i.e. for $I^{coll} = 0$, equation

$$Hx \frac{\partial f}{\partial x} = 0 \quad (54)$$

is solved as

$$f = f(y, x_{in}), \quad (55)$$

where x_{in} is an initial value of the scale factor. Thus the distribution function maintains its initial form in terms of variables x and y . For massless particles with non-zero chemical potential: $f = f_{eq}(y, \xi)$, if they ever were in equilibrium. Here $T \sim 1/a$ and $\xi = \mu/T = \text{const}$. Initially equilibrium distribution of massless particles maintains its equilibrium form even after interactions are switched off, as is observed in CMB.

Let us check this important statement. The l.h.s. of the kinetic equation can be written as:

$$(\partial_t - Hp\partial_p) f_{eq} \left[\frac{E - \mu(t)}{T(t)} \right] = \left[-\frac{\dot{T}}{T} \frac{E - \mu}{T} - \frac{\dot{\mu}}{T} - \frac{Hp}{T} \right] \frac{df_{eq}}{dx}. \quad (56)$$

The factor in square brackets vanishes if $\dot{\mu} = \dot{T}/T = -H$, which is true in the expanding universe, and if $E(\dot{T}/T) = -Hp$ which can be and is true only for $E = p$ i.e. for $m = 0$.

If the particle mass is non-zero and if the interaction is switched off at $T \gg m$, the distribution looks as an equilibrium one but in terms of p/T but not E/T .

Problem 11. Find the distribution of massive particles decoupled at $T < m$. What if decoupling is non-instantaneous?

In the equilibrium state and for vanishing chemical potentials entropy, S , in comoving volume is conserved:

$$\frac{dS}{dt} \equiv \frac{d}{dt} \left(a^3 \frac{\rho + P}{T} \right) = 0 \quad (57)$$

In fact this equality is more general. It is true for any distribution function $f = f(E/T)$, satisfying the condition of the covariant energy conservation, $\dot{\rho} = -3H(\rho + P)$ with arbitrary $T(t)$. So we find:

$$\frac{d}{dt} \left(a^3 \frac{\rho + P}{T} \right) = a^3 \left[\frac{\rho + P}{T} \left(3H - \frac{\dot{T}}{T} - 3H \right) + \frac{\dot{P}}{T} \right], \quad (58)$$

where the pressure is:

$$P = \int \frac{d^3q}{(2\pi)^3} \frac{q^2}{3E} f \left(\frac{E}{T} \right). \quad (59)$$

From this expression we can find \dot{P} (remember that only T depends upon time) and integrate by parts to obtain:

$$\dot{P} = \frac{\dot{T}}{T} (\rho + P), \quad (60)$$

which leads to the conservation law (57).

Let us return now to the kinetic equation and specify the collision integral for an arbitrary process: $i + Y \leftrightarrow Z$:

$$I_i^{coll} = \frac{(2\pi)^4}{2E_i} \sum_{Z,Y} \int d\nu_Z d\nu_Y \delta^4(p_i + p_Y - p_Z) [|A(Z \rightarrow i + Y)|^2 \prod_Z f \prod_{i+Y} (1 \pm f) - \quad (61) \\ |A(i + Y \rightarrow Z)|^2 f_i \prod_Y f \prod_Z (1 \pm f)],$$

where Y and Z are arbitrary, generally multi-particle states, $\prod_Y f$ is the product of phase space densities of particles forming the state Y , and

$$d\nu_Y = \prod_Y \frac{d^3p}{(2\pi)^3 2E} \quad (62)$$

The signs '+' or '-' in $\prod(1 \pm f)$ are chosen for bosons and fermions respectively.

Equilibrium distributions by definition annihilate the collision integral:

$$I^{coll}[f^{(eq)}] = 0 \quad (63)$$

The standard Bose/Fermi distributions do that. It is easy to check that this is indeed true in T-invariant theory where the detailed balance condition holds,

$$|A_{if}(p)|^2 = |A_{fi}(p')|^2 \quad (64)$$

where p' is time reversed momentum, i.e. with the opposite sign of the space coordinate with respect to p .

So after an evident change of variables in the collision integral we can factor out $|A_{if}|^2$ and the integrand would be proportional to:

$$\prod f_{in} \prod (1 \pm f_{fin}) - \prod f_{fin} \prod (1 \pm f_{in}) = 0 \quad (65)$$

It is easy to check that the functions f_{eq} (37) annihilate the collision integrals due to conservation of energy

$$\sum E_{in} = \sum E_{fin} \quad (66)$$

and if chemical potentials satisfy:

$$\sum \mu_{in} = \sum \mu_{fin}. \quad (67)$$

This condition is enforced by reactions.

Since we know that CP-invariance is broken and (mostly) believe that CPT invariance holds, we must conclude that the invariance with respect to time reversal, T-transformation, is broken as well. It means, in particular, that the detailed balance condition is invalid, $|A_{if}|^2 \neq |A_{fi}|^2$. Now a natural question arises: would the usual equilibrium distributions survive in T-violating theory? Let us check what happens with the collision integral for $f = f_{eq}$. Due to eq. (65) the integrand is proportional to

$$I_{coll} \sim \Pi f_{in} (1 \pm f_{fin}) (|A_{if}|^2 - |A_{fi}|^2). \quad (68)$$

The last factor is non-vanishing if T -invariance is broken. However, due to S-matrix unitarity (or hermicity of the Hamiltonian) breaking of T -invariance is observable only if several processes participate and though each separate term is non-zero, the sum over all relevant processes vanishes [7].

It can be proven using S-matrix unitarity condition, $SS^\dagger = 1$. If as usually we introduce scattering matrix: $S = (I + iT)$, then it satisfies:

$$i(T_{if} - T_{fi}^\dagger) = - \sum_n T_{in} T_{nf}^\dagger = - \sum_n T_{in}^\dagger T_{nf} \quad (69)$$

Summation over n includes integration over phase space. Instead of detailed balance a new condition of cyclic balance [7]

$$\sum_k \int d\tau_k (|A_{ki}|^2 - |A_{ik}|^2) = 0 \quad (70)$$

ensures vanishing of I_{coll} on $f = f_{eq}$. Here $d\tau_k$ includes Bose/Fermi enhancement/suppression factors. For validity of this relation full unitarity is not necessary. Normalisation of probability $\sum_f w_{if} = 1$ plus CPT invariance are sufficient. If CPT is broken, then the additional condition $\sum_f w_{if} = 1$, $\sum_f w_{fi} = 1$ would save the standard equilibrium statistics. However, in the case that nothing above is true, equilibrium distributions would differ from the canonical ones. If the theory does not respects sacred principles of unitarity, hermicity, etc., the Pandora box of disasters would be open and equilibrium might deviate very much from the standard case or even not exist.

6 Freezing of species

As we discussed in the previous section, primordial plasma is typically in thermal equilibrium state in the early universe. When T dropped down to the so called decoupling

temperature, T_d , (sometimes it is called freezing temperature, T_f) the interaction with plasma effectively switched off and the particles started to behave as free, non-interacting ones. There are two types of decoupling:

1. Relativistic freezing, when $T_d > m$. This is realised e.g. for neutrinos, or some other hypothetical weakly interacting particles. Such particles by definition make hot dark matter (HDM) or warm dark matter (WDM).
2. Non-relativistic freezing, when $T_d < m$. Such particles make cold dark matter (CDM).

6.1 Relativistic freezing

Neutrinos decoupled at temperatures much larger than their mass, $T \gg m_\nu$. It can be seen from the decoupling condition that the weak interaction rate became smaller than the expansion rate $\sigma_W n < H$ at:

$$G_F^2 E^2 T^3 \sim T^2 / M_{Pl}, \quad (71)$$

where $G_F = 1.166 \times 10^{-5} \text{ GeV}^{-2}$ is the Fermi coupling constant.

Since the energy of relativistic particles is equal by an order of magnitude to the plasma temperature $E \sim T$ we obtain:

$$T_f \sim m_N (10^{10} m_N / M_{Pl})^{1/3} \sim \text{MeV}. \quad (72)$$

More accurate calculations are necessary to establish if T_f is larger or smaller than m_e , which is important for the calculations of the number density, n_ν , of relic neutrinos at the present time and for the cosmological bound on their mass, m_ν .

For more accurate calculations of the decoupling temperature from the electron-positron plasma we will use the kinetic equation in Boltzmann approximation with only direct reactions with electrons, i.e. νe elastic scattering and $\nu \bar{\nu}$ -annihilation taken into account:

$$Hx \frac{\partial f_\nu}{f_\nu \partial x} = - \frac{80 G_F^2 (g_L^2 + g_R^2) y}{3\pi^3 x^5}, \quad (73)$$

where we define $x = \text{MeV}/T$.

It is clear from this equation that the freezing temperature, T_f , depends upon the neutrino momentum $y = p/T$, and this can distort the spectrum of the decoupled neutrinos, as we see in what follows.

For the average value of the neutrino momentum, $y = 3$, the temperature of decoupling of neutrinos from e^\pm is

$$T_{\nu_e} = 1.87 \text{ MeV}, \text{ and } T_{\nu_\mu, \nu_\tau} = 3.12 \text{ MeV}. \quad (74)$$

Problem 12. Derive equation (73). Find decoupling temperature of annihilation, $\nu \bar{\nu} \leftrightarrow e^+ e^-$ which changes the number density of neutrinos.

Answer: $T_{\nu_e} \approx 3 \text{ MeV}$ and $T_{\nu_\mu, \nu_\tau} \approx 5 \text{ MeV}$.

To take into account all reactions experienced by neutrinos including elastic νe and all $\nu \nu$ scattering we need to make the substitution: $(g_L^2 + g_R^2) \rightarrow (1 + g_L^2 + g_R^2)$ in eq. (73) and find that neutrinos started to propagate freely in the universe when the temperature dropped below

$$T_{\nu_e} = 1.34 \text{ MeV} \text{ and } T_{\nu_\mu, \nu_\tau} = 1.5 \text{ MeV}. \quad (75)$$

6.2 Gershtein-Zeldovich(GZ) bound

Consideration of thermal equilibrium and entropy conservation permitted Gerstein and Zeldovich to derive famous cosmological bound on neutrino mass [8]. Sometimes this bound is called Cowsic-McLelland bound but this is not just because paper [9] has been published 6 years after Gerstein and Zeldovich and contained a couple of inaccurate statements which resulted in overestimation of the bound by factor 22/3.

We need to calculate the ratio n_ν/n_γ at the present time. The known from observations number density of photons in CMB, see eq. (43), allows to determine the cosmological number density of unobservable neutrinos. At neutrino decoupling the ratio is determined by thermal equilibrium:

$$n_\nu = n_{\bar{\nu}} = (3/8)n_\gamma \quad (76)$$

After decoupling n_ν was conserved in the comoving volume, i.e. $n_\nu a^3 = \text{const}$ but $n_\gamma a^3$ rises due to e^+e^- -annihilation into photons. At first sight the rise of n_γ is difficult to calculate but entropy conservation (57) makes the calculations trivial.

After neutrino decoupling the number density of neutrinos fall down as cosmological volume, $n_\nu \sim 1/a^3$. Before e^+e^- -annihilation but after neutrino decoupling, say at $T \sim 1\text{MeV}$, the entropy of photons and electron-positron pairs was:

$$S_{in} \sim (2 + 7/2)T_{in}^3 a_{in}^3. \quad (77)$$

After annihilation it became:

$$S_{fin} \sim 2T_{fin}^3 a_{fin}^3. \quad (78)$$

Since $n_\gamma \sim T^3$, and $S_{in} = S_{fin}$, the ratio of number densities of neutrino and photons dropped by the factor 4/11. If in the course of subsequent cosmological evolution the numbers of photons and neutrinos conserved in the comoving volume, the number density of neutrinos today must be

$$n_\nu + n_{\bar{\nu}} = \frac{3}{11} n_\gamma = 112/\text{cm}^3 \quad (79)$$

This result is obtained under assumption of vanishingly small chemical potentials of neutrinos, in other words for $n_\nu = n_{\bar{\nu}}$.

Energy density of neutrinos today should be smaller than the total energy density of matter, ρ_m . This leads to the following upper bound on the sum of masses of all neutrino species:

$$\sum m_{\nu_j} < 94 \text{ eV } \Omega h^2. \quad (80)$$

Since $h^2 \approx 0.5$, $\Omega_m \approx 0.25$, and masses of different neutrinos are nearly equal, as follows from the data on neutrino oscillations, we find $m_\nu < 5 \text{ eV}$.

This limit may be further strengthen, if one takes into account that cosmological structure formation would be inhibited at small scales if $\Omega_{HDM} > 0.3 \Omega_{CDM}$. Hence $m_\nu < 1.7 \text{ eV}$. Recent combined analysis of CMB and LSS leads to the bound:

$$m_\nu < 0.3 \text{ eV}. \quad (81)$$

For more detail and reviews see ref. [10] If the masses of neutrinos are close to this upper limit, their contribution into cosmological energy density at the present time would be non-negligible, $\Omega_\nu \sim 0.02$, comparable to that of baryons, $\Omega_b \approx 0.04$.

One may argue that though massless neutrinos are 100% left-handed, i.e. they have only one helicity state, massive neutrinos have both spin states and hence their cosmological number density should be twice larger than calculated. However, it is not so because right-handed states did not reach equilibrium and their contribution may be neglected.

Next question is how robust is the GZ bound. Is it possible to modify the standard picture to avoid or weaken it. The bound is based on the following assumptions:

1. Thermal equilibrium between ν , e^\pm , γ at $T \sim \text{MeV}$. If the universe never was at $T \geq \text{MeV}$, neutrinos might be under-abundant and the bound would be much weaker. However, successful description of light element production at BBN makes it difficult or impossible to eliminate equilibrium neutrinos at the MeV phase in the universe evolution.
2. Negligible lepton asymmetry. Non-zero lepton asymmetry would result in larger number/energy density of neutrinos plus antineutrinos and the bound would be stronger.
3. No extra production of CMB photons after neutrino decoupling. Strictly speaking this is not excluded but strongly constrained. If the extra photons were created before BBN terminated, they might distort abundances of light elements. Late time creation of extra photons, after BBN, would lead to distortion of the energy spectrum of CMB and there is only very small freedom, not sufficient to change n_ν/n_γ essentially.
4. Neutrino stability on the cosmological scale, $\tau_\nu > t_U$. If neutrino decays into another normal neutrino, e.g. $\mu_\mu \rightarrow \nu_e + X$, the total number of neutrinos does not change and the limit on the mass of the lightest neutrino remains undisturbed, but heavier neutrinos are allowed. If the decay goes into a new lighter fermions, e.g. sterile neutrino, the bound may be weakened for all neutrino species.
5. No late-time annihilation of $\nu + \bar{\nu}$ into a pair of (pseudo)goldstone bosons, e.g. majorons. For noticeable annihilation too strong coupling of neutrinos to majorons is necessary which is probably excluded by astrophysics.

6.3 Distortion of neutrino spectrum

As we mentioned above, see eqs. (54-56), massless particles keep their equilibrium spectrum even after the interaction is switched off. E.g. the spectrum of CMB photons is the equilibrium one with the precision better than 10^{-4} . However, it happened not to be true for neutrinos. The point is that neutrino decoupling is not instantaneous and for some time there coexist two components of plasma with different temperatures, weakly interacting with each other. Indeed, due to e^+e^- annihilation the photon temperature rises with respect to the neutrino temperature as, $T_\gamma/T_\nu \approx 1.4$.

Due to residual interaction of neutrinos with hotter electrons and positrons some energy is transferred to colder neutrino sector. More energetic neutrinos decoupled later. As a result the spectrum is distorted [11]:

$$\delta f_{\nu_e}/f_{eq} \approx 3 \times 10^{-4} \frac{E}{T} \left(\frac{11E}{4T} - 3 \right) \quad (82)$$

This analytical estimate was confirmed by precise numerical solution of the integro-differential kinetic equation; for discussion and the list of references see review [12].

This effect leads to an increase of the effective number of neutrino species:

$$\Delta N_\nu = 0.03 + 0.01. \quad (83)$$

The last 0.01 comes from plasma corrections [13], which diminish n_e and n_γ with respect to unperturbed quantities at the same temperature. An increase of the number of neutrino species has negligible effect on BBN, $\sim 10^{-4}$, but may be noticeable in CMB measurements by the recently launched Planck mission.

Note in conclusion of this section that though we mentioned above that neutrino temperature is approximately 1.4 times smaller than the temperature of photons, i.e. today it should be 1.95 K, would neutrino be massless, the distribution of neutrinos has the non-equilibrium form:

$$f_\nu \approx [\exp(p/T_\nu) + 1]^{-1}, \quad (84)$$

i.e the magnitude of neutrino momentum enters instead of energy and so the parameter T_ν does not have meaning of temperature. The correction (82) is neglected here.

6.4 Non-relativistic freezing

If particles have sufficiently strong interactions, they decouple from primordial plasma at temperatures much smaller than their mass, $T_f < m_h$. After that their number density stopped falling down according to the Boltzmann suppression law but remains constant in the comoving volume. The number density of heavy particles at decoupling is given by

$$n_h/n_\gamma \approx (m_h/T_f)^{3/2} e^{-m_h/T_f} \ll 1, \quad (85)$$

so such particles may have masses much larger than permitted by GZ bound and can make cosmological interesting cold dark matter. The frozen number density of such particles is determined by the cross-section of their annihilation and is given by a simple expression, see e.g. [14]:

$$\frac{n_h}{n_\gamma} \approx \frac{(m_h/T_f)}{\langle \sigma_{ann} v \rangle m_{Pl} m_h}, \quad (86)$$

where $m_h/T_f \approx \ln(\langle \sigma_{ann} v \rangle m_{Pl} m_h) \sim (10 - 50)$.

We derive this expression in what follows. One can solve kinetic equation governing evolution of the number density of heavy particles numerically but it is instructive to make analytic calculations. Moreover, the results are pretty accurate. Analytic calculations of frozen abundances are usually done under the following assumptions:

1. Boltzmann statistics. It is usually a good approximation for heavy particles at $T < m$.
2. It is assumed that heavy particles are in kinetic, but not chemical, equilibrium, i.e. their distribution function has the form:

$$f_h = e^{-E/T+\xi(t)}, \quad (87)$$

where ξ is the effective chemical potential normalised to temperature, $\xi = \mu/T$. Chemical equilibrium is enforced by annihilation which needs a partner whose number density is exponentially suppressed, while kinetic equilibrium demands encounter with abundant massless particles. That's why chemical equilibrium stopped to be maintained much earlier than the kinetic one.

3. The products of annihilation are in complete thermal equilibrium.
4. Charge asymmetry of heavy particles is negligible and thus the chemical potentials for particles and antiparticles are equal, $\xi = +\bar{\xi}$. Annihilation would be much more efficient in the case of non-zero charge asymmetry.

Kinetic equation under this assumption becomes an ordinary differential equation, which has been derived in 1965 by Zeldovich [15] and used for the calculation of the frozen number density of non-confined massive quarks in ref. [16]. In 1978 the equation was applied to the calculations of the frozen number densities of stable heavy leptons in ref. [17, 18] and after that it got the name Lee-Weinberg equation, though it would be more proper to call it Zeldovich equation.

The equation has the following simple form:

$$\dot{n}_h + 3Hn_h = \langle \sigma_{ann} v \rangle (n_h^{(eq)2} - n_h^2), \quad (88)$$

where n_h is the number density of heavy particles, $n_h^{(eq)}$ is its equilibrium value, and $\langle \sigma_{ann} v \rangle$ is thermally averaged annihilation cross-section multiplied by velocity of the annihilating particles:

$$\langle \sigma_{ann} v \rangle = \frac{(2\pi)^4}{(n_h^{eq})^2} \int \overline{dp}_h \overline{dp}_{\bar{h}} \int \overline{dp}_f \overline{dp}_{f'} \delta^4(P_{in} - P_{fin}) |A_{ann}|^2 e^{-(E_p+E_{p'})/T}, \quad (89)$$

where $\overline{dp} = d^3p/[2E(2\pi)^3]$.

The integration in eq. (89) can be taken down to one variable and we find [19]:

$$\langle \sigma_{ann} v \rangle = \frac{x}{8m_h^5 K_2^2(x)} \int_{4m_h^2}^{\infty} ds (s - 4m_h^2) \sigma_{ann}(s) \sqrt{s} K_1\left(\frac{x\sqrt{s}}{m_h}\right) \quad (90)$$

where $x = m_h/T$ and $s = (p + \bar{p})^2$. Usually $x \gg 1$ and $\sigma_{ann} v \rightarrow const$ near threshold, so thermally averaged $\langle \sigma v \rangle$ is reduced just to the threshold value of σv . The expression above can be useful if cross-section noticeably changes near threshold, e.g. in the case of resonance annihilation.

For derivation of eq. (88) we start with the general kinetic equation:

$$\partial_t f - H p \partial_p f = I_{el} + I_{ann}, \quad (91)$$

where we take into account only two-body processes with heavy particles, I_{el} and I_{ann} are respectively collision integrals for elastic scattering and annihilation. At $T < m_h$ the former is much larger than the latter because of exponential suppression of the number density of heavy particles, $f_h \sim \exp(-m_h/T)$. Since I_{el} is large, it enforces kinetic equilibrium, i.e. canonical distribution over energy:

$$f_h = \exp[-E/T + \xi(t)]. \quad (92)$$

With such a form of f_h we can integrate both sides of eq. (91) over $\bar{d}p$ and the large elastic collision integral disappears, but the trace of it remains in the distribution (92).

As the last step we express $\xi(t)$ through n_h :

$$\exp(\xi) = n_h/n_{eq} \quad (93)$$

and arrive to eq. (88).

The equation can be solved analytically, approximately but quite accurately. Usually at high temperatures, $T \geq m_h$, the annihilation rate is high:

$$\sigma_{ann} n_h / H \gg 1 \quad (94)$$

and thus the equilibrium with respect to annihilation is maintained, $n_h = n_{eq} + \delta n$, where δn is small. It is convenient to introduce dimensionless ratio of number density to entropy $r = n_h/S$, so the effects of expansion disappear from the equation:

$$\dot{n} + 3Hn = S\dot{r}. \quad (95)$$

Recall that S is conserved in comoving volume, eq. (57).

By assumption r weakly deviated from equilibrium, so we can write $r = r_{eq}(1 + \delta r)$, where $\delta r \ll 1$. In this limit the solution of eq. (88) can be found in stationary point approximation:

$$\delta r \approx -\frac{H x r'_{eq}}{2\sigma v S r_{eq}^2} \quad (96)$$

Since r_{eq} exponentially drops down, δr rises and at some moment δr would reach unity. After that we will use another approximation, neglecting r_{eq}^2 , integrate equation for r and obtain the final result (86).

Another way to solve equation (88) is to transform this Ricatti type equation to the second order Schroedinger type one and to integrate the latter in quasi-classical approximation.

Problem 13. Derive all above, in particular kinetic equation and solve it.

Problem 14. Find frozen number densities of protons and electrons in charge symmetric universe. *Answer:* $n_p/n_\gamma \approx 10^{-19}$, $n_e/n_\gamma \approx 10^{-16}$.

Problem 15. What number density would have anti-protons if $(n_p - n_{\bar{p}})/n_\gamma = 10^{-9}$.

Let us apply the obtained results for calculation of the frozen number density of lightest supersymmetric particle (LSP) which must be stable if R-parity is conserved and is a popular candidate for dark matter. The annihilation cross-section is estimated as:

$$\sigma v \sim \alpha^2/m_S^2 \quad (97)$$

Correspondingly the energy density of LSP would be:

$$\rho_{SUSY} = m_S n_S \approx \frac{n_\gamma m_S^2 \ln(\alpha^2 M_{Pl}/m_S)}{M_{Pl}} \quad (98)$$

For $m_S = 100$ GeV, which is a reasonable value for minimal supersymmetric model, we find:

$$\Omega_{SUSY} \approx 0.05. \quad (99)$$

It is very close to the observed 0.25 and makes LSP a natural candidate for DM.

Another interesting example is the frozen number density of magnetic monopoles, which may exist in spontaneously broken gauge theories containing $O(3)$ subgroup [20]. The cross-section of the monopole-antimonopole annihilation can be estimated as

$$\sigma_{ann} v \sim g^2/M_M^2, \quad (100)$$

where M_M is the monopole mass. Correspondingly the present day energy density of magnetic monopoles would be [21]:

$$\rho_M = \frac{n_\gamma M_M^2}{g^2 M_{Pl}}. \quad (101)$$

In fact slow diffusion of monopoles in cosmic plasma would slightly diminish the result but not much.

If $M_M \sim 10^{17}$ GeV, as predicted by grand unified theories, the monopoles would overclose the universe by about 24 orders of magnitude assuming that their initial abundance was close to the thermal equilibrium one. This problem played a driving role for the suggestion of inflationary cosmology.

7 Big bang nucleosynthesis

BBN is one of the pillars of the standard cosmological model. It describes creation of light elements, 2H , 3He , 4He , and 7Li , in the early universe when she was between 1 sec to 200 sec old and the temperature ran in the interval from 1 MeV down to 60-70 keV. The calculated abundances of light elements are in a good agreement with the observation. This proves that our understanding of the universe when it was so young, is basically correct.

The first stage of BBN is the freezing of the neutron-to-proton ratio, which determines the number density of neutrons for the second phase when formation of light elements took place. The n/p -freezing happened at $T \approx 1$ MeV and $t \approx 1$ s, while the light element formation occurred much later at $T \approx 65$ keV and $t \approx 200$ s.

The neutron-to-proton ratio is determined by the reactions:

$$n + e^+ \leftrightarrow p + \bar{\nu}_e \quad (102)$$

$$n + \nu_e \leftrightarrow p + e^-, \quad (103)$$

which frozen at $T \approx 0.7$ MeV (see below). After that r_{np} remained almost constant, slowly decreasing due to the neutron decay:

$$n \leftrightarrow p + e^- + \bar{\nu}_e, \quad (104)$$

whose life-time is $\tau_n = 886$ s. Since the formation of light elements started at $t \approx 200$ s, the decrease of r_{np} due to decay was essential. However, the decay is not important for $(n - p)$ freezing.

It is more convenient to consider the ratio of neutron number density normalised to total baryon number density, $r = n_n/(n_p + n_n)$, because latter is conserved in comoving volume at BBN epoch since baryonic number was conserved at low temperatures.

Kinetic equation which governs the neutron to baryon ratio can be obtained from the general kinetic equation (50) with collision integral given by eq. (62) in the limit of non-relativistic nucleons:

$$\dot{r} = \frac{(1 + 3g_A^2)G_F^2}{2\pi^3} [A - (A + B)r] \quad (105)$$

where $g_A = -1.267$ is the axial coupling constant of $(n - p)$ weak current and the coefficients A and B are

$$\begin{aligned} A &= \int_0^\infty dE_\nu K f_e(E_e) [1 - f_\nu(E_\nu)] |_{E_e=E_\nu+\Delta m} + \int_{m_e}^\infty dE_e K f_{\bar{\nu}}(E_\nu) [1 - f_{\bar{e}}(E_e)] |_{E_\nu=E_e+\Delta m} \\ &+ \int_{m_e}^{\Delta m} dE_e K f_{\bar{\nu}}(E_\nu) f_e(E_e) |_{E_\nu+E_e=\Delta m}, \\ B &= \int_0^\infty dE_\nu K f_\nu(E_\nu) [1 - f_e(E_e)] |_{E_e=E_\nu+\Delta m} + \int_{m_e}^\infty dE_e K f_{\bar{e}}(E_e) [1 - f_{\bar{\nu}}(E_\nu)] |_{E_\nu=E_e+\Delta m} \\ &+ \int_{m_e}^{\Delta m} dE_e K [1 - f_{\bar{\nu}}(E_\nu)] [1 - f_e(E_e)] |_{E_\nu+E_e=\Delta m}. \end{aligned} \quad (106)$$

where $K = E_\nu^2 E_e p_e$ and we included terms describing neutron decay, the last ones in expressions for A and B . In precision calculations relativistic corrections and all form-factors of $(n - p)$ -transformations are taken into account.

The expressions for A and B take very simple form if e^\pm and ν_e are in thermal equilibrium with equal temperatures. In this case $A = B \exp(-\Delta m/T)$. Moreover, in essential range of temperature m_e can be neglected and:

$$B = 48T^5 + 24(\Delta m)T^4 + 4(\Delta m)^2T^3. \quad (107)$$

In this approximation equation (105) can be easily solved numerically.

We can make an estimate of the freezing temperature using the following simple considerations. The reaction rate versus Hubble rate is:

$$\frac{\Gamma_{np}}{H} = \frac{(1 + 3g_A^2)G_F^2 B / 2\pi^3}{T^2 \sqrt{g_*} / 0.6 m_{Pl}} \quad (108)$$

The n/p freezing temperature can be approximately found from the condition $\Gamma_{np}/H = 1$, that is

$$T_{np} = 0.7 \left(\frac{g_*}{10.75} \right)^{1/6} \text{ MeV} \quad (109)$$

and correspondingly $(n/p)_f = \exp(-\Delta m/T_{np}) \approx 0.135$. Note that T_{np} depends upon the number of neutrino species through g_* , see discussion after eq. (53).

When T drops down to $T_{BBN} = 60 - 70$ keV, practically all neutrons quickly form ${}^4\text{He}$ (about 25% by mass), ${}^2\text{H}$ (3×10^{-5} by number), ${}^3\text{He}$ (similar to ${}^2\text{H}$), and ${}^7\text{Li}$ ($10^{-9} - 10^{-10}$). The calculated abundances span 9 orders of magnitude and well agree with the data.

Question: why T_{BBN} is much smaller than nuclear binding energy, $E_b \sim \text{MeV}$? The answer is below in this section.

Almost all frozen neutrons, except for those which decayed before the onset of light element formation at $T = T_{BBN} \approx 65$ keV, form ${}^4\text{He}$ because of its largest binding energy equal to 7 MeV/nucleon. Correspondingly the mass fraction of ${}^4\text{He}$ can be estimated as:

$$Y = 2(n/p)/[1 + (n/p)] \approx 24\%. \quad (110)$$

in a good agreement with the data.

It is interesting that a small variation of the Fermi coupling constant would strongly change the amount of the produced ${}^4\text{He}$ and correspondingly the star properties. The stars might either have a deficit of helium or of hydrogen.

Problem 16. Find the range of variation of G_F which is in agreement with $(25 \pm 1)\%$ of the mass fraction of ${}^4\text{He}$. Find the same for g_* or the number of neutrino families.

The light element formation proceeded through the chain of reactions: $p(n, \gamma)d$, $d(p\gamma){}^3\text{He}$, $d(d, n){}^3\text{He}$, $d(d, p)t$, $t(d, n){}^4\text{He}$, etc. All reaction go through formation of deuterium, because due to low baryon density two body processes dominate. An absence of a stable nuclei with $A = 5$ results in suppression of heavier nuclei production.

Let us turn now to the calculations of the temperature when the light elements were created. Naively one should expect this temperature to be close to the nucleus binding energy $T \sim E_b$. However, a large number of photons make it possible to destroy the produced nuclei on the tail of their energy distribution, despite the Boltzmann suppression. So one would expect that the temperature of nucleus formation is smaller than the binding energy by the logarithm of the ratio n_B/n_γ . We will derive now the Saha equation and see that this is indeed that case.

The equilibrium density of deuterium is determined by the equality of chemical potentials, $\mu_d = \mu_p + \mu_n$:

$$n_D = 3 \left(\frac{m_D T}{2\pi} \right)^{3/2} e^{[(-m_D + \mu_p + \mu_n)/T]}, \quad (111)$$

where the chemical potential can be expressed through the number density as

$$e^{\mu_p/T} = \frac{1}{2} n_p \left(\frac{2\pi}{m_p T} \right)^{3/2} e^{m_p/T} \quad (112)$$

Thus in equilibrium we can express the number density of deuterium through the number densities of protons and neutrons:

$$n_d = \frac{3}{4} n_n n_p e^{B_D/T} \left(\frac{2\pi m_d}{m_p m_n T} \right)^{3/2}, \quad (113)$$

where $B_D = m_p + m_n - m_D = 2.224$ MeV is the binding energy of deuterium and $n_p = \eta n_\gamma$ with $\eta \equiv \beta = 6 \cdot 10^{-10} \ll 1$. Notation η is used at consideration of BBN, while the same or similar quantity is denoted as β when baryogenesis is considered.

The number density of deuterium, n_d , becomes comparable to n_n at

$$T_{BBN} \approx \frac{B_D}{\ln(16/3\eta) + 1.5 \ln(m_p/4\pi T)} = \frac{0.064 \text{ MeV}}{1 - 0.029 \ln \eta_{10}}, \quad (114)$$

where $\eta_{10} = 10^{10} \eta$. At higher temperatures n_D (and densities of other nuclei) are much smaller than n_n and thus T_{BBN} can be considered as the temperature when the formation of light nuclei started. Because of the exponential dependence on T the light nuclei formation proceeded during quite short time interval.

Note, that due to the same effect, the hydrogen recombination took place at $T \sim 0.1$ eV which is much smaller than the hydrogen binding energy, $E_H = 14.6$ eV.

In Fig. 1, taken from ref. [22], the calculated abundances of light elements as functions of the baryon-to-photon ratio are presented.

Helium-4 slowly rises with η , because for larger η the moment of BBN becomes earlier, see eq. (114), and less neutrons decayed.

Deuterium abundance quickly drops down with rising η because the probability of processing of deuterium to heavier more tightly bound nuclei, ${}^4\text{He}$, is larger with larger baryonic number density. So less deuterium survives at larger η . High sensitivity of the primordial deuterium abundance to η allowed it to be the best way to measure the cosmological amount of baryons before more accurate CMB measurements became available. This is why deuterium was called ‘‘baryometer’’.

Lithium-7 is formed in two competing processes with different dependence on η . At low $\eta_{10} < 3$ the production predominantly goes through the reaction ${}^3\text{H}({}^4\text{He}, \gamma){}^7\text{Li}$. On the other hand, ${}^7\text{Li}$ can be destroyed by collisions with protons and with rising η destruction becomes more efficient and ${}^7\text{Li}$ drops down. At larger $\eta_{10} > 3$ the dominant process of creation is ${}^3\text{He}({}^4\text{He}, \gamma){}^7\text{Be}$. Destruction of ${}^7\text{Be}$ by protons is less efficient because ${}^7\text{Be}$ has larger binding energy than ${}^7\text{Li}$. So ${}^7\text{Be}$ production rises with rising η . At lower T , when formation of atoms became non-negligible, ${}^7\text{Be}$ could capture electron and decay into ${}^7\text{Li}$ and neutrino.

The BBN calculations are based on pretty well known low energy nuclear physics and theoretical uncertainties would not play a significant role in comparison of theory with observations, if we were able to observe these light elements at the epoch of their creation. However, we observe them now, while the results are obtained for very young universe, about 300 seconds old. So evolutionary effects should be taken into account. We will briefly describe the problems of comparison of theory with the data below. For more detailed discussion see reviews [22, 23].

Helium-4 is very tightly bound nuclei and so it is not destroyed in the course of cosmological evolution. Hence the observed abundance of ${}^4\text{He}$ should be larger than the

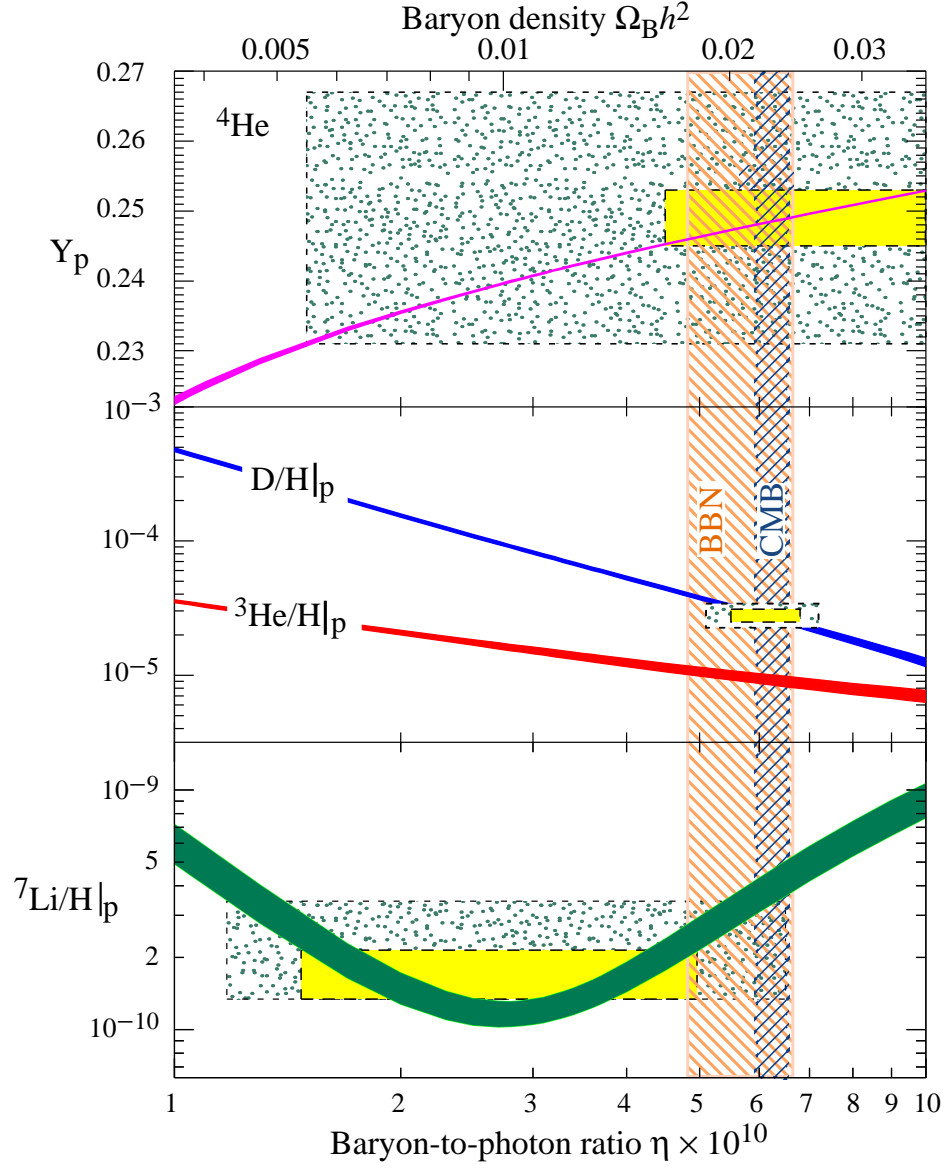


Figure 1: The abundances of He^4 , D, He^3 and Li^7 predicted by the standard model of BBN. The bands show the 95% CL range. Boxes indicate the observed light element abundances (smaller boxes: $\pm 2\sigma$ statistical errors; larger boxes: $\pm 2\sigma$ statistical and systematic errors). The narrow vertical band indicates the CMB measure of the cosmic baryon density, while the wider band indicates the BBN concordance range (both at 95% CL).

primordial one. With the existing observation means ${}^4\text{He}$ can be observed only at low red-shifts in chemically evolved regions with the abundance which may be quite different from the primordial one. To deduce the primordial abundance of ${}^4\text{He}$ one needs to extrapolate to zero metallicity. Namely the regions where ${}^4\text{He}$ is observed are contaminated by heavier elements. One can study the correlation of this elements with ${}^4\text{He}$ and extrapolate (linearly) the data to zero values of the metals (in astronomy all heavier than helium are called metals). Another source of uncertainty is not very well known fraction of ionised helium with respect to the total amount. This described by the so called ionisation correction, which is another source of uncertainty.

Deuterium could be destroyed in the course of evolution in poorly controlled manner. Fortunately, in contrast to helium, the deuterium line can be observed at large red-shifts, $z \sim 1$, i.e. at the earlier stages of the cosmological evolution. If the clouds, where deuterium is observed, are not contaminated by heavier elements, there is a good chance that deuterium is primordial. However, the deuterium line is shifted from the hydrogen one only by 80 km/sec and the peculiar motion of the cloud may induce an essential systematic error. The average value of deuterium abundance is in good agreement with the value of η determined from CMB, but the individual values are rather strongly dispersed from $1.6 \cdot 10^{-5}$ up to $3.5 \cdot 10^{-5}$. It would be interesting to understand the origin of such strong dispersion.

Primordial lithium visibly creates a potential problem for BBN, but ${}^7\text{Li}$ is difficult to observe (in first generation stars) and maybe it is premature to worry about the disagreement.

As a whole the data and theory are in a good agreement. Still there seems to be some “small clouds”. It would be very interesting if these clouds indicate new physics but most probably the resolution of the problems can be found in the traditional way when more accurate data are accumulated and better understood.

8 Role of neutrinos in BBN

BBN is sensitive to any form of energy which was present in the universe during light element formation. Indeed the universe cooling rate can be determined by equating two expressions for the cosmological energy density, namely the critical energy density and energy density of relativistic plasma with temperature T :

$$\rho = \frac{3m_{Pl}^2}{32\pi t^2} = \frac{\pi^2}{30} g_* T^4, \quad (115)$$

compare to eq. (53). The factor $g_* = 10.75 + 1.75\Delta N_\nu$ count the contributions from photons, e^\pm , neutrinos, and any other form of energy parametrized by ΔN_ν . Usually ΔN_ν is called the number of extra neutrinos, though the corresponding additional energy may have nothing to do with neutrinos. If $\Delta N_\nu = 1$, the additional energy is equal to the equilibrium energy density of one family of neutrinos plus antineutrinos with negligible mass (at BBN). In particular, if neutrinos are degenerate, i.e. their chemical potential μ

is non-zero, the additional energy density corresponds to

$$\Delta N_\nu = \frac{15}{7} \left[\left(\frac{\xi}{\pi} \right)^4 + 2 \left(\frac{\xi}{\pi} \right)^2 \right], \quad (116)$$

where $\xi = \mu/T$.

Problem 17. Derive eq. (116). Use eq. (45) and, if necessary, consult e.g. book [24], chapter V, sec. 58 for the calculation of the integral.

There are two effects induced by variation of g_* . First, larger ΔN_ν leads to earlier n/p -freezing and higher n/p -ratio, see eq. (109). Second, with larger g_* the BBN temperature (114) would be reached faster and more neutrons could survive against decay prior the onset of the light element formation. Both effects work in the same direction and $\Delta N_\nu = 1$ would lead to an increase of ${}^4\text{He}$ by 5%. Depending upon the data analysis the existing observational limit is

$$\Delta N_\nu < 0.3 - 0.5. \quad (117)$$

According to ref. [23], $N_\nu \approx 2.5$ seems to be the best fit. What is it, a problem or an “experimental” error?

Extra energy in electronic neutrinos have an additional and stronger effect on BBN because they can shift the equilibrium value of n/p -ratio:

$$(n/p)_{eq} = \exp \left(-\frac{\delta m}{T} - \xi_e \right) \quad (118)$$

Hence the bounds on lepton asymmetries depend upon the neutrino flavour and are much more restrictive for ν_e , than for $\nu_{\mu,\tau}$:

$$|\xi_{\mu,\tau}| < 2.5, \quad |\xi_e| < 0.1, \quad (119)$$

if a compensation between effects induced by $\xi_{\mu,\tau}$ and ξ_e is allowed. In absence of compensation the bounds are somewhat stronger. However, this results were obtained in the case of weak mixing between different neutrino flavours. In real case of large mixing angle solution to neutrino anomalies the bounds on all chemical potentials are equal and quite strong, see below, sec. 9.

The results presented above are valid for the equilibrium distributions of neutrinos. If ν_μ or ν_τ were out of equilibrium at $(n - p)$ -freezing their only effect would be a change of the energy density or, what is the same a contribution into ΔN_ν . As for ν_e , their deviations from equilibrium would directly change the freezing temperature, T_{np} . The effect depends upon the distortion of the neutrino energy spectrum. In particular, an excess of ν_e at high energies would shift T_{np} to lower values and would lead to a smaller n/p -ratio, which is opposite to the discussed above effect from an increase of the total energy density.

Sensitivity of BBN to additional energy at $T \sim 1$ MeV is used for deriving bounds on the number density of new (light) particles or new interactions of, say, neutrinos. As is discussed above, the parameter g_* , which describes the contribution of different species into cosmological energy density, cannot differ from the canonical value 10.75 more than

by 1. If neutrinos are massive there should be additional right-handed states in addition to the usual left-handed ones. If right-handed neutrinos are created by the canonical weak interactions, their production probability is proportional to m_ν^2 . The most favourable period for their production took place at $T \sim 100$ GeV through decays of W or Z bosons. The probability of production of ν_R is

$$\Gamma_R = \frac{\dot{n}_{\nu_R}}{n_\nu} = 10 \left(\frac{m_\nu}{T} \right)^2 \frac{\Gamma_W^\nu n_W + \gamma_Z^\nu n_Z}{T^3}. \quad (120)$$

One can check that ν_R have never been produced abundantly if m_ν respects the GZ-bound.

Right-handed neutrinos could be produced directly if there exist right-handed current induced e.g. by exchange of right-handed intermediate bosons, W_R . At some early stage of cosmological evolution ν_R might be abundantly created and this would endanger successful predictions of BBN. To avoid this problem ν_R should decouple before the QCD phase transition (p.t.). In this case their number and energy densities would be diluted by the entropy factor, as e.g. number density of the usual neutrinos were diluted by e^+e^- -annihilation discussed in sec. 6.2. The number of species above the QCD p.t. is 58.25, which includes three quark families (u,d,s) with three colours and 8 gluons with two spin states plus the usual 10.75 from e^\pm , γ , and ν . The energy dilution factor is $(10.75/58.25)^{4/3} \approx 0.105$. That is if three ν_R had equilibrium energy density before QCD p.t., their energy density at BBN would make 0.3 of the energy density of the one normal neutrino.

Since the ratio of the production rate of ν_R to the Hubble parameter is equal to:

$$\Gamma_R/H = (T/T_W)^3 (m_M/M_{W_R})^4, \quad (121)$$

ν_R would decouple before QCD p.t. if

$$\frac{m_{W_R}}{m_{W_L}} > \left(\frac{T_{QCD}}{200 \text{ MeV}} \right)^{4/3}. \quad (122)$$

Here T_W is the decoupling temperature of the normal neutrinos which is taken to be 3 MeV. Thus $m_{W_R} > (\text{a few}) \text{ TeV}$, which is an order of magnitude better than the direct experimental limit.

When/if the precision in BBN will reach the level $\Delta N_\nu < 0.1 - 0.2$, the bound on m_{W_R} would be much stronger than above, namely about 10^4 TeV , because we will need to move to the electroweak phase transition or higher.

Using similar arguments we can find a bound on the mixing between W_R and W_L :

$$W_1 = \cos \theta W_L + \sin \theta W_R \quad (123)$$

The probability of production of ν_R at $T \leq T_{QCD}$ in this case is given by:

$$r = \Gamma_R/H = \sin^2 \theta \left(\frac{T_{QCD}}{200 \text{ MeV}} \right)^3. \quad (124)$$

The condition $r < 0.3$ results in $\sin^2 \theta < 10^{-3}$, or smaller. The effect does not vanish for $m_{W_2} \rightarrow \infty$. Due to the presence of ν_{eR} the $(n-p)$ -transformation would be more efficient

and it would shift T_{np} to smaller values, thus compensating the increase of g_* , but the effect is small, $\sim \theta^2$. However, it may open window for very large $\theta \sim 1$.

If neutrino has a non-zero magnetic moment, ν_R would be produced in electromagnetic interactions:

$$e^\pm + \nu_L \rightarrow e^\pm + \nu_R \quad (125)$$

Demanding $\Delta N_\nu < 0.5$ we obtain:

$$\mu_\nu < 3 \times 10^{-10} \mu_B \quad (126)$$

If there existed primordial magnetic field, then ν_R could be produced by the spin precession in this field and the bound on the magnetic moment would be:

$$\mu_\nu < 10^{-6} \mu_B (B_{primord}/\text{Gauss})^{-1} \quad (127)$$

If there exist intergalactic magnetic fields with the strength $B_{int-gal} \sim 10^{-6}$ Gauss and if these fields were generated in the early universe and evolved adiabatically, then:

$$\mu_\nu < 10^{-19} \mu_B. \quad (128)$$

If there exists mirror matter which is similar or identical to ours, BBN demands that the temperature of the mirror stuff should be smaller than the temperature of the usual matter roughly by factor 2, see e.g. ref. [25].

More detail about the material of this and the next section can be found in review [12].

9 Neutrino oscillations in the early universe

Neutrino oscillations in medium are modified in the same way as light propagation. It can be described by refraction index or what is the same (up to energy factor) by effective potential. In cosmological plasma the effective potential contains two terms [26]:

$$V_{eff}^a = \pm C_1 \eta G_F T^3 + C_2^a \frac{G_F^2 T^4 E}{\alpha}, \quad (129)$$

where $C_j \sim 1$ and η is the plasma charge asymmetry:

$$\eta^{(e)} = 2\eta_{\nu_e} + \eta_{\nu_\mu} + \eta_{\nu_\tau} + \eta_e - \eta_n/2 \quad (\text{for } \nu_e) \quad (130)$$

$$\eta^{(\mu)} = 2\eta_{\nu_\mu} + \eta_{\nu_e} + \eta_{\nu_\tau} - \eta_n/2 \quad (\text{for } \nu_\mu) \quad (131)$$

Effective potential is proportional to the amplitude of forward elastic scattering (the same as the optical refraction index) and is usually calculated in the lowest order in the coupling constant. In our case it is first order in G_F .

In the local, 4-fermion, limit the Lagrangian describing elastic neutrino interaction has the form

$$\mathcal{L}_\nu = (G_F/\sqrt{2}) \sum_f g_f \bar{\psi}_\nu \gamma_\alpha (1 + \gamma_5) \psi_\nu \bar{\psi}_f \gamma_\alpha (1 + \gamma_5) \psi_f, \quad (132)$$

where f are fermions with which neutrino interacts (they include nucleons, charged leptons, and neutrinos) and g_f is the proper coupling constant. This interaction form can be read-off any textbook on weak interaction.

The first term comes from the averaging of the current $J_\alpha \bar{\psi}_l \gamma_\alpha (1 + \gamma_5) \psi_l$ over medium. Since the cosmological plasma is assumed to be homogeneous and isotropic, the average value of the space component of the current vanishes but $\langle J_0 \rangle \neq 0$, if the plasma is charge asymmetric. The result is proportional to the difference of the number densities of particles and antiparticles. This is the first term in eq. (129). The second term arises from non-locality of weak interactions due to W or Z boson exchange. At low energies the effect is proportional to $q^2/m_{W,Z}^2$, where q is the momentum transfer. The second term looks formally as being of the second order on the coupling constant, but it is imply because the intermediate boson mass is written as $m^{-2} \sim G_F/\alpha$, where α is the fine structure constant. In stellar interior V_{eff} is dominated by the first term, while in cosmology the second term takes over at $T \geq 10$ MeV.

There is another significant complication in cosmology because in contrast to stars (except for supernovae at dense stage), the of neutrino absorption and production and of coherence breaking due to scattering are significant. Because of that the standard wave function description is not adequate, the system is essentially open and the density matrix formalism should be applied [27]. The equation for density matrix has the following form:

$$\begin{aligned} \dot{\rho} = \left(\frac{\partial}{\partial t} - H p \frac{\partial}{\partial p} \right) \rho = & i [\mathcal{H}_m + V_{eff}, \rho] + \int d\tau (\bar{\nu}, l, \bar{l}) \left(f_l f_{\bar{l}} A A^+ - \frac{1}{2} \{ \rho, A \bar{\rho} A^+ \} \right) + \\ & \int d\tau (l, \nu', l') \left(f_{\nu'} B \rho' B^+ - \frac{1}{2} f_l \{ \rho, B B^+ \} \right). \end{aligned} \quad (133)$$

A and B are matrix amplitudes of scattering and annihilation respectively. The equation looks very complicated but numerical solution is possible. Moreover, in some cases, e.g. if there is MSW-resonance, even an accurate analytical solution can be found.

Let us discuss now how neutrino oscillation may influence BBN bounds on lepton asymmetry. It seems that for large lepton asymmetry, L , the oscillations are expected to be inhibited due to large first term in effective potential (129). This term is non-negligible even if $L \sim 10^{-9}$ and for $L \sim 1$ it might kill the oscillations at all. This is indeed the case for mixing between active and sterile neutrinos. However, for mixed active neutrinos off-diagonal terms in effective potential stimulate oscillations even in presence of large L [28].

If the oscillations are not suppressed, individual leptonic numbers would not be conserved and e.g. initial asymmetry in, say, muonic sector would be equally redistributed between all three flavours: L_e , L_μ , and L_τ . Thus the strict BBN bound on L_e , eq. (119), became valid for all three leptonic numbers. The efficiency of redistribution of initially large lepton asymmetry by the oscillations was calculated in ref. [29], for more related references and analytical calculations see review [12]. It has been shown that for weak mixing (LOW solution for the solar neutrino anomaly) the redistribution of leptonic numbers is insignificant, as is presented in fig. 2. However, for large mixing angle solution the equilibration of all leptonic numbers proceeded quite efficiently and they all became equal when the neutron freezing occurred, see fig. 3. As a result, the following quite restrictive

bound on asymmetry of all neutrino flavours can be obtained [29]:

$$|\xi_a| < 0.07. \quad (134)$$

If this is the case, neutrino degeneracy cannot have a noticeable cosmological impact, in particular on LSS and CMBR. The bound can be relaxed if neutrinos have new interaction with light Majorons. The potential induced by the majoron exchange inhibits early oscillations and different lepton numbers do not equalise [30].

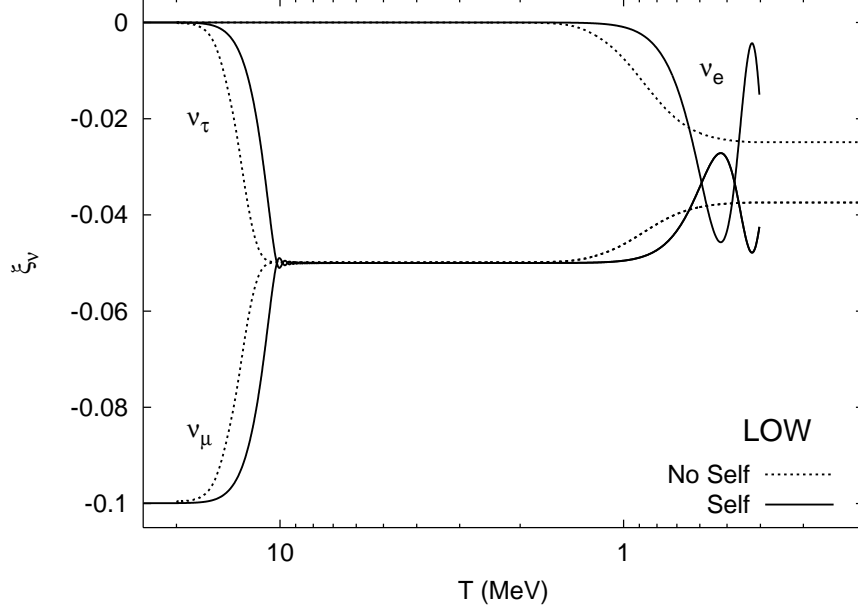


Figure 2: Evolution of lepton charge asymmetry for LOW solution to solar anomaly.

If there exists one or several sterile neutrino, ν_s , their mixing with the usual active ones would lead to several effects potentially observable in BBN:

1. An excitation of additional neutrino species leading to positive ΔN_ν .
2. Modification of spectrum of ν_e , because the oscillation probability depends upon neutrino energy.
3. Generation of large lepton asymmetry in ν_e .

The impact of these processes on BBN was discussed in a large numbers of works, see review [12]. In (non-realistic) case that active nus are not mixed and the mixing between ν_s and ν_a of a certain flavour is not resonance, the following bounds can be derived:

$$(\delta m_{\nu_e \nu_s}^2 / \text{eV}^2) \sin^4 2\theta^{\nu_e \nu_s} = 3.16 \cdot 10^{-5} \ln^2(1 - \Delta N_\nu), \quad (135)$$

$$(\delta m_{\nu_\mu \nu_s}^2 / \text{eV}^2) \sin^4 2\theta^{\nu_\mu \nu_s} = 1.74 \cdot 10^{-5} \ln^2(1 - \Delta N_\nu). \quad (136)$$

They are noticeably stronger than the bounds obtained from direct experiment.

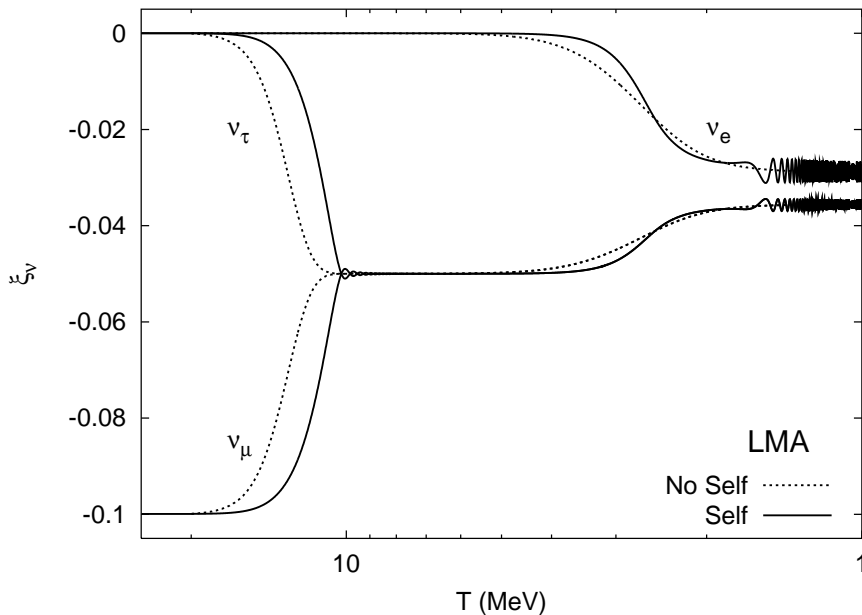


Figure 3: Evolution of lepton charge asymmetry for LMA solution to solar anomaly.

If the realistic mixing between $\nu_{e,\mu,\tau}$ are taken into account there are no simple analytical expressions for the BBN bounds, but they have been obtained numerically in ref. [31] both in non-resonance case where they are similar to those above, eq. (135,136), and for resonance case, where the bounds are much stronger.

10 Inflation

10.1 General features

Inflation is a period of exponential (or more generally accelerated) expansion of the very early universe, with approximately constant Hubble parameter:

$$a(t) \sim \exp[H_i t]. \quad (137)$$

It is the earliest time in the universe history about which we can say that it surely existed.

Inflation is easy to realise e.g. by a scalar field, inflaton, with the energy-momentum tensor:

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - (1/2)g_{\mu\nu} [\partial_\alpha \phi \partial^\alpha \phi - U(\phi)] \quad (138)$$

If field ϕ slowly evolves, i.e. if $(\partial_\mu \phi)^2 \ll U(\phi)$, then $T_{\mu\nu} \sim g_{\mu\nu}$ and the vacuum-like equation of state (27) would be realised.

The natural question, why do we need such a regime, has a very simple answer: inflation is the only *known* way to create the observed universe suitable for life. Inflation

naturally solves previously insolvable problems of Friedman cosmology:

1. Flatness. Without inflation the fine-tuning of $|\Omega - 1|$ should be 10^{-15} at BBN and 10^{-60} at the Planck era, see eq. (30). Otherwise the universe would re-collapse in much less than 10^{10} years, or expand too fast to make structures. Since $|\Omega - 1| \sim \exp(-2H_i t_i)$, the duration of inflation should be:

$$H_i t_i > 70. \quad (139)$$

More precisely the minimum duration of inflation may be somewhat smaller depending upon the heating temperature after inflation.

2. Causality. The distance that CMB photons can propagate in the Friedman universe, before they stopped interacting, was calculated in sec. 3.2, eq. 28. CMB photons stopped to interact with the surrounding medium after the hydrogen recombination at the redshift $z_{rec} \approx 10^3$, see eq. (114) and below. At this moment the universe age was $t \approx 10^{13}$ s. Thus it is the maximal size of the region d_c , which could be connected by interactions in Friedman cosmology. After recombination d_c rises due to the cosmological expansion by $z_{rec} = 10^3$ and today becomes $\sim 10^{16}$ s. The angular size of this path in the sky is:

$$\theta_{max} = \frac{10^{16}}{2\pi t_U} \approx 1^\circ, \quad (140)$$

and regions outside this size should not know anything about each other. On the other hand, CMB comes practically the same from all the sky. At inflationary stage the causally connected region is exponentially large, so inflation could make all the observed universe causally connected if $H_i t_i > 70$, the same as above.

3. Inflation explains the origin of the initial push which induced cosmological expansion by antigravity of an almost constant scalar field, since for such a state $P \approx -\rho$ and according the eq. (15) expansion speeds up and not slows down as naturally expected with attractive gravity.
4. Inflation makes the universe almost homogeneous and isotropic at the present-day Hubble scale. Indeed, any perturbation with $\delta/\rho/\rho \sim 1$ with wave length λ would transform into the perturbation with the same amplitude but with exponentially larger wave length:

$$\lambda \rightarrow \exp(H_i t_i) \lambda \quad (141)$$

In other words, perturbations at fixed scale exponentially smooth down.

5. Though inflation kills pre-existing perturbations, it creates its own small inhomogeneities, at the level 10^{-5} , but at astronomically large scales, which become seeds of large scale structure (LSS) formation. Inflation predicts adiabatic Gaussian density perturbations with almost flat Harrison-Zeldovich spectrum. In terms of dimensionless gravitational potential the spectrum does not contain any dimensional parameter and has approximately a simple form: $\delta\Psi \sim \delta k/k$. Deviations from flat spectrum agree with the data.

Problem 18. According to the existing models, the natural duration of inflation is much larger than $H_i t_i = 70$, hence one should expect $|\Omega - 1|$ negligibly small but in reality inflation predicts at horizon scale today: $|\Omega - 1| \sim 10^{-4}$. Why?

The idea of inflation was probably the most important breakthrough in cosmology of the XX century after the big bang one. Historically first paper where exponential expansion was invoked for solution of some problems of Friedman cosmology was that by Kazanas [32], who suggested that exponential regime could solve the problem of the observed homogeneity and isotropy of the universe. A few months later a famous paper by Guth “Inflationary universe: A possible solution to the horizon and flatness problems” was published [33]. This work has initiated a stream of papers which remains unabated to the present day. In both scenarios the vacuum-like energy, which might dominate at first order phase transition, was suggested as a driving force of exponential expansion. It was soon understood that such mechanism was not satisfactory because it would create an inhomogeneous universe consisting of many relatively small bubbles in exponentially expanding vacuum-like background. The first workable mechanism of inflation was suggested by Linde [34] and Abrecht and Steinhardt [35]. Probably the most beautiful inflationary mechanism, the so called chaotic inflation, was proposed by Linde [36]. For a review on these and discussed below issues, see refs. [37].

There was significant “pre-inflationary” literature directly related to the subject. The idea that the universe avoided singularity and underwent exponential period during which the mass of the cosmological matter rose by tens orders of magnitude was discussed by Gliner [38] and Gliner and Dymnikova [39]. De Sitter like (exponentially expanding) non-singular cosmology was considered by Gurovich and Starobinsky [40] and by Starobinsky [41]. In the last paper an important result was obtained that during “initial” exponentially expanding stage gravitational waves were produced which may be observable at the present time. If observed, it would be one of the strongest “experimental” indications to primordial inflation.

Another prediction of inflation is the spectrum of primordial density perturbations which is already verified by the data. The pioneering calculations of the spectrum have been done by Mukhanov and Chibisov [42] and confirmed by many subsequent studies [37].

It was shown in the paper by Sato [43] that exponential expansion induced by first order phase transition would never be terminated for certain under-critical values of the parameters. This happened to be a serious shortcoming of suggested later first inflationary scenarios. It was also noticed by Sato [44] that exponential expansion might permit astronomically interesting antimatter domains.

10.2 Models of inflation

There are now many dozens of different mechanisms of inflation, see reviews [37], and we do not have time to talk about them. We will discuss here only one which looks most economic and simple; it originated from the Linde’s suggestion of chaotic inflation [36].

Let us assume that there existed in the very early universe a small piece where some scalar field (we call it inflaton) is almost constant. In other words all derivatives, $\partial\phi$ are small. We will see that this piece of the universe would expand exponentially and spatial derivatives would smooth further down. The size of this smooth part of the universe should be larger than the inverse Hubble parameter. It is difficult to evaluate probability of such a state but it may be unnecessary. If it existed (with whatever small probability), our large universe would evolve out of this microscopically small piece.

If spatial derivatives can be neglected, field ϕ satisfies the following equation of motion:

$$\ddot{\phi} + 3H\dot{\phi} + U'(\phi) = 0 \quad (142)$$

The second term takes into account cosmological expansion and is similar to the liquid friction in Newtonian mechanics. If H is in some sense large (we will specify below the proper conditions), then the equation can be reduced to the first order one:

$$\dot{\phi} = -U'/3H. \quad (143)$$

Intuitively it is clear that in the case of large friction velocity is proportional to the force. This is the so called slow roll approximation which works pretty well in many inflationary scenarios.

If the cosmological energy density, ρ , is dominated by slow varying inflaton field ϕ , then the Hubble parameter is equal to

$$H^2 = \frac{8\pi U}{3m_{Pl}^2} \quad (144)$$

From this expression we can estimate the number of e-foldings while ϕ “lives” high in the potential $U(\phi)$ and the expansion is approximately exponential:

$$N = \int H dt = \frac{8\pi}{m_{Pl}^2} \int \frac{d\phi U(\phi)}{U'(\phi)}. \quad (145)$$

For the power law potential, $U(\phi) = g\phi^n$, we find:

$$N = \frac{4\pi}{n m_{Pl}^2} (\phi_{in}^2 - \phi_{fin}^2) \approx \frac{4\pi}{n m_{Pl}^2} \phi_{in}^2. \quad (146)$$

For successful inflation $N > 65-70$ is necessary. It implies $\phi_{in} > 2.5n^{1/2}m_{Pl}$. At first sight it looks disturbing if something is larger than m_{pl} . However this is not the case to worry about, because the observable quantity is the energy density of ϕ and it would remain much smaller than m_{Pl}^4 because of small mass of m_ϕ and the self-coupling constants, as we see in what follows.

For the validity of the slow roll approximation the following two conditions are to be fulfilled:

$$\ddot{\phi} \ll 3H\dot{\phi}. \quad (147)$$

and

$$\dot{\phi}^2 \ll 2U(\phi) \quad (148)$$

These conditions are satisfied if

$$\left| \frac{U''}{U} \right| \ll \frac{8\pi}{3m_{Pl}^2} \quad (149)$$

E.g. for massive free field $U = m^2\phi^2/2$ (harmonic potential) the slow roll approximation would be valid if:

$$\phi^2 > (4\pi/3) m_{Pl}^2 \quad (150)$$

With ϕ exactly at the lower limit the number of e-foldings is not enough but a slightly larger ϕ would do the job. The harmonic potential would not exceed the Planck value if

$$\phi^2 < m_{Pl}^4/m_\phi^2 \quad (151)$$

If we take ϕ equal to the upper bound, $\phi_{in} = m_{Pl}^2/m_\phi$, and $m_\phi \sim 10^{-6}m_{Pl}$, which is demanded by the condition of sufficiently small density perturbations, the number of e-folding would be huge: $N = 10^{13}$. The characteristic time when all this happened is tiny, $t_{inf} \sim 10^{-31}$ s.

Problem 19. Find N and t_{inf} for $U = \lambda\phi^4$, assuming $\lambda = 10^{-12}$ and initial $U(\phi) = m_{Pl}^4$.

10.3 Particle production by inflaton

During inflation the curvature of the inflaton potential should be smaller than the Hubble parameter, $U''(\phi) < H_I$, see the slow roll conditions above. In this regime inflaton monotonically but slowly moved down to the minimum of $U(\phi)$. Simultaneously decreases the Hubble parameter. At some moment the second derivative in eq. (142) became non-negligible and ϕ started to oscillate near minimum with frequency $\omega > H$. The expansion regime changed from the exponential to matter dominated one, if $U(\phi) = m_\phi^2\phi^2/2$, or to relativistic regime, if $U(\phi) = \lambda\phi^4/4$.

Problem 20. Solve eq. (142) with potential $U = m^2\phi^2$ and find cosmological scale factor $a(t)$ and $H(t)$ assuming that the cosmological energy is dominated by ϕ . Use expression (138) for the energy-momentum tensor of ϕ .

As is well known, a time varying field produces particles to which it is coupled and especially efficiently those whose mass is smaller than the frequency of the oscillations. It is exactly what happened with the inflaton. Empty, cold universe filled with oscillating ϕ exploded with creation of hot relativistic particles. It is almost as is described in the Bible: “Let there be light”. This moment is proper to call big bang, though it may be not generally accepted terminology.

Particle production by inflaton have been first calculated perturbatively in ref. [45]. It was shown that due to weakness of inflaton coupling to matter, perturbative particle creation is rather slow and the temperature of the universe heating after inflation is relatively small, much smaller than $\rho_{fin}^{1/4}$, where ρ_{fin} is the inflaton energy density to the end of inflation. Let us consider an example when the inflaton is coupled to fermions through the Yukawa coupling:

$$L_{int} = g\phi(t)\bar{\psi}\psi \quad (152)$$

where $\phi(t)$ is supposed to be classical field satisfying eq. (142):

$$\phi(t) = \frac{m_{Pl}}{\sqrt{3\pi}m_\phi} \frac{\sin m_\phi(t + t_0)}{t + t_0} \quad (153)$$

In the lowest order of perturbation theory the amplitude of production of pair of fermions with momenta k_1 and k_2 is

$$A(k_1, k_2) = g \int d^4x \sigma(t) \langle k_1, k_2 | \bar{\psi} \psi | 0 \rangle = \frac{(2\pi)^3 g}{\sqrt{4E_1 E_2}} \delta(\mathbf{k}_1 + \mathbf{k}_2) \bar{u}(k_1) v(k_2) \tilde{\phi}(E_1 + E_2), \quad (154)$$

where the standard decomposition of ψ and $\bar{\psi}$ in creation-annihilation operators is used, \bar{u} and v are the Dirac spinors, $E_i = |k_i|$ is the particle energy, and

$$\tilde{\phi}(\omega) = \int dt e^{i\omega t} \phi(t). \quad (155)$$

The probability of particle production per unit volume is

$$N_f \equiv \frac{W}{V} = \frac{1}{V} \int \frac{d^3 k_1 d^3 k_2}{(2\pi)^6} |A|^2 = \frac{g^2}{\pi^2} \int_{E>0} dE E^2 |\tilde{\phi}(2E)|^2. \quad (156)$$

The volume factor V as usually comes from the square of δ -function:

$$[\delta(\mathbf{k}_1 + \mathbf{k}_2)]^2 = \delta(\mathbf{k}_1 + \mathbf{k}_2) V / (2\pi)^3 \quad (157)$$

If $\omega \gg t^{-1}$ then the integration over time in the interval $\Delta t \gg \omega^{-1}$ in eq. (155) gives $\delta(\omega - 2E)$ and the square of it would be $\delta t \delta(\omega - 2E) / (2\pi)$. So we obtain for the rate of the fermion production per unit time and unit volume:

$$\dot{N}_f = \frac{N_f}{\Delta t} = \frac{m_1^2 \omega^2}{8\pi} = \frac{g^2 m_{Pl}^2}{24\pi^2 (t + t_0)^2} \quad (158)$$

This corresponds to the following decay rate of the field ϕ :

$$\Gamma_\phi = \frac{\dot{N}_f}{N_\phi} = \frac{g^2}{4\pi} m_\phi. \quad (159)$$

This is, as one can expect, the decay width of the ϕ -meson.

Because of the particle production $\phi(t)$ should decrease faster than it given by eq. (153). For $\Gamma_\phi \ll \omega$ this can be taken into account by the substitution $\phi(t) \rightarrow \phi(t) \exp(-\Gamma_\phi t)$.

The rate of thermalisation of the produced fermions is as a rule larger than the expansion rate. In this case the temperature of the plasma can be evaluated as follows. Assume that the particles are produced instantly at the moment when the Hubble parameter $H = 1/2t$ becomes equal to the decay rate Γ_ϕ and that $t \gg t_0$. The energy density of the produced fermions is

$$\rho_f = \frac{3\Gamma_\phi^2 m_{Pl}^2}{8\pi} = \frac{3g^4}{128\pi^3} m_{Pl}^2 m_\phi^2 \quad (160)$$

and correspondingly the temperature of the universe heating is

$$T_h = \left(\frac{30\rho_f}{\pi^2 g_*} \right)^{1/4} = \left(\frac{90}{128\pi^5 g_*} \right)^{1/4} g \sqrt{m_{Pl} m_\phi}. \quad (161)$$

For a more accurate evaluation of T_h let us take into account non instant character of the particle production and the decrease of the amplitude of the oscillations $\phi_0(t)$ not only because of the Universe expansion but also due to the particle production. With these factors taken into account the energy density of the produced fermions satisfies the equation

$$\dot{\rho}_f = \Gamma_\phi \rho_\phi - 4H\rho_f, \quad (162)$$

where Γ_ϕ is given by eq. (159) and $\rho_\phi = m_\phi^2 \phi_0^2(t)$ is the energy density of oscillating $\phi(t)$. We assume also that the total energy density is equal to the critical one:

$$\rho_\phi + \rho_f = \frac{3}{8\pi} m_{Pl}^2 H^2 = \frac{m_{Pl}^2}{6\pi(t+t_0)^2} \quad (163)$$

In the last equation it is assumed also that the non-relativistic expansion law is valid, that is $\rho_\phi > \rho_f$. Thus eq. (162) takes the form:

$$\dot{\rho}_f = \Gamma_\phi \rho_c - (4H + \Gamma_\phi) \rho_f. \quad (164)$$

Integrating it with the initial condition $\rho_f(0) = 0$ we obtain for the early MD-stage:

$$\rho_f = \frac{\Gamma^2 m_{Pl}^2 e^{-\Gamma t}}{6\pi[\Gamma(t+t_0)]^{8/3}} \int_0^{\Gamma t} dx e^x x^{2/3} \quad (165)$$

This equation is valid till the energy density of the produced fermions becomes non-negligible, e.g. till $\rho_f \approx \rho_\phi \approx \rho_c/2$. This is realised at $\Gamma t = 1.3$. At that moment $\rho_f \approx g^4 m_\phi^2 m_{Pl}^2 / 500\pi^3$ and the temperature is

$$T_h \approx \left(\frac{3}{50\pi^5 g_*} \right)^{1/4} g \sqrt{m_\phi m_{Pl}} \quad (166)$$

This is about twice smaller than result (161).

In many interesting cases perturbative approximation to particle production do not adequately describe physics of the process. Non-perturbative calculations have been pioneered in ref. [46] and further developed in ref. [47] with an emphasis on the parametric resonance enhancement. Due to that the inflaton decay can proceed much faster than expected from perturbation theory and the initial particle state might be far from thermal equilibrium. In particular, heavy particles with mass larger than would-be temperature could be produced. It was also noticed [48] that production of heavy particles by gravitational field at the final stage of inflation might be essential as well.

We conclude here that the heating temperature after inflation is model dependent and most probably is not large, $T_h < E_{GUT} \sim 10^{15}$ GeV, i.e. GUT era probably never existed. It solves the problem of overabundant magnetic monopoles. Initial hot universe might be far from thermal equilibrium and very heavy particles could be produced by the cosmological gravitational field [48].

10.4 Inflationary density perturbations

The increase of quantum fluctuations of scalar field in the exponentially expanding space-time gives rise to density inhomogeneities in the Universe. Physically this phenomenon is connected with different moments of the end of inflation in different space points (in an appropriate reference frame) due to small spatial fluctuations of the inflaton. Below we will follow the simple presentation of review [49]. For detailed rigorous treatment see book [50]

Exponential expansion transforms microscopically small wave lengths of quantum fluctuations into astronomically large ones and so produce natural candidates for initial density inhomogeneities which could be the seeds of the large scale structure formation. In a sense this task is even over-fulfilled because the inhomogeneities $\delta\rho$ prove to be too large for natural values of the parameters of the theory. In particular to get $\delta\rho/\rho \approx 10^{-4}$ on the galactic scale the self-coupling of the inflaton should be $\lambda < 10^{-12}$.

Since we expect that the density fluctuations have originated from quantum fluctuation during inflation, we need to say a few words about scalar field quantisation in De Sitter space-time. We assume that the field is massless because we are interested in $m_\phi \ll H$.

Field ϕ is quantised in the De Sitter space-time along the same lines as in flat space-time. The starting point is the expansion in creation-annihilation operators:

$$\phi(x, t) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} [a_k e^{-ikx} \phi(t) + h.c.], \quad (167)$$

where $\omega_k = \sqrt{k^2 + m^2}$ and $\phi_k(t)$ satisfies the Furrier transformed free equation of motion (in curved space-time):

$$\ddot{\phi}_k + 3H\dot{\phi}_k + \frac{k^2}{a^2}\phi_k = 0. \quad (168)$$

Operators a_k and $a_{k'}^+$ obey the standard commutation relation

$$[a_k, a_{k'}^+] = \delta^{(3)}(k - k'). \quad (169)$$

For massless scalar field the solution of eq. (168) is expressed through the Bessel functions:

$$\phi_k(t) = C_{1k}(k\tau)^{3/2} [J_{-3/2}(k\tau) + C_{2k}J_{3/2}(k\tau)], \quad (170)$$

where $\tau = \exp(-Ht)/H$ is the conformal time and the coefficients C_{ik} are determined by the matching ϕ_k to the corresponding expression in flat space-time, $\phi_k \rightarrow \exp(-i\omega_k t)$ for $H \rightarrow 0$ and $\tau \rightarrow H^{-1} - t$. Substituting asymptotic values of the Bessel functions for $k\tau \rightarrow \infty$ into eq. (170) we find

$$\begin{aligned} C_{1k} &= \sqrt{\frac{\pi}{2}} \frac{H}{k} e^{i\alpha_k} \\ C_{2k} &= -i, \end{aligned} \quad (171)$$

where α is a constant phase.

In the limit of large t (or $\tau \rightarrow 0$) one easily finds:

$$|\phi_k|^2 \approx \frac{H^2}{k^2} \left(1 + \frac{k^2}{H^2} e^{-2Ht} \right) \quad (172)$$

Let us extract from the inflaton field $\phi(x, t)$ classical homogeneous part $\phi_0(t)$:

$$\phi(x, t) = \phi_0(t) + \delta\phi(x, t), \quad (173)$$

where $\delta\phi(x, t)$ describes small quantum fluctuations. On inflationary stage it satisfies the equation:

$$\delta\ddot{\phi} + 3H\delta\dot{\phi} - e^{-2Ht}\partial_i^2\delta\phi - \frac{\partial^2 V(\phi_0)}{\partial\phi^2}\delta\phi = 0. \quad (174)$$

For large Ht the third term in the equation can be neglected and $\delta\phi$ satisfies the same equation as $\dot{\phi}_0$ does, see eq. (142). Equation (174) has two solutions. One of them decreases as $\exp(-3Ht)$ for $\partial^2 V/\partial\phi^2 \ll H^2$. The second solution varies relatively slowly. So at large t the first solution can be neglected and we can write

$$\delta\phi(x, t) = -\delta\tau(x)\dot{\phi}_0(t). \quad (175)$$

If ϕ is small this is equivalent to x -dependent retardation of the classical field motion to the equilibrium point:

$$\phi(x, t) = \phi_0(t - \delta\tau(x)). \quad (176)$$

Correspondingly inflation ends at different moments in different space points. This is the physical reason for generation of density perturbations. Since the energy density in the universe during inflation is dominated by the inflaton field ϕ , one can write $\rho(x, t) = \rho(t - \delta\tau(x))$ forgetting possible subtleties connected with the freedom in co-ordinate choice. Thus we get

$$\frac{\delta\rho}{\rho} = -\delta\tau\frac{\dot{\rho}}{\rho} = 4H\delta\tau(x). \quad (177)$$

At the last step the relation $\dot{\rho}/\rho = -4H$ has been used. It is valid on RD-stage which by assumption was formed in the heated universe when inflation was over. It is convenient to make the Furrier transform with a different integration measure:

$$\langle \delta\phi(x, t)^2 \rangle = \int \frac{d^3k}{k^3} |\Delta\phi(k, t)|^2. \quad (178)$$

The density fluctuations can be now expressed in terms of $\Delta\phi_k$ as

$$\frac{\delta\rho}{\rho} = 4H \frac{\Delta\phi_k}{\phi_0}. \quad (179)$$

The Fourier amplitude of the fluctuations is evaluated with the help of eqs. (170,172) as:

$$\Delta\phi(k, t) = \frac{H}{4\pi^{3/2}} \left(1 + \frac{k^2}{H^2} e^{-2Ht} \right)^{1/2}. \quad (180)$$

Recall that this expression is valid for massless free field. Since $\delta\phi(x, t)$ satisfies eq. (174) the approximation is valid till $k^2 \exp(-2Ht) \geq |\partial^2 U / \partial \phi^2|$. On the other hand, the proportionality condition (175) is valid in the opposite limit when one can neglect the decreasing part of the solution. So for the evaluation of the density fluctuations one has to substitute expression (180) in the boundary region:

$$t = t_b(k) = \frac{1}{2H} \ln \frac{k^2}{|\partial^2 U / \partial \phi^2|} \quad (181)$$

One should remember that $U(\phi)$ depends on t through $\phi = \phi_0(t)$ so eq. (181) implicitly determines t_b . Finely we find:

$$\frac{\Delta\rho}{\rho} \big|_h = \frac{H^2}{\pi^{3/2} \dot{\phi}_0[t_b(k)]} \left[1 + \frac{1}{H^2} \frac{\partial^2 V(\phi_0(t_b(k)))}{\partial \phi^2} \right]^{1/2}. \quad (182)$$

This expression for the fluctuation spectrum is valid till the fluctuation wave length reaches the horizon. This is indicated by sub-h in the left hand side.

Since $\phi(t)$ is a slowly varying function of time, the fluctuation spectrum weakly depends on k . This flat spectrum is known as the Harrison-Zel'dovich spectrum. For a satisfactory description of the universe structure the fluctuations should have the value $\Delta\rho/\rho|_h \approx 10^{-4}$.

As an example let us consider density perturbations in the model with the potential $U(\phi) = -\lambda\phi^4/4$. This potential is not bounded from below, so we can trust it only for sufficiently low ϕ . Homogeneous classical field $\phi_0(t)$ satisfies the equation:

$$\ddot{\phi} + 3H\dot{\phi} - \lambda\phi^3 = 0. \quad (183)$$

We assume that the term $\ddot{\phi}$ can be neglected, see subsection 10.2, and check the validity of this approximation on the explicit solution. The latter has the form:

$$\phi(t) = \left(\frac{3H}{2\lambda} \right)^{1/2} (t_f - t)^{-1/2}, \quad (184)$$

where t_f is approximately the moment when inflation ended. In a realistic model $U(\phi)$ is bounded from below, so $\phi(t)$ does not tend to infinity. Still solution (184) gives a satisfactory approximation up to t slightly less than t_f . After that moment the rise of ϕ should turn into oscillations around the equilibrium point.

The neglect of $\ddot{\phi}$ is justified if the condition

$$\frac{\ddot{\phi}}{3H\dot{\phi}} = \frac{1}{2H(t_f - t)} \ll 1 \quad (185)$$

is fulfilled. Substituting solution (184) into eq. (182) we arrive to

$$\frac{\Delta\rho}{r h o} \big|_h = \left(\frac{8}{3\pi^3} \right)^{1/2} \lambda^{1/2} [H(t_f - t_b)]^{3/2}, \quad (186)$$

where in accordance with eq. (181):

$$Ht_b = \ln \frac{k}{H} + \frac{1}{2} \ln \frac{2H(t_f - t_b)}{9}. \quad (187)$$

The second term in this expression is evidently small. Ht_b can be approximately expressed through the comoving wave vector k and the corresponding to it present-day physical scale l_0 as:

$$l_0 = \frac{2\pi a(t_0)}{ka(t_f)} e^{Ht_f} \approx \frac{2\pi T_h}{kT_0} e^{Ht_f}, \quad (188)$$

where $a(t)$ is the scale factor, T_0 is the present-day value of the CMB temperature, and T_h is the heating temperature at the end of inflation. It is taken into account that the physical momentum changes inversely proportionally to the scale factor. In particular during inflationary stage $p = k \exp(-Ht)$ and during Friedman stage p basically decreases as inverse temperature.

As a result

$$Ht_f = \ln \frac{kl_0 T_0}{2\pi T_R} \quad (189)$$

and finally we obtain

$$\frac{\Delta\rho}{\rho} \big|_h = \left(\frac{8}{3\pi^3} \right) \lambda^{1/2} \ln^{3/2} \frac{l_0}{b}, \quad (190)$$

where $b = (2\pi T_h/HT_0)$. For $T_h = 10^{15}$ GeV which is possibly rather high, but not unreasonable, we obtain $b \approx 15 m \approx 1.5 \cdot 10^{-15}$ years. Hence the density fluctuations on the galactic scale $l_0 = 10^6$ years are

$$\Delta\rho/\rho \approx 10^2 \lambda^{1/2} \quad (191)$$

So that λ should be tiny, $\lambda \approx 10^{-12}$, to give rise to a proper value of the fluctuations, $\Delta\rho/\rho \approx 10^{-4}$. This is a common shortcoming of inflationary models. Such a small value means in particular that the inflaton should be a gauge singlet or more complicated scenarios are necessary. Otherwise the interactions with gauge bosons would generate too big effective coupling, $\lambda_{eff}\phi^4$ with $\lambda_{eff} \approx \alpha^2 \approx 10^{-4}$.

The perturbation spectrum is not exactly flat but slightly deviates from the Harrison-Zeldovich one. This is a general feature of all inflationary models.

Generation of gravitational waves during inflation can be considered along the same lines. Moreover, the equation of motion of spin eigenstates of gravitons coincides with the equation of motion of massless scalar field. It may be instructive to note that gravitational waves are not created in eternal De Sitter background. One can check that the Bogolyubov coefficients which describe particle production are trivial. But when expansion is changed from exponential to, say, power law, the production of particles and waves becomes possible.

10.5 Inflationary conclusion

Inflation seems to be practically an experimental fact. It nicely explains the observed features of the universe. In particular, the observed spatial flatness today, $\Omega = 1 \pm (\sim 10^{-2})$ and very exact fine-tuning at, say, BBN, $\Omega = 1 \pm 10^{-15}$, naturally arises because of exponential expansion in the early universe. The observed spectrum of density fluctuations is predicted to be nearly flat, but not exactly flat. This is another successful prediction of inflation. The only “missing link” is not yet observed, namely, long gravitational waves, which may be accessible to LISA. However, one should keep in mind that though an observation of such waves will be a strong argument in favour of inflation but if they are not observed, inflation still will not be killed, because the amplitude of inflationary gravitational waves is model dependent and may be so small that they will escape observation by the near future antennas.

11 Baryogenesis

The observed part of the universe is 100% C(CP)-asymmetric. Up to now no astronomically significant objects consisting antimatter have been detected. There is only matter and no antimatter, except for a small number of antiprotons and positrons most probably of secondary origin. From the bounds of the flux of 100 MeV gamma rays one can conclude that the nearest galaxy, if dominated by antimatter, should be at least at ~ 10 Mpc [51]. However we cannot say much about galaxies outside of our super-cluster. Observed colliding galaxies at any distance or galaxies in the common cloud of intergalactic gas are of the same kind of matter (or antimatter?). In particular, the fraction of antimatter in two colliding galaxies in Bullet Cluster is bounded by $n_{\bar{B}}/n_B < 3 \times 10^{-6}$ [52]. In charge symmetric universe the nearest antimatter domain should be practically at the cosmological horizon, $l_B > \text{Gpc}$, because of very efficient annihilation at an early stage [53]. Still smaller clumps of antimatter are allowed in our neighbourhood.

However, one should bear in mind that these bounds are true if antimatter makes exactly the same type objects as the *observed* matter. For example, compact objects made of antimatter may escape observations and be quite abundant and almost at hand.

A natural question arises in this connection: is matter predominance accidental (a result of asymmetric initial conditions) or dynamical? Inflation proves that it is dynamical, originated from some physical processes with non-conserved baryon number. Sufficient inflation should last at least (~ 70 Hubble times) with practically constant Hubble parameter. This could happen only if the energy density was approximately constant, see eq. (13). However, if baryons are conserved, the energy density associated with baryonic number cannot be constant and inflation could last at most 4-5 Hubble times. Indeed if baryonic charge were conserved then it would remain constant in the comoving volume, i.e. $B \sim 1/a^3$. The energy density of bearers of this baryonic number cannot stay constant as well but should evolve as $\rho_B \sim 1/a^n$, where $n = 3$ for non-relativistic matter and $n = 4$ for relativistic matter.

Using the observed value of the cosmological density of the baryons and assuming its conservation, we find that at the early relativistic epoch (RD), e.g. at BBN the energy density related to this baryonic number should be $\rho_B \approx 10^{-7} \rho_{tot}$. During evolution at

RD-stage the ratio ρ_B/ρ_{tot} remained practically constant. Let us go backward in time till inflation. At inflation $\rho_{tot} \approx const$, but baryons excluded. The energy density of the latter, if their number is conserved, should evolve as $\rho_B \sim \exp(-4Ht)$. It means that 4-5 Hubble times back sub-dominant baryons were dominant and $\rho_{tot} \approx \rho_B$, which could not stay constant more than 4-5 Hubble times. This is surely insufficient for solution of cosmological problems discussed in the previous section.

It was suggested by Sakharov in 1967 [54] that the cosmological baryon asymmetry could be generated dynamically in the early universe if the following three conditions are fulfilled:

1. Non-conservation of baryons
2. Breaking of C and CP symmetries.
3. Deviation from thermal equilibrium.

Note in passing that none of them is obligatory – we present examples below. But first let us discuss normal baryogenesis and validity of these three conditions.

I. Non-conservation of baryons is justified theoretically. Grand unified theories, non-minimal SUSY, and electroweak (EW) theory predict that baryonic number is not conserved, $\Delta B \neq 0$. However, at the present time this prediction is not confirmed by direct experiment. Despite an extensive search, only upper bounds on proton life-time and period of neutron-antineutron oscillations are established. The only “experimental piece of data” in favour of baryon non-conservation is our universe: we exist, ergo baryons are not conserved. Half of century ago from the same experimental fact, our existence, an opposite conclusion of baryon conservation was deduced. Theory is an important input in understanding of what we observe.

II. C and CP violation are discovered and confirmed in direct experiments. At the first part of the XXth century the common belief was that physics was invariant with respect to separate action of all three transformations: mirror reflection, P, charge conjugation, C, and time reversal, T. The weakest link in this chain of discrete symmetries was P, found to be broken in 1956 [55].

It was immediately assumed that the world was symmetric with respect to the combined transformation from particles to mirror reflected antiparticles, CP. Both P and C are 100% broken in weak interactions but still some symmetry between particles and antiparticles was saved. This symmetry crashed down pretty soon, in 1964 [56]. After this discovery life in the universe became possible.

Why CP-breaking is necessary for generation of cosmological baryon asymmetry but C-breaking is not enough? A formal answer to this question is the following. Let us assume first that C is conserved and that the universe was initially in C eigenstate, i.e.

$$C|u\rangle = \eta|u\rangle \quad (192)$$

where $|u\rangle$ is the wave function of the universe and $|\eta| = 1$ is a constant. This means, in particular, that the universe had initially all zero charges because charge operator anticommute with C-transformation. May some non-zero charge, e.g. B , be generated dynamically? The answer is negative because due to C-invariance the Hamiltonian of the system commutes with C-operator, $[C, \mathcal{H}] = 0$. The time evolution of B is given by:

$$B(t) = \langle u|e^{-i\mathcal{H}t}J_0^B e^{i\mathcal{H}t}|u\rangle. \quad (193)$$

Let us insert into this equation the unity operator $I = C^{-1}C$:

$$B(t) = \langle u | I e^{-i\mathcal{H}t} I J_0^B I e^{i\mathcal{H}t} I | u \rangle = -B(t), \quad (194)$$

taken that $C J_0^B C^{-1} = -J_0^B$. Thus in C-conserving theory $B(t) = B_{in} = 0$.

The same arguments with CP instead of C prove that charge asymmetry cannot be generated, if CP is conserved and the universe is an eigenstate of CP :

$$CP|u\rangle = \eta|u\rangle. \quad (195)$$

In rotating universe charge asymmetry might be generated even if CP is conserved. Global rotation can be transformed into baryonic charge! However, it seems difficult to realise such an idea. New long-range interactions, possibly very unusual, are needed.

For B-generation in elementary (local) processes no assumption about the universe state is necessary: if CP is conserved, no asymmetry is generated through particle decays or reactions. We will discuss this below in concrete examples.

At the present time only CPT-symmetry survived destruction. It is the only one which has rigorous theoretical justification, CPT-theorem, based on solid ground: of Lorenz-invariance, canonical spin-statistics relation, and positive definite energy. Still models without CPT are considered, e.g. for explanation of some neutrino anomalies and for baryogenesis.

III. Thermal equilibrium is always broken for massive particles, but usually very little. To estimate the effect let us approximate the collision integral in kinetic equation (50) as $I_{coll} = \Gamma(f_{eq} - f)$, where Γ is the interaction rate. Let us assume that Γ is large so that the deviation from equilibrium is small, $\delta f / f_{eq} \ll 1$. Substituting into the l.h.s. of kinetic equation f_{eq} , as is done in eq. (56), we find:

$$\frac{\delta f}{f_{eq}} \approx \frac{Hm^2}{\Gamma TE} \sim \frac{Tm^2}{\Gamma Em_{Pl}}. \quad (196)$$

Since the Planck mass is very large the deviation from equilibrium might be significant only at large temperatures or tiny Γ . However, if the fundamental gravity scale is near TeV [57], equilibrium could be strongly broken even at the electroweak scale.

Another source of deviation from equilibrium in the cosmological plasma could be first order phase transition from, say, unbroken to broken symmetry phase in non-abelian gauge theories with spontaneous symmetry violation. There might be a rather long non-equilibrium period of coexisting two phases.

There are plenty scenarios of baryogenesis each of them one way or other performing a rather modest task explaining only one number, the observed asymmetry:

$$\beta = (n_B - n_{\bar{B}})/n_\gamma = 6 \cdot 10^{-10}, \quad (197)$$

found from the analysis of two independent measurements: of light element abundances created at BBN and of angular fluctuations of CMB.

It is a great challenge to astronomers to check if β is constant or it may vary at different space points, $\beta = \beta(x)$. What is characteristic scale l_B of variation of baryonic number density? May there be astronomically large domains of antimatter nearby or only very far

away? Answers to these questions depend, in particular, upon mechanism of CP violation realized in cosmology, which are described below. For more detail see lectures [58]. There are three possibilities for CP-breaking in cosmology:

1. *Explicit*, realized by complex coupling constants in Lagrangian, in particular, complex Yukawa couplings transformed by the vacuum expectation value of the Higgs field $\langle\phi\rangle \neq 0$ into non-vanishing phase in CKM-mixing matrix. However, in the minimal standard model (MSM) based on $SU(3) \times SU(2) \times U(1)$ CP-violation at $T \sim \text{TeV}$ is too weak, at least by 10 orders of magnitude, to allow for generation of the observed baryon asymmetry. Indeed, CP-violation in MSM is absent for two quark families because the phase in quark mass matrix can be rotated away. So at least three families are necessary. It could be an anthropic explanation why we need three generations.

If masses of different up or down quarks are equal, CP violation can be also rotated away because the unit matrix is invariant with respect to unitary transformations. If the mass matrix is diagonal in the same representation as flavour matrix, CP-violation can also be rotated away. Thus CP-breaking is proportional to the product of the mixing angles and to the mass differences of all down and all up quarks:

$$A_- \sim \sin \theta_{12} \sin \theta_{23} \sin \theta_{31} \sin \delta (m_t^2 - m_u^2)(m_t^2 - m_c^2)(m_c^2 - m_u^2) \\ (m_b^2 - m_s^2)(m_b^2 - m_d^2)(m_s^2 - m_d^2) / M^{12}. \quad (198)$$

At high $T \geq \text{TeV}$, where electroweak baryon nonconservation is operative, the characteristic mass $M \sim 100 \text{ GeV}$ and $A_- \sim 10^{-19}$. So for successful baryogenesis an extension of MSM is necessary.

2. *Spontaneous CP violation* [59], which could be realized by a complex scalar field Φ with CP-symmetric potential, with two separated minima at $\langle\Phi\rangle = \pm f$. The Lagrangian is supposed to be CP-invariant but these two vacuum states have the opposite signs of CP-violation. Such CP-breaking is locally indistinguishable from the explicit one but globally leads to charge symmetric universe with equal amount of matter and antimatter. As we mentioned at the beginning of this section, the antimatter domain should be very far at $l_B \geq \text{Gpc}$. Moreover, there is another problem with this mechanism, namely domain walls between matter and antimatter domains could destroy the observed homogeneity and isotropy of the universe [60]. To avoid the problem a mechanism of the wall destruction is necessary.

3 *Stochastic or dynamical*. If a complex scalar field χ was displaced from its equilibrium point in the potential, e.g. by quantum fluctuations at inflation, and did not relaxed down to equilibrium at baryogenesis, it would create CP-violation proportional to the amplitude of the field but without problems of spontaneous CP-violation. Later, after baryogenesis was over, χ would relax down to zero. So domain walls do not appear. Inhomogeneous $\beta(x)$ with domains of matter and antimatter can be created with such CP-violation. Their size depends upon the details of the scenario.

There is a long but probably incomplete list of different scenarios of baryogenesis (BG):

1. Heavy particle decays [54].
2. Electroweak BG [61]. Too weak in MSM but may work with TeV gravity.
3. Baryo-through-leptogenesis [62].
4. SUSY condensate BG [63].
5. Spontaneous BG [64].

6. BG by PBH evaporation [65].

7. Space separation of B and \bar{B} at astronomically large distances [66], which is probably not effective. However anti-baryons might be removed from our into higher dimensions [67] or predominantly accumulated inside quark nuggets [68].

7. BG due to CPT violation [71].

In all these scenarios new physics beyond minimal standard model is necessary. In what follows we will very briefly describe some of these scenarios. More details can be found in the reviews [72].

Heavy particle decay BG is naturally realized in grand unification theories, GUTs, where gauge bosons X with mass around $10^{16} - 10^{15}$ GeV are present. These bosons can decay e.g. into, qq and $\bar{q}l$ pairs where baryon number is evidently not conserved. Due to large mass of X the deviation from equilibrium could be significant, CP-violation might be sufficiently large (we know nothing about it) and the mechanism could be efficient enough to generate the observed asymmetry. The problem with GUTs is that the temperatures of the GUT scale might not be reachable after inflation. On the other hand, baryogenesis might proceed with under-abundant X -bosons created out of equilibrium.

Particles and antiparticles can have different decay rates into charge conjugated channels if C and CP are broken, while the total widths are equal due to CPT invariance. If only C is broken, but CP is not, then partial widths, summed over spins, are the same because CP-invariance implies:

$$\Gamma(X \rightarrow f, \sigma) = (\bar{X} \rightarrow \bar{f}, -\sigma). \quad (199)$$

If both C and CP are broken, partial widths may be different, but the effect takes place happen in higher orders of perturbation theory. In lowest order the amplitudes of charged conjugated processes must be equal, $A = \bar{A}^*$, because of hermicity of Lagrangian. The same would be also true for higher order contributions if they were real. An imaginary part is generated by re-scattering in the final state (with non-conservation of B or L), as can be seen from the S-matrix unitarity condition:

$$i(T_{if} - T_{if}^\dagger) = - \sum_n T_{in} T_{nf}^\dagger = - \sum_n T_{in}^\dagger T_{nf} \quad (200)$$

Let us consider an example of X -boson decays into the channels:

$$\begin{aligned} X &\rightarrow qq, \quad X \rightarrow q\bar{l}, \\ \bar{X} &\rightarrow \bar{q}\bar{q}, \quad \bar{X} \rightarrow \bar{q}l. \end{aligned} \quad (201)$$

and assume that the partial widths are different due to C and CP violation:

$$\begin{aligned} \Gamma_{X \rightarrow qq} &= (1 + \Delta_q)\Gamma_q, \quad \Gamma_{X \rightarrow q\bar{l}} = (1 - \Delta_l)\Gamma_l, \\ \Gamma_{\bar{X} \rightarrow \bar{q}\bar{q}} &= (1 - \Delta_q)\Gamma_q, \quad \Gamma_{\bar{X} \rightarrow \bar{q}l} = (1 + \Delta_l)\Gamma_l. \end{aligned} \quad (202)$$

Here $\Gamma \sim \alpha$ and $\Delta \sim \alpha$, where $\alpha \sim 1/50$ is the fine structure constant at GUT scale. The asymmetry is proportional to $\beta \sim (2/3)(2\Delta_q - \Delta_l)$. Its magnitude can be roughly estimated as

$$\beta \sim \frac{\delta f}{f} \frac{\Delta \Gamma}{\Gamma} \sim \frac{m}{m_{Pl}} \quad (203)$$

Small numerical coefficients omitted here would diminish the result. For example, the subsequent entropy dilution by about 1/100 is not included. For successful lepto/baryogenesis the mass of the decaying particle should be larger than 10^{10} GeV, or $m_{Pl} \ll 10^{19}$ GeV).

Problem 21. How the charge asymmetry generated in heavy particle decay vanishes in equilibrium? It is stated in the literature that the inverse decay does the job. However one can see that it is not so because using CPT, one finds:

$$\begin{aligned}\Gamma_{\bar{q}\bar{q}\rightarrow\bar{X}} &= (1 + \Delta_q)\Gamma_q, \quad \Gamma_{\bar{q}l\rightarrow\bar{X}} = (1 - \Delta_l)\Gamma_l, \\ \Gamma_{qq\rightarrow X} &= (1 - \Delta_q)\Gamma_q, \quad \Gamma_{q\bar{l}\rightarrow X} = (1 + \Delta_l)\Gamma_l.\end{aligned}\tag{204}$$

Thus direct and inverse decays produce the same sign of baryon asymmetry!

Electroweak baryogenesis is very attractive because all the necessary ingredients are present in the minimal standard model. CP is known to be broken, but, as we have seen above, very weakly. Baryonic number is non-conserved because of nonabelian chiral anomaly. At zero T baryon nonconservation is exponentially suppressed as $\exp(-2\pi/\alpha)$ [73], because of barrier penetration between different vacua. However, it is argued that at high T it is possible to go over the barrier, by formation of classical field configuration, sphalerons. We do not know how to calculate the probability of production of large coherent field configurations in elementary particle collision but lattice simulations show that their production might be efficient and sphalerons could have thermal equilibrium abundance. Unfortunately the deviation from equilibrium of massive particles at EW scale is tiny and the first order electroweak phase transitions seems to be excluded because of heavy Higgs boson. So electroweak baryogenesis, though very attractive is not efficient enough, but it may be operative with TeV gravity.

Baryo-through-leptogenesis is probably the most popular mechanism today. It started from creation of lepton asymmetry by L -nonconserving decays of heavy, $m \sim 10^{10}$ GeV, Majorana neutrino, analogously to GUT, and subsequent transformation of the lepton asymmetry into baryonic asymmetry by CP symmetric and B non-conserving and $(B - L)$ conserving electroweak processes. The mass matrix of three flavour light and heavy Majorana neutrinos has 6 independent phases, three in the sector of light neutrinos and 3 in heavy ones. They are unknown and allowed to be of order unity.

Primordial black hole evaporation does not demand B -nonconservation at particle physics level for generation of the baryon asymmetry. Of course, thermal evaporation cannot create any charge asymmetry. However the spectrum is not exactly black but is modified due to propagation of the produced particles in gravitational field of BH. Moreover, an interaction among the produced particles is essential. Let us assume that a meson A is created at the horizon and decays as:

$$A \rightarrow H + \bar{L} \text{ and } A \rightarrow \bar{H} + L,\tag{205}$$

where H and L are a heavy and light baryons, e.g. t and u quarks, respectively. Due to CP-violation the branching ratios of these decays may be different. Back-capture of H by gravitational field of the black hole is larger than that of L . Thus some net baryon asymmetry in the external world could be created. If the cosmological energy density of black holes at production was small, $\rho_{BH}/\rho_{tot} = \epsilon \ll 1$, then at red-shift $z = 1/\epsilon$,

with respect to the production moment, the non-relativistic BHs would dominate. Their evaporation could provide the necessary baryon asymmetry and reheat the universe.

For a numerical estimate let us present some simple formulae, omitting numerical factors of order 1 – 10. The black hole temperature is essentially given by the only available parameter with dimension of length, i.e. by its gravitational radius:

$$T_{BH} \sim 1/r_g \sim m_{Pl}^2/M_{BH}. \quad (206)$$

The luminosity of the body with temperature T_{BH} and radius r_g is:

$$L_{BH} \sim T^4 r_g^2 \sim m_{Pl}^4/M_{BH}^2. \quad (207)$$

Correspondingly the BH life-time is equal to:

$$\tau_{BH} \sim M_{BH}^3/m_{Pl}^4. \quad (208)$$

For example, if $M_{BH} = 10^{15}\text{g}$, its life-time is equal to the universe age, $\tau_{BH} \approx t_U \sim 10^{10}$ years. For our case much lighter BHs are needed. Let us assume that primordial BHs were formed at $T_{BH} = 10^{14}\text{ GeV}$ and their mass was equal to the mass inside the cosmological horizon at that moment, $M_{BH} = m_{Pl}^2 t \approx 10^4\text{g}$. The life-time of such BHs would be $\tau_{BH} \sim 10^{-16}$ sec, which corresponds to cosmological temperature $T \sim 10^5\text{ GeV}$ and red-shift from the moment when horizon mass was equal to M_{BH} , was about 10^9 . In other words, if the mass fraction of BHs at production was 10^{-9} , then at the moment of their evaporation they would dominate the cosmological energy density and could create observed baryon asymmetry even if the fraction of the baryon number density was small in comparison with the total number density of the evaporated particles.

Spontaneous baryogenesis may operate in thermal equilibrium. Explicit CP-violation is not obligatory. It is assumed that a global $U(1)$ -symmetry associated with baryonic number is spontaneously broken. The Higgs-like scalar boson acquires non-zero vacuum expectation value and its phase becomes massless Goldstone boson, $\phi = \eta \exp(i\theta)$. In the broken phase the Lagrangian can be written as:

$$\mathcal{L} = \eta^2(\partial\theta)^2 + \partial_\mu\theta j_\mu^B - V(\theta) + i\bar{Q}\gamma_\mu\partial_\mu Q i\bar{L}\gamma_\mu\partial_\mu L + (g\eta\bar{Q}L + h.c.). \quad (209)$$

In the case of homogeneous $\theta(t)$ the second term looks like chemical potential, $\dot{\theta}n_N$. However, in reality it is not true, because chemical potential is introduced into Hamiltonian but for derivative coupling $\mathcal{L} \neq \mathcal{H}$.

If the potential $V(\theta) = 0$, i.e. in purely Goldstone case, we can integrate the equation of motion:

$$2\eta^2\partial^2\theta = -\partial_\mu j_\mu^B \quad (210)$$

and obtain:

$$\Delta n_B = -\eta^2\Delta\dot{\theta}, \quad (211)$$

i.e. non-zero baryon asymmetry in thermal equilibrium and without explicit CP-violation. The latter is created by initial $\dot{\theta} \neq 0$.

In realistic situation $\dot{\theta}$ is small (because inflation kills all motion) and the pseudogoldstone case, i.e. non-zero $V(\theta)$ could be more efficient. Now the equation of motion for θ takes the form

$$\eta^2 \ddot{\theta} + 3H\dot{\theta} + V'(\theta) = \partial_\mu j_\mu^B, \quad (212)$$

where $V(\theta) \approx m^2 \eta^2 [-1 + (\theta - \pi)^2]$ and $j_\mu^B = \bar{\psi} \gamma_\mu \psi$. Initially θ is uniform in $[0, 2\pi]$ and after inflation it started to oscillate around minimum.

The second necessary equation is that for the quantum baryonic Dirac field:

$$(i\partial + m) \psi = -g\eta l + (\partial_\mu \theta) \gamma_\mu \psi \quad (213)$$

The solution to this equation can be found in one-loop approximation for $\psi(\theta)$ in external classical field θ . Then this solution, $\psi^\dagger \psi = F(\theta)$, should be substituted into eq. (212). In this way a closed equation for $\theta(t)$ can be obtained. The solution oscillates with alternating baryonic number giving the net result for the baryon number density

$$n_B \sim \eta^2 \Gamma_{\Delta B} (\Delta\theta)^3. \quad (214)$$

The SUSY baryonic condensate scenario will be discussed in more detail here because with simple modification it allows for creation of astronomically significant antimatter [69].

The basic features of this scenario are the following. SUSY predicts existence of scalars with non-zero baryonic number. Such bosons may condense along flat directions of the potential:

$$U_\lambda(\chi) = \lambda |\chi|^4 (1 - \cos 4\theta), \quad (215)$$

where $\chi = |\chi| \exp(i\theta)$. In SUSY models with high energy scale the baryonic number is naturally non-conserved. It is reflected by the non-sphericity of potential (215). Due to infrared instability of massless ($m \ll H$) fields in de Sitter space-time, χ can travel away from zero along the flat directions, $\theta = 0, \pi/2, \pi, 3\pi/2$. We can also add a mass term to the potential:

$$U_m(\chi) = m^2 |\chi|^2 [1 - \cos(2\theta + 2\alpha)], \quad (216)$$

where $m = |m|e^\alpha$. If $\alpha \neq 0$, then C and CP are explicitly broken, though it is not necessary for baryogenesis,

“Initially” (as a result of inflation) χ was pushed away from origin and when inflation was over it started to evolve down to equilibrium point, $\chi = 0$, according to the equation of the Newtonian mechanics:

$$\ddot{\chi} + 3H\dot{\chi} + U'(\chi) = 0. \quad (217)$$

The baryonic number of χ :

$$B_\chi = \dot{\theta} |\chi|^2 \quad (218)$$

is analogous to mechanical angular momentum. Using this mechanical analogy, and having the picture of potential $U(\chi)$ is easy to visualise the solution of the equation of motion without explicitly solving it.

The baryonic number of χ is accumulated in its “rotational” motion, induced by quantum fluctuations in orthogonal to valley direction. When χ decays its baryonic charge is transferred to that of quarks through B-conserving processes. If the mass term is absent or symmetric with respect to the phase rotation of χ this scenario leads to globally charge symmetric universe. The domain size l_B is determined by the size of the region with a definite sign of $\dot{\theta}$. Usually l_B would be too small if no special efforts are done.

If $m \neq 0$, the angular momentum, B, is generated by a different direction of the mass valley at low χ . If CP-odd phase α is small but non-vanishing, both baryonic and antibaryonic regions are possible with dominance of one of them. In this case matter and antimatter domain may exist but globally $B \neq 0$.

Now let us modify the model by adding general renormalizable coupling of χ to inflaton field Φ :

$$\lambda_\Phi |\chi|^2 (\Phi - \Phi_1)^2, \quad (219)$$

where Φ_1 is the value of Φ which it passed during inflation, not too long before its end. It is a free adjustable parameter.

Because of this coupling the gates to the valley would be open only for a short time when Φ was close to Φ_1 . So the probability for χ to reach a large value would be small. As a result we will have the following picture of the universe. The bulk of space would have normal homogeneous baryon asymmetry, $\beta = 6 \cdot 10^{-10}$, with small bubbles having large $\beta \sim 1$. In the simplest version of the scenario the high B regions should be almost symmetric with respect to baryons and antibaryons.

The mass spectrum of such baryon rich bubbles is practically model independent (it is determined by inflation) and has simple log-normal form:

$$\frac{dN}{dM} = C_0 \exp [-C_1 \ln^2 (M/M_0)] \quad (220)$$

Such object could make primordial black holes, quasars, disperse clouds of antimatter, and unusual stars and anti-stars, all not too far from us in the Galaxy. Phenomenological implications of this mechanism of antimatter creation and observational bounds are discussed in ref. [70]. If such mechanism is realized in nature the attempts for search of cosmic antimatter have non-zero chances to be successful.

It is worth noting that primordial nucleosynthesis in high B domains proceeded with large ratio n_B/n_γ . So the outcome of light and heavier element abundances could be much different from the predictions of the standard BBN. Thus the regions in the sky with abnormal chemistry would be first candidates to search for cosmic antimatter through 100 MeV photons or through the positron annihilation line.

References

- [1] C. Brans, R.H. Dicke, Phys. Rev. 124 (1961) 925.
- [2] S. Deser, B. Laurent, Ann. Phys. 50 (1968) 76;
L.P. Grishchuk, UFN, 160 (1990) 147;

- S.V. Babak, L.P. Grishchuk, Phys. Rev. D61 (1999) 024038; Int. J. Mod. Phys. D12 (2003) 1905.
- [3] L.D. Landau, E.M. Lifshitz, v.2, *Field theory*.
 - [4] A. Friedman, Z. Phys. 10 (1922) 377.
 - [5] E. Hubble, (1929) Proc. Nat. Acad. Sci. 15 (1929) 168.
 - [6] H.P. Robertson, Astrophys. J. 82 (1935) 248; 83 (1936) 187 and 257;
A.G. Walker, Proc. London Math. Soc. 2 42 (1937) 90.
 - [7] A.D. Dolgov, Pisma Zh. Eksp. Teor. Fiz. 29 (1979) 254.
 - [8] S.S. Gerstein, Ya.B. Zeldovich, Pis'ma Zh. Eksp. Teor. Fiz. 4 (1966) 74; JETP Lett. 4 (1966) 120.
 - [9] R. Cowsik, J. McClelland, Phys. Rev. Lett. 29 (1972) 669.
 - [10] J. Lesgourgues, S. Pastor, Phys. Rept. 429 (2006) 307;
S. Hannestad, Ann. Rev. Nucl. Part. Sci. 56 (2006) 137;
K. Ichikawa, M. Fukugita, and M. Kawasaki, Phys. Rev. D71 (2005) 043001;
M. Fukugita, K. Ichikawa, M. Kawasaki, O. Lahav, Phys. Rev. D74 (2006) 027302;
A.D. Dolgov, Phys. Atom. Nucl. 71 (2008) 2152.
 - [11] A.D. Dolgov, M. Fukugita, Pisma Zh. Eksp. Teor. Fiz. 56 (1992) 129; [JETP Lett.56 (1992) 123];
A.D. Dolgov, M. Fukugita, Phys. Rev. D46 (1992) 5378;
 - [12] A.D. Dolgov, Phys. Rept. 370 (2002) 333.
 - [13] R.E. Lopez, S. Dodelson, A. Heckler, M.S. Turner, Phys. Rev. Lett.82 (1999) 3952.
 - [14] A.D. Dolgov, Ya. B. Zeldovich, Rev. Mod. Phys. 53 (1981) 1; Uspekhi Fiz. Nauk, 139 (1980) 559.
 - [15] Ya.B. Zeldovich, Advances Astron. and Astophys. v. 3, Academic Press, 1965, p. 242.
 - [16] Ya.B. Zeldovich, L.B. Okun, S.B. Pikelner, Uspekhi Fiz. Nauk, 87 (1965) 113.
 - [17] B.W. Lee, S. Weinberg, Phys. Rev. Lett. 39 (1977) 165.
 - [18] M.I. Vysotsky, A.D. Dolgov, Ya.B. Zeldovich, JETP Lett. 26 (1977) 188, Pisma Zh. Eksp. Teor. Fiz. 26 (1977) 200.
 - [19] P. Gondolo, G. Gelmini, Nucl. Phys. B360 (1991) 145.
 - [20] A.M. Polyakov, JETP Lett. 20 (1974) 194, Pisma Zh. Eksp. Teor. Fiz. 20 (1974) 430;
G. 't Hooft, Nucl. Phys. B79 (1974) 276.

- [21] Ya.B. Zeldovich, M.Yu. Khlopov, Phys. Lett. B79 (1978) 239;
J. Preskill, Phys. Rev. Lett. 43 (1979) 1365.
- [22] C. Amsler et al., Physics Letters B667 (2008) 1, review on BBN by B.D. Fields, S. Sarkar.
- [23] V. Simha, G. Steigman, CAP 0806 (2008) 016.
- [24] L.D. Landau, E.M. Lifshitz, v. 5, *Statistical Physics*.
- [25] Z. Berezhiani, D. Comelli, F.L. Villante, Phys. Lett. B503 (2001) 362;
For a review see: Z. Berezhiani, Int. J. Mod. Phys. A19 (2004) 3775.
- [26] G. Sigl, G. Raffelt, Nucl. Phys. B 406 (1993) 423.
- [27] A.D. Dolgov, Sov. J. Nucl. Phys. 33 (1981) 700, Yad.Fiz.33 (1981) 1309;
D. Notzold, G. Raffelt, Nucl. Phys. B307 (1988) 924.
- [28] J. Pantaleone, Phys. Lett. B 287 (1992) 128.
- [29] A.D. Dolgov, S.H. Hansen, S. Pastor, S.T. Petcov, G.G. Raffelt, D.V. Semikoz, Nucl. Phys. B632 (2002) 363.
- [30] A.D. Dolgov, F. Takahashi, Nucl. Phys. B688 (2004) 189.
- [31] A.D. Dolgov, F.L. Villante, Nucl. Phys. B679 (2004) 261.
- [32] D. Kazanas, Astrophys. J. 241 (1980) L59.
- [33] A.H. Guth, Phys.Rev. D23 (1981) 347.
- [34] A.D. Linde, Phys. Lett. B108 (1982) 389.
- [35] A.J. Albrecht, P.J. Steinhardt, Phys. Rev. Lett. 48 (1982) 1220.
- [36] A.D. Linde, Phys. Lett. B129 (1983) 177.
- [37] A.D. Linde, *Particle physics and inflationary cosmology*, Chur, Switzerland: Harwood (1990) (Contemporary concepts in physics, 5); e-Print: hep-th/0503203;
W.H. Kinney, *TASI Lectures on Inflation*, e-Print: arXiv:0902.1529.
- [38] E.B. Gliner, ZhETF, 49 (1965) 542.
- [39] E.B. Gliner, I.G. Dymnikova, Astronomical Journal Letters, 1 (1975) 7.
- [40] V.Ts. Gurovich, A.A. Starobinsky Sov. Phys. JETP 50 (1979) 844; Zh. Eksp. Teor. Fiz. 77 (1979) 1683.
- [41] A.A. Starobinsky, JETP Lett.30 (1979) 682; Pisma Zh. Eksp. Teor. Fiz. 30 (1979) 719.

- [42] V.F. Mukhanov, G.V. Chibisov JETP Lett. 33 (1981) 532; Pisma Zh. Eksp. Teor. Fiz. 33 (1981) 549.
- [43] K. Sato, Mon. Not. Roy. Astron. Soc. 195 (1981) 467.
- [44] K. Sato, Phys. Lett. B99 (1981) 66.
- [45] A.J. Albrecht, P.J. Steinhardt, M.S. Turner, F. Wilczek, Phys. Rev. Lett. 48 (1982) 1437;
A.D. Dolgov, A.D. Linde, Phys. Lett. B116 (1982) 329;
L.F. Abbott, E. Farhi, M.B. Wise, Phys. Lett. B117 (1982) 29.
- [46] A.D. Dolgov, D.P. Kirilova, Sov. J. Nucl. Phys. 51 (1990) 172; Yad. Fiz. 51 (1990) 273;
J.H. Traschen, R.H. Brandenberger, Phys. Rev. D42 (1990) 2491.
- [47] L. Kofman, A.D. Linde, A.A. Starobinsky, Phys. Rev. Lett. 73 (1994) 3195; Phys. Rev. D56 (1997) 3258.
- [48] V. Kuzmin, I. Tkachev, Phys. Rev. D59 (1999) 123006.
- [49] S.K. Blau, A.H. Guth, *300 Years of Gravitation*, editors: S.W. Hawking, W. Israel, Cambridge Univ. Press (1987) p. 524.
- [50] V. Mukhanov, *Physical Foundations of Cosmology*, Cambridge, UK: Univ. Pr. (2005).
- [51] G. Steigman, Ann. Rev. Astron. Astrophys. 14 (1976) 339.
- [52] G. Steigman, arXiv:0808.1122.
- [53] A.G. Cohen, A. De Rujula, S.L. Glashow, Astrophys. J., 495 (1998) 495.
- [54] A.D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. 5 (1967) 32, JETP Lett. 5 (1967) 24.
- [55] T.D. Lee, C.-N. Yang, Phys. Rev., 104 (1956) 254;
C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes, R.P. Hudson, Phys. Rev., 105 (1957) 1413.
- [56] J.H. Christenson, J.W. Cronin, V.L. Fitch, R. Turlay, Phys. Rev. Lett., 13 (1964) 138.
- [57] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Lett. B 429 (1998) 263 (1998).
- [58] A.D. Dolgov, Lectures given at International School of Physics "Enrico Fermi": CP Violation: From Quarks to Leptons, Varenna, Italy, 19-29 Jul 2005 in "Varenna 2005, CP violation" p. 407, e-Print: hep-ph/0511213.
- [59] T.D. Lee, Phys. Rept. 9 (1974) 143.
- [60] Ya.B. Zeldovich, I.Yu. Kobzarev, L.B. Okun, Zh. Eksp. Teor. Fiz. 67 (1974) 3; Sov. Phys. JETP 40 (1974) 1.

- [61] V.A. Kuzmin, V.A. Rubakov, M.E. Shaposhnikov, Phys. Lett., B155 (1985) 36.
- [62] M. Fukugita, T. Yanagita, Phys. Lett., B174 (1986) 45.
- [63] I. Affleck, M. Dine, Nucl. Phys., B249 (1985) 361.
- [64] A.G. Cohen, D.B. Kaplan, Nucl. Phys., B308 (1988) 913.
- [65] Ya. B. Zeldovich, Pisma Zh. Eksp. Teor. Fiz., 24 (1976) 29;
A.D. Dolgov, ZhETF, 79 (1980) 337; Phys. Rev. D24 (1981) 1042.
- [66] R. Omnes, Phys. Rev. Lett. 23 (1969) 38; Phys. Rev. D1 (1969) 723.
- [67] G.R. Dvali, G. Gabadadze, Phys. Lett. B460 (1999) 47.
- [68] D.H. Oaknin, A. Zhitnitsky, Phys. Rev. D71 (2005) 023519;
D.H. Oaknin, A. Zhitnitsky, Phys. Rev. Lett. 94 (2005) 101301.
- [69] A. Dolgov, J. Silk, Phys. Rev. D47 (1993) 4244;
A.D. Dolgov, M. Kawasaki, N. Kevlishvili, Nucl. Phys. B807 (2009) 229.
- [70] C. Bambi, A.D. Dolgov, Nucl. Phys. B784 (2007) 132.
- [71] A.D. Dolgov, arXiv:0903.4318 and references therein.
- [72] A.D. Dolgov, Phys. Repts 222 (1992) No. 6;
A.D. Dolgov, Surveys in High Energy Physics, 13 (1998) 83, hep-ph/9707419;
V.A. Rubakov, M.E. Shaposhnikov, Usp. Fiz. Nauk, 166 (1996) 493, hep-ph/9603208;
A. Riotto, M. Trodden, Ann. Rev. Nucl. Part. Sci. 49 (1999), 35, hep-ph/9901362;
M. Dine, A. Kusenko, Rev. Mod. Phys. 76 (2004) 1.
- [73] G. 't Hooft, Phys. Rev. Lett., 37 (1976) 8; Phys. Rev., D14 (1976) 3432, *Erratum-
ibid.*, D18 (1978) 2199.