An Introduction to Standard Cosmology

Jorge L. Cervantes–Cota

Instituto Nacional de Investigaciones Nucleares (ININ)
Departamento de Física, Apartado Postal 18-1027, Col. Escandón, 11801 México D.F. jorge@nuclear.inin.mx

Abstract. A short introduction to the Standard Big Bang model is provided, presenting its physical model, and emphasizing its long–standing problems such as the horizon, flatness, baryon asymmetry, among others. Next, an introduction to the inflationary cosmology is presented to elucidate a solution to some of the above–mentioned problems. It is shown that the inflationary scenario succeeds in explaining what the standard Big Bang model cannot, passing the tests of the high precision experimental constraints which have been performed since last decade. This contribution should serve as an introduction to the standard ideas and scenarios which will be used in the forthcoming lectures of this book.

1 On the Standard Big Bang Model

We would like to begin our study by reviewing some basic aspects of the the standard hot Big Bang model (SBB), paying attention to what particle physics theories would bring about in the very early Universe. Our primary focus is to present the achievements of the SBB, but also some difficulties or conundrums that cannot be understood without the incorporation of other concepts, such as extensions to both gravity and particle physics theories, which will give rise to an inflationary scenario.

1.1 FRW Models

The SBB is based on Einstein's general relativity (GR) theory, which can be derived from the Einstein–Hilbert Lagrangian:

$$\mathcal{L} = \frac{1}{16\pi G} R\sqrt{-g} \ , \tag{1}$$

where R is the Ricci scalar, G the Newton constant, and $g = |g_{\mu\nu}|$ the determinant of the metric tensor; for our geometric conventions see the table provided in [68] (cover page), here we have used "-" for the metric $\mathbf{g}$, "+" for Riemann, and "-" for Einstein.

By performing the metric variation of this equation, one obtains the Einstein's well known field equations

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

J.L. Cervantes–Cota, An Introduction to Standard Cosmology, Lect. Notes Phys. **646**, 7–51 (2004)
http://www.springerlink.com/

where $R_{\mu\nu}$ is the Ricci tensor and $T_{\mu\nu}$ is the energy–momentum stress tensor. The left hand side (l.h.s.) of this equation represents the geometry, whereas the right hand side (r.h.s.) accounts for the fluid(s) present. In GR the space–time is four dimensional (three spatial dimensions plus time), and since both tensors are symmetric, (2) represents a collection of ten coupled, partial differential equations.

Once one is provided with the gravity theory, one should introduce a symmetry through the metric tensor. In cosmology one assumes a simple metric tensor according to the *cosmological principle* which states that the Universe is both homogeneous and isotropic. This turns out to be in very good agreement with the observed very–large–scale structure of the Universe. This homogeneous and isotropic space–time symmetry was originally studied by Friedmann, Robertson, and Walker (FRW); see [38, 81, 96]. The symmetry is encoded in the special form of the following line element:

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right] , \tag{3}$$

where t is the time variable, r–θ–ϕ are polar coordinates, which can be adjusted so that the constant curvature takes the values $k = 0, +1$, or -1 for a flat, closed, or open space, respectively. $a(t)$ is the scale factor of the Universe.

The FRW solutions to the Einstein equations (2) represent a cornerstone in the development of modern cosmology, since with them it is possible to understand the expansion of Universe, as was realized in the late 20s through Hubble's law of expansion [53]. With this metric, the GR cosmological field equations are,

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{4}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \ , \tag{5}$$

where H is the Hubble parameter; ρ and p are the density and pressure of the perfect fluid considered; that is, $T_{\mu\nu} = \rho u_\mu u_\nu + p(u_\mu u_\nu - g_{\mu\nu})$, where $u_\mu = \delta^0_\mu$ is the four–velocity of the fluid in co–moving coordinates, i.e. in coordinates that are moving with the expansion. Equations (4–5) can also be deduced within Newtonian cosmology, but there the pressure is not a source of gravitation; see the contribution of E. Copeland in this book.

The energy–momentum tensor conservation, ${T_\mu}^\nu{}_{;\nu} = 0$, is valid and from it one obtains that

$$\dot{\rho} + 3H(\rho + p) = 0 \ . \tag{6}$$

If one assumes further a barotropic equation of state for the fluid, $\omega =$ const.,

$$\frac{p}{\rho} = \omega = \begin{cases} \frac{1}{3} & \text{for} \quad \text{radiation and/or ultra} - \text{relativistic matter} \\ 0 & \text{for} \quad \text{dust} \\ 1 & \text{for} \quad \text{stiff fluid} \\ -1 & \text{for} \quad \text{vacuum energy} \end{cases} \tag{7}$$

to integrate (6), it yields

$$\rho = \frac{M_\omega}{a^{3(1+\omega)}} \ , \tag{8}$$

where M_ω is the integration constant and is differently dimensioned by considering different $\omega-$fluids. Equations (4), (5), and (6) are not linearly independent, only two of them are. That is, one can derive, e.g., (5) from (4) and (6). Note that these equations are time symmetric, the interchange $t \to -t$ leaves the equations the same.

Let us very briefly recall which $\omega-$values are needed to describe the different epochs of the Universe's evolution. At very early times, the Universe is believed to have experienced a huge expansion due to some cosmological constant ($\Lambda = 8\pi G\rho$, where $\rho =$ const.) or vacuum energy. This epoch, to be fully described later on, is roughly characterized by an equation of state $\omega = -1$. After inflation, the vacuum energy decays in some particle content, a process called *reheating* [3, 33, 1], after which the Universe is filled with a "fluid" of radiation or of ultra–relativistic matter where the material content of the Universe consisted of photons, neutrinos, electrons, and other massive particles with very high kinetic energy. During this epoch the assumption $\omega = 1/3$ is valid. After some Universe cooling, some massive particles decayed and others survived (protons, neutrons, electrons) and their masses eventually surpassed the radiation components (photons, neutrinos). From that epoch until very recent times, the matter content dominated and effectively produced no pressure on the expansion and, therefore, one accepts a model filled with dust, i.e. $\omega = 0$. Until the mid 90s we thought that a dust model would be representative for the current energy content of the Universe. Recent measurements (see contribution of A. Filippenko in this book), however, indicate that as of recently the Universe is again experiencing a huge expansion rate. It is believed that a kind of cosmological constant, or vacuum energy, is the largest energy contribution to the expansion of the Universe at present. Thus the cosmological constant is the generic factor of an inflationary solution, see the $k = 0$ solution below, (10), which is believed to be characteristic of both the very early inflationary epoch and today.

Finally, a stiff model, $\omega = 1$, is sometimes considered in order to describe very dense matter under very high pressures.

The ordinary differential equations system described above needs a set of either initial conditions or boundary conditions to be integrated. One can assume a set of two initial values, say, $(\rho(t_*), \dot{a}(t_*)) \equiv (\rho_*, \dot{a}_*)$ at some (initial) time t_*, in order to determine its evolution. Its full analysis has been reviewed by many authors [97, 68]. Here, in order to show some early Universe consequences we take $k = 0$, justified as follows: From (4) and (8) one notes that the expansion rate, given by the Hubble parameter, is dominated by the density term as $a(t) \to 0$, since $\rho \sim 1/a^{3(1+\omega)} > k/a^2$ for $\omega > -1/3$; that is, the flat solution is very well fitted at the very beginning of time. Furthermore, recent Cosmic Microwave Background Radiation (CMBR) [1] measurements [26, 6, 48, 13] are consistent with $k \sim 0$. Therefore, assuming $k = 0$, (4) implies

$$a(t) = [6\pi G M_\omega (1+\omega)^2]^{\frac{1}{3(1+\omega)}} (t-t_*)^{\frac{2}{3(1+\omega)}}$$
$$= \begin{cases} (\frac{32}{3}\pi G M_{\frac{1}{3}})^{1/4} \; (t-t_*)^{1/2} & \text{for} \quad \omega = \frac{1}{3} \text{ radiation} \\ (6\pi G M_0)^{1/3} \; (t-t_*)^{2/3} & \text{for} \quad \omega = 0 \text{ dust} \\ (24\pi G M_1)^{1/6} \; (t-t_*)^{1/3} & \text{for} \quad \omega = 1 \text{ stiff fluid} \end{cases} \tag{9}$$

and

$$a(t) = a_* e^{Ht} \qquad \text{for} \qquad \omega = -1 \text{ vacuum energy} \tag{10}$$

where the letters with a subindex "$*$" are integration constants, representing quantities evaluated at the beginning of times, $t = t_*$.

From (9) one can immediately see that at $t = t_*$, $a_* = 0$ and from (8), $\rho_* = \infty$; that is, the solution has a singularity at that time, presumably at the Universe's beginning; this initial cosmological singularity is also called *Big Bang* singularity. As the Universe expands the Hubble parameter evolves as $H \sim 1/t$, i.e. the expansion rate decreases; whereas the matter–energy content acts as an expanding agent, cf. (4), it also decelerates the expansion in an asymptotically decreasing manner, cf. (5) and (8). In that way, H^{-1} represents an upper limit to the age of the Universe; for instance, $H^{-1} = 2t$ for $\omega = 1/3$ and $H^{-1} = 3t/2$ for $\omega = 0$, t being the Universe's age.

The solution (10) is inflationary and has no singularity. This solution is such that the Hubble parameter is indeed a constant. A fundamental ingredient of inflation is that the r.h.s. of (5) remains positive, $\ddot{a} > 0$. This is performed when the *inflation pressure* is negative [18], $\rho + 3p < 0$. In this way, one does not have necessarily to impose the strong condition $\omega = -1$, but it suffices that $\omega < -1/3$, in order to have a moderate inflationary solution; for example, $\omega = -2/3$ it implies $a = a_* t^2$, a mild power-law inflation. The issue of inflation will be discussed in Sect. 2.

1.2 The Physical Scenario

So far we have obtained some exact solutions for Einstein's cosmology. Now, to achieve a more physical scenario one considers the Universe filled with a plasma of particles and their antiparticles. This was originally done by G. Gamow [40], who first considered a hot Big Bang model for the Universe's

[1] The CMBR is also sometimes referred to in this book as Cosmic Microwave Background (CMB) or Cosmic Background Radiation(CBR).

beginning, which was later qualitatively confirmed by Penzias and Wilson [73] and interpreted by Dicke et. al. [29]. Furthermore, with the development of modern particle physics theories in the 70s it was unavoidable to think about a physical scenario for the early Universe which should include even the "new" physics. It was also realized that the physics described by GR should not be applied beyond Planck (Pl) initial conditions, because there the quantum corrections to the metric tensor become very important, a theory which is still in progress. Thus, we make some assumption at some early time, $t \gtrsim t_{Pl}$: the Universe was filled with a plasma of relativistic particles, including quarks, leptons, and gauge and Higgs bosons, all in thermal equilibrium at a very high temperature, T, with some gauge symmetry dictated by a particle physics theory.

Now, in order to work in that direction one introduces some thermodynamic considerations necessary for the description of the physical content of the Universe, which we would like to present here. Assuming an ideal-gas approximation, the number density n_i of the particles of type i, with a momentum q, is given by a Fermi or Bose distribution [60]:

$$n_i = \frac{g_i}{2\pi^2} \int \frac{q^2 dq}{e^{(E_i - \mu_i)/T} \pm 1} \ , \tag{11}$$

where $E_i = \sqrt{m_i^2 + q^2}$ is the particle energy, μ_i is the chemical potential, the sign $(+)$ applies for fermions and $(-)$ for bosons, and g_i is the number of spin states. One has that $g_i = 2$ for photons, quarks, baryons, electrons, muons, taus, and their antiparticles, but $g_i = 1$ for neutrinos because they are only left-handed. For the particles existing in the early Universe one usually assumes that $\mu_i = 0$: one expects that in any particle reaction the μ_i are conserved, just as the charge, energy, spin, and lepton and baryon number are conserved as well. The number density of photons (n_γ), which can be created and/or annihilated after some particle collisions, must not be conserved and its distribution with $\mu_\gamma = 0, \ E = q = h\nu$, reduces to the Planck one. For other constituents, in order to determine the μ_i, one needs n_i; one notes from (11) that for large $\mu_i > 0, \ n_i$ is large too. One does not know n_i, but from nucleosynthesis that [72]

$$\eta \equiv \frac{n_B}{n_\gamma} \equiv \frac{n_{\text{baryons}} - n_{\text{anti-baryons}}}{n_\gamma} \approx (3-4) \times 10^{-10} \ . \tag{12}$$

The smallness of the baryon number density, n_B, relative to the photon's, suggests that n_{leptons} may also be small compared with n_γ. Therefore, one takes for granted that $\mu_i = 0$ for all particles. Why the ratio n_B/n_γ is so small, but not zero, is one of the puzzles of the SBB. This ratio is also often called $\eta \equiv n_B/n_\gamma$.

The above approximation allows one to treat the density and pressure of all particles as a function of the temperature only. According to the second law of thermodynamics, one has [97]

$$dS(V,T) = \frac{1}{T}[d(\rho V) + pdV] \ , \tag{13}$$

where S is the entropy in a volume $V \sim a^3(t)$ with $\rho = \rho(T)$, $p = p(T)$ in equilibrium. Furthermore, the integrability condition $\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T}$ is also valid, which turns out to be

$$\frac{dp}{dT} = \frac{\rho + p}{T} \ . \tag{14}$$

Additionally, the energy conservation law equation (6) leads to

$$a^3(t)\frac{dp}{dt} = \frac{d}{dt}[a^3(t)(\rho + p)] \tag{15}$$

and using (14), the latter takes the form

$$\frac{d}{dt}[\frac{a^3(t)}{T}(\rho + p)] = 0 \ . \tag{16}$$

Using (14), (13) can be written as

$$dS(V,T) = \frac{1}{T}d[(\rho + p)V] - \frac{V}{T^2}(\rho + p)dT \ . \tag{17}$$

Then, (16) together with (17) imply that the entropy

$$S = \frac{a^3}{T}[\rho + p] = \text{const.} \tag{18}$$

is a constant of motion.

The density and pressure are given by

$$\rho \equiv \int E_i n_i dq \quad \text{and} \quad p \equiv \int \frac{q^2}{3E_i} n_i dq \ . \tag{19}$$

For photons or ultra–relativistic fluids, $E = q$, these equations become such that

$$p = \frac{1}{3}\rho \ , \tag{20}$$

confirming (7), and after integrating (14), it results that

$$\rho = bT^4 \tag{21}$$

with the constant of integration, b. In a real scenario there are many relativistic particles present, each of which contributes as in (21). Summing up all of them, $\rho = \sum_i \rho_i$ and $p = \sum_i p_i$ over all relativistic species, it results that $b(T) = \frac{\pi^2}{30}(N_B + \frac{7}{8}N_F)$, which depends on the number of effective relativistic degrees of freedom of bosons (N_B) and fermions (N_F). Therefore, this quantity varies with the temperature; different $i-$species remain relativistic until some characteristic temperature $T \approx m_i$, after which the value N_{F_i} (or N_{B_i}) no longer contributes to $b(T)$. The factor 7/8 accounts for the different statistics the particles have, see (11). In the standard model of particles physics $b \approx 1$ for $T \ll 1$ MeV and $b \approx 35$ for $T > 300$ GeV. Additionally, for relativistic particles one obtains from (11) that

$$n = cT^3, \quad \text{with } c = \frac{\zeta(3)}{\pi^2}(N_B + \frac{3}{4}N_F) \ . \tag{22}$$

where $\zeta(3) \approx 1.2$ is the Riemann zeta function of 3. Currently, $n_\gamma \approx \frac{422}{\text{cm}^3} T^3_{2.75}$, where $T_{2.75} \equiv \frac{T_{\gamma 0}}{2.75°K}$; the subscript "0" refers to quantities evaluated at present time.

From (18), using (20) and (21), one concludes that $T \sim 1/a(t)$ and from the $\omega = 1/3$ solution in (9) one arrives at the result

$$T = \sqrt[4]{\frac{M_{\frac{1}{3}}}{b}} \frac{1}{a(t)} = \sqrt[4]{\frac{3}{32\pi G b}} \frac{1}{(t - t_*)^{\frac{1}{2}}} \ , \tag{23}$$

a decreasing temperature behavior as the Universe expands. Therefore, initially at the Big Bang $t = t_*$ implies $T_* = \infty$, the Universe was very hot.

The entropy for an effective relativistic fluid is given by (18) together with (20) and (21):

$$S = \frac{4}{3}\ b\ (a\ T)^3 = \text{const.} \tag{24}$$

Combining this with (23), one can compute the value of $M_{\frac{1}{3}}$ to be $M_{\frac{1}{3}} = (\frac{3}{4}S)^{4/3}/b^{1/3} \approx 10^{116}$, since $b \approx 35$ and the photon entropy $S_0 = \frac{4}{3}\ b\ (a_0\ T_0)^3 \approx 10^{88}$ for the currently evaluated quantities $a_0 = d_H(t_0) = 10^{28}$cm and $T_{\gamma_0} = 2.7\ °K$. For later convenience, we define the entropy per unit volume, *entropy density*, to be $s \equiv S/V = \frac{4}{3}\frac{\pi^2}{30}(N_B + \frac{7}{8}N_F)T^3$; thus, currently $s \approx 7n_\gamma$. The nucleosynthesis bound on η, (12), implies that $n_B/s \approx (4–6) \times 10^{-11}$.

Now we consider particles in their non–relativistic limit ($m \gg T$). From (11) one obtains for both bosons and fermions that

$$n = g \left(\frac{mT}{2\pi}\right)^{3/2} e^{-m/T} \ . \tag{25}$$

The abundance of equilibrium massive particles decreases exponentially once they become non–relativistic; this situation is referred as *in equilibrium annihilation.* Their density and pressure are given through (19) and (25) by

$$\begin{aligned} \rho &= nm \\ p &= nT \ll \rho \ . \end{aligned} \tag{26}$$

Therefore, the entropy given by (18) for non–relativistic particles through (25) and (26) also diminishes exponentially during their in equilibrium annihilation. The entropy of these particles is transferred to that of relativistic components by augmenting their temperature. Hence, the constant total entropy is essentially the same as that given by (24), but the $i-$species contributing to it are just those which are in equilibrium and maintain their relativistic behavior, that is, particles without mass such as photons.

Having introduced the abundances of the different particle types, we would like to comment on the equilibrium conditions for the constituents of the Universe as it evolves. This is especially of importance in order to have an idea whether or not a given $i-$species disappears or decouples from the primordial brew. To see this, let us consider n_i when the Universe's temperature, T, is such that (a) $T \gg m_i$, during the ultra–relativistic stage of some particles of type i and (b) $T \ll m_i$, when the particles i are non–relativistic; both cases originally in thermal equilibrium. From (22) one obtains for case (a) that $n_i \sim T^3$; the total number of particles, $\sim n_i a^3$, remains constant. Whereas for case (b), from (25), $n_i \sim T^{3/2} e^{-m_i/T}$, i.e. when the Universe temperature goes down below m_i, the number density of the $i-$species significantly diminishes; an "in equilibrium annihilation" occurs. Let us take as an example the neutron–proton annihilation: one then has

$$\frac{n_n}{n_p} \sim e^{\frac{m_p - m_n}{T}} = e^{-\frac{1.5\times 10^{10} \ ^\circ K}{T}} \tag{27}$$

which drops with the temperature, from nearly 1 at $T \geq 10^{12} \ ^\circ K$ to about 5/6 at $T \approx 10^{11} \ ^\circ K$, and 3/5 at $T \approx 3 \times 10^{10} \ ^\circ K$ [70]. If this is forever valid, one ends up without massive particles, meaning that our Universe should have consisted only of radiative components; our own existence contradicts that! Therefore, the in–equilibrium annihilation eventually stopped. The quest is now to freeze out this ratio (to be $n_n/n_p \approx 1/6$) [2] in order to leave some hadrons for posteriorly achieving successful nucleosynthesis. The answer comes by comparing the Universe expansion rate, H, with particle physics reaction rates, Γ. Hence, for $H < \Gamma$, the particles interact with each other faster than the Universe expansion rate and then equilibrium is established. For $H > \Gamma$ the particles cease to interact effectively and then thermal equilibrium drops out [3]. In this way, the more interacting the particles are, the longer they remain in equilibrium annihilation and, therefore, the lower their number densities are after some time; e.g., baryons vanish first, then charged leptons,

[2] Due to neutron decays, until the time when nucleosynthesis begins, n_n/n_p reduces to 1/7.

[3] This is only approximately true; a proper account of this involves a Boltzmann equation analysis. In doing so a numerical integration should be carried out in which annihilation rates are balanced with inverse processes; see for example [90, 60].

neutral leptons, etc. Finally, the particle numbers of (massless) photons and neutrinos remain constant, as it was mentioned above; see Fig. 1. Note that if interactions of an $i-$species freeze out when it is still relativistic, then its abundance can be significant at present.

It is worth mentioning that if the Universe were to expand faster, then the temperature of decoupling at $H \sim \Gamma$ would be higher, thus the fixed ratio n_n/n_p would be greater, thus leading to profound implications in the nucleosynthesis of the light elements. For instance the Helium, ^{4}He, abundance should be higher. Therefore, the expansion of the Universe cannot arbitrarily be augmented during the equilibrium era of some particles. Furthermore, if a particle species is still highly relativistic ($T \gg m_i$) or highly non–relativistic ($T \ll m_i$) when decoupling from primordial plasma occurs, it maintains an equilibrium distribution; the former characterized by $T_r a$ =const. and the latter by $T_m a^2$ =const., cf. (30).

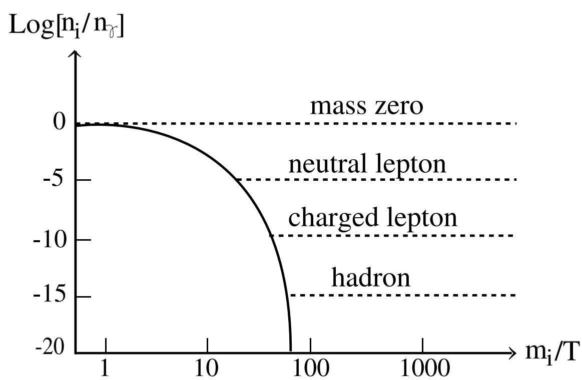

Fig. 1. The evolution of the particle density of different $i-$species. If an $i-$species is in equilibrium its abundance diminishes exponentially after the particle becomes non–relativistic (solid line). However, interactions of an $i-$species can freeze out, causing the particle species to decouple from equilibrium and maintain its abundance (dashed line). (Figure adapted from Kolb and Turner 1990).

There are also some other examples of decoupling, like the neutrino decoupling: during nucleosynthesis there exist reactions like $\nu\bar{\nu} \longleftrightarrow e^+e^-$, which maintain neutrinos efficiently coupled to the original plasma ($\Gamma > H$) until about 1 MeV, since $\frac{\Gamma}{H} \approx \left(\frac{T}{\mathrm{MeV}}\right)^3$. Below 1 MeV reactions are no longer efficient and neutrinos decouple and continue evolving with a temperature $T_\nu \sim 1/a$. Then, at $T \gtrsim m_e = 0.51$MeV the particles in equilibrium are photons (with $N_B = 2$) and electron and positron pairs (with $N_F = 4$) which contribute to the entropy with $b(T) = \frac{\pi^2}{30} \cdot (11/2)$. Later, when the temperature drops to $T \ll m_e$, the reactions are no longer efficient ($\Gamma < H$) and after the $e^\pm$ pair annihilation there are only photons in equilibrium with

$b(T) = \frac{\pi^2}{30} \cdot (2)$. Since the total entropy, $S = \frac{4}{3}b(aT)^3$, must be conserved, the decrease in $b(T)$ must be balanced with an increase in the radiation temperature; this gives a result of $\frac{T_\gamma}{T_\nu} = \left(\frac{11}{4}\right)^{1/3}$, which should remain to the present day, implying the existence of a cosmic background of neutrinos with a temperature today of $T_{\nu_0} = 1.96\ °K$.

Another example of this is the gravitation decoupling, which should also be present if gravitons were in thermal equilibrium at the Planck time and then decoupled. The present–day background of temperature should be characterized at most by $T_{\rm grav.} = \left(\frac{4}{107}\right)^{1/3} \approx 0.91\ °K$.

For the matter dominated era we have stressed that effectively $p = 0$; next we will see the reason for this: First consider an ideal gas (like atomic Hydrogen) with mass m, then $\rho = nm + \frac{3}{2}nT_m$ and $p = nT_m$. From (15) one obtains, equivalently, that

$$\frac{d}{da}(\rho a^3(t)) = -3pa^2(t) \tag{28}$$

and substituting the above ρ and p, one obtains

$$\frac{d}{da}(nma^3(t) + \frac{3}{2}nT_m a^3(t)) = -3nT_m a^2(t) \tag{29}$$

where $nma^3(t)$ is a const. This equation yields

$$T_m a^2(t) = \text{const.}\ , \tag{30}$$

that matter temperature drops faster than that of radiation as the Universe expands; see (23). Now, if one considers both radiation and matter, it is valid that $\rho = nm + \frac{3}{2}nT_m + bT_r^4$ and $p = nT_m + \frac{1}{3}bT_r^4$; the source of the Universe's expansion is proportional to $\rho + p = nm + \frac{5}{2}nT_m + \frac{4}{3}bT_r^4$; the first term dominates the second, precisely because T_m decreases very rapidly. The third term diminishes as $\sim 1/a^4$, whereas the first as $\sim 1/a^3$, and after the time of densities equality (eq.), $\rho_m = \rho_r$, the matter density term is greater than the others, which is why one assumes no pressure for that era.

From now on, when we refer to the temperature, T, it should be related to the radiation temperature.

The detailed description of the thermal evolution of the Universe for the different particle types, depending on their masses, cross-sections, etc., is well described in many textbooks, going from the physics known in the early 70s [97] to the late 80s [60], or late 90s [62], and therefore it will not be presented here. However, we notice that as the Universe cools down, a series of spontaneous symmetry–breaking phase transitions are expected to occur. The type and/or nature of these transitions depend on the specific particle physics theory considered. Among the most popular are Grand Unification theories (GUT) which bring together all known interactions except for gravity. One could also be more modest and just consider the standard model of particle physics or some extensions of it. Ultimately, one should decide, in constructing a cosmological theory, according to which energy scale one wants to use to describe physics. For instance, at a temperature between 10^{14} GeV to 10^{16} GeV the transition of the $SU(5)$ GUT should take place -if this theory is valid- in which a Higgs field breaks this symmetry to $SU(3)_C \times SU(2)_W \times U(1)_{HC}$, a process through which some bosons acquire their masses. Due to the gauge symmetry, there are color (C), weak (W) and hypercharge (HC) conservation, as the subindexes indicate. Later on, when the Universe evolves to about a few hundred GeV, the electroweak phase transition takes place, in which a second Higgs field breaks the symmetry $SU(3)_C \times SU(2)_W \times U(1)_{HC}$ to $SU(3)_C \times U(1)_{EM}$; through this second breaking the fermions also acquire their masses. At this stage, there are only color and electromagnetic (EM) charge conservation, due to the gauge symmetry. Afterwards, at a temperature of about 100 to 300 MeV the Universe should undergo a transition associated to the chiral symmetry–breaking and color confinement, from which baryons and mesons are formed out of quarks. Subsequently, at approximately 10 MeV the synthesis of light elements (nucleosynthesis) begins, producing most of the observed Hydrogen and Helium observed in the present day, along with abundances of some other light elements. The nucleosynthesis represents the earliest scenario tested in the SBB. After some time, matter dominates, over radiation components, in the Universe, and the large scale structure (galaxies, clusters, superclusters, voids, etc.) begins to form. At about 1 eV the *recombination* takes place; that is, the Hydrogen ions and electrons combine to compose neutral Hydrogen atoms, then matter and EM radiation decouple from each other. At this moment the surface of last scattering (ls) of the CMBR evolves as an imprint of the Universe at that time. In Fig. 2 the main events of the SSB are sketched.

Let us go back to the FRW cosmological equations. In observing the two terms involved in (4), the matter term $8\pi G\rho/3$ and the curvature term k/a^2, one should be aware of the validity of the approximation $8\pi G\rho/3 > k/a^2$. Let us for the moment elucidate that k is tiny but different from zero. Then, eventually when the energy density has diminished enough due to the expansion, $8\pi G\rho/3 \sim k/a^2$, and further on the Universe will be dominated by its curvature. Let us consider this case, but for both $k = \pm 1$ separately. First, take $k = -1$, then $H = 1/a$ and the solution is $a \sim t$, that is, the Universe expands forever. Otherwise, for $k = +1$, at the moment of maximum expansion, say[4] $\tau_c/2$, $8\pi G\rho/3 = k/a^2$, the Universe stops its expansion and then the scale factor begins to decrease. The solution given by the negative square root of (4) again ends with a singularity but now at $t = \tau_c > t_*$, where

[4] τ_c stands for τ_{collapse}. The lifetime of such an Universe, the time of a cycle, is just twice the time of maximal expansion, because the solution is time symmetric.

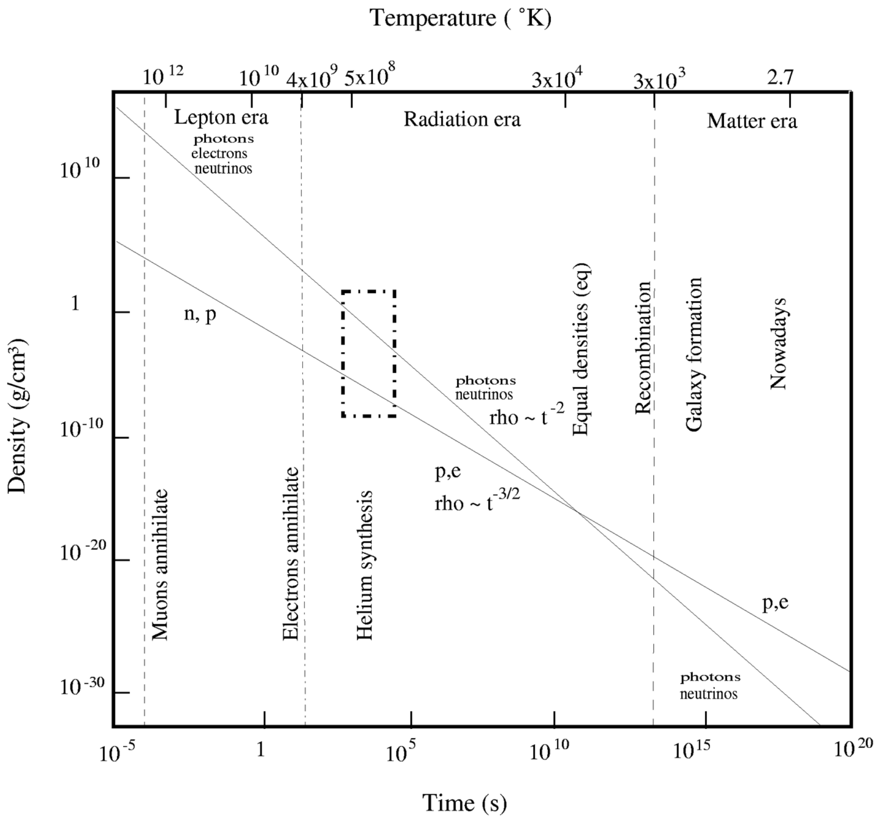

Fig. 2. The thermal history of the SBB (Figure adapted from Harrison 1970).

$\rho = \infty$, $T = \infty$, and $a = 0$; this is the so-called Big Crunch. For instance, the lifetime τ_c for a model filled with cold dark matter is

$$\tau_c \approx MG \approx \frac{M}{M_{Pl}} 10^{-43} \text{ s} ; \tag{31}$$

that is, if our Universe were dark–matter dominated and with a closed curvature, then it must presently have (10^{17} s) a mass $M > 10^{60} M_{Pl} \approx 10^{55}$gr.

For a radiation model, in terms of its entropy, from (18) and (23), one obtains that

$$\tau_c \approx \frac{S^{2/3}}{M_{Pl}} \approx S^{2/3} 10^{-43} \text{ s} , \tag{32}$$

in this case $S > 10^{90}$, a huge entropy! In trying to understand such big numbers one is forced to recognize some problems with the SBB. Next, we present some of them.

1.3 Problems of the Standard Big Bang Model

In considering a theory of the Universe one is open to think about an Universe's "arena" as general as possible. In doing so one finds a large list of problems to be understood. However, not all of them are of the same nature. For instance, some problems arise as a result of computations, others by implementing a physical scenario for GUT or from the conception itself of how the Universe should have begun; that is, apart from the choice of initial conditions, $(\rho_*, \dot{a}_*)$, are the number of dimensions or the global topology. In this section we list some of these problems, emphasizing those for which the inflationary Universe offers an explanation:

Dimensionality

Why should the Universe have four space–time dimensions, at least locally in our surroundings? A first attempt to consider theories in more dimensions was carried out by Kaluza and Klein [54, 56], who tried, unsuccessfully, to unify gravity with the electromagnetic interaction. However, from that we learned to use more than four dimensions for unifying meanings.

Other theories such as fundamental strings are conceivable in $D-$dimensions, but by demanding Lorentz invariance of the quantized bosonic string theory one has to choose $D = 26$, or in fermionic strings $D = 10$. Yet, there arises the problem of compactifying the $D-4$ dimensions, to a compact space whose size is of the order of the Planck length (l_{Pl}). There is no unique procedure. The compactification can be achieved in a number of ways, many of them casting different particle content in their low energy effective Lagrangians. In addition, there exists no compelling principle which would determine the space–time dimension to be four. All dimensions below D seem to be on an equal footing [66].

Euclidicity

What is the global geometry of the Universe? The space geometry is almost perfectly Euclidean on large scales, but on very small scales -say, slightly smaller than the Planckian- GR is not any more tractable, as quantum fluctuations of the metric make it impossible to extend a classical formalism. Within GR one understands a large–scale euclidicity, but not at the very small scale, even though the only natural length in GR is $l_{Pl} = \sqrt{G}$. Why this? Naturally, it is tempting to go beyond GR, a theory which is not yet completed.

Singularity

As we have already mentioned, at $t = t_*$ the scale factor is $a = 0$, the density $\rho = \infty$ and $T = \infty$, see (9), (8) and (23). It can also be shown that the curvature tensor $R^{\mu}{}_{\nu\gamma\delta} = \infty$ at that time. There exists no such a theory to explain gravity as $a(t)$ approaches zero. In fact, one expects GR to be valid as far as $a(t) \to l_{Pl}$; in going beyond this limit the problems mentioned in the Euclidicity item appear.

Homogeneity and Isotropy

The large–scale structure of the Universe seems to be very homogeneous and isotropic. However, looking on small scales, the isotropy and homogeneity break down: there exist planets, stars, compact objects, galaxies, clusters...a large–scale structure. Hence, it is tempting to consider more general inhomogeneous and anisotropic models, which should explain, as a consequence of their evolution, the currently observed large–scale structure along with the isotropy limits observed in the CMBR, in x–ray backgrounds (e.g. quasars at high redshift), and in number counts in faint radio sources.

In GR, without the aid of a cosmological constant or inflation, Collins and Hawking [25] examined the question in terms of an "initial conditions" analysis. They obtained that the set of spatially homogeneous cosmological models approaching isotropy in the limit of infinite times is of measure zero in the space of all spatially homogeneous models. This in turn implies that the isotropy of the models is unstable to homogeneous and anisotropic perturbations. However, their definition of isotropization demands asymptotic stability of the isotropic solution. An asymptotic stability analysis of Bianchi models in GR [10] shows, e.g., that in the Bianchi type VII_h the anisotropy will not exactly vanish but can be bounded. In this sense, the open FRW model may be stable. Attempts to understand this question in other gravity theories, such as Brans–Dicke theory, shed some light on the solution [20].

Horizon

The region of space which can be connected to some other region by causal physical processes, at most through the propagation of light with $ds^2 = 0$, defines the *causal* or *particle horizon*, d_H. For the FRW equation (3), in spherical coordinates with θ, ϕ =const. and after redefining r, this means that [82, 97]:

$$\int_0^t \frac{dt}{a(t)} = \int_0^{r_H} \frac{dr}{\sqrt{1-kr^2}}$$

$$d_H(t) \equiv a(t) \int_0^{r_H} \frac{dr}{\sqrt{1-kr^2}} \quad . \tag{33}$$

In order to analyze the whole horizon evolution, from the present (t_0) to the Planck time (t_{Pl}), we first compute the horizon for the matter dominated era $t_{\rm eq.} \leq t \leq t_0$ and secondly for the radiation era $t \leq t_{\rm eq.}$, because they are differently determined by (9), where we set $t_* = 0$ for convenience. For the matter epoch one has $a(t) = a_0(t/t_0)^{2/3}$, then the first equation above gives $r_H = \frac{3}{a_0}(t_0^2\ t)^{1/3}$; from the second equation one obtains the horizon $d_H(t) = 3t = 2H^{-1}$. For the radiation period, one finds that $r_H = \frac{2}{a_{\rm eq.}}(t_{\rm eq.}\ t)^{1/2}$ and $d_H(t) = 2t = H^{-1}$. We see for the matter dominated era that the causal horizon is twice the Hubble distance, H^{-1}, and that they are equal to each other during the radiation dominated era; therefore, one uses them interchangeably. It is clearly seen for both eras that as $t \to 0$, the Universe is causally disconnected, being $a(t) > d_H(t)$.

The evolution of a typical co–moving distance scale, L, due to the Universe expansion is given by $L(t) = L_0\frac{a(t)}{a_0}$. Next, let us compare the past evolution of that scale with the corresponding traced by the horizon, $d_H(t) = d_{H_0}(t/t_0)$, where $d_{H_0} = 3t_0$, for the matter dominated era. Then, one finds that

$$\frac{d_H}{L} = \frac{d_{H_0}}{L_0}\left(\frac{t}{t_0}\right)^{1/3} \qquad \text{for} \qquad t_{\rm eq.} \le t \le t_0\ . \tag{34}$$

Now consider the typical scale, L_0, to be the present observed particle horizon, $L_0 = d_{H_0}$. Then, the amount by which the three dimensional horizon was smaller than the "volume" $L^3(t)$ is determined by the following relation:

$$\left(\frac{d_H}{L}\right)^3 = \frac{t}{t_0} = \left(\frac{T_0}{T}\right)^{3/2} \qquad \text{for} \qquad t_{\rm eq.} \le t \le t_0\ , \tag{35}$$

in which we have made use of (9) and (24). At the time when the CMBR was last scattered (ls) one has then that $\left(\frac{d_H}{L}\right)^3_{\rm ls} \approx 10^{-5}$; that is, there were approximately one hundred thousand small horizon regions without causal connection! But, on the other hand, by that time the CMBR was already highly isotropic. Thus, one has to take for granted that the initial conditions for all the 10^5 volume horizons were fine tuned so as to account for the present observed large angle CMBR levels of isotropy, with $\delta T/T \sim 10^{-5}$. This is the horizon problem.

One can go further and compute the number of disconnected regions up to the Planck epoch. But first, one needs to evaluate the ratio d_H/L when the radiation and matter densities equal (eq.) each other; this is $\left(\frac{d_H}{L}\right)^3_{\rm eq.} \approx 10^{-6}$. Up until this time, one has to use the radiation solution given by

$$\frac{d_H}{L} = \frac{d_{H_{\rm eq.}}}{L_{\rm eq.}}\left(\frac{t}{t_{\rm eq.}}\right)^{1/2} \qquad \text{for} \qquad t \le t_{\rm eq.}\ , \tag{36}$$

again taking as the typical scale that of the horizon at that time, which is given by $\left(\frac{d_H}{L}\right)^3_{\rm eq.} \approx 10^{-6}$. Then, one finds that

$$\left(\frac{d_H}{L}\right)^3 = 10^{-6}\left(\frac{t}{t_{\rm eq.}}\right)^{3/2} = 10^{-6}\left(\frac{T_{\rm eq.}}{T}\right)^3 \qquad \text{for} \qquad t \le t_{\rm eq.}\ , \tag{37}$$

which at the time of nucleosynthesis (ns) is $\left(\frac{d_H}{L}\right)^3_{\rm ns} \approx 10^{-24}$; then one has to tune the initial conditions even finer (than at $t = t_{eq}$) to explain the homogeneous Universe element composition. Further, at Planck time it yields $\left(\frac{d_H}{L}\right)^3_{Pl} \approx 10^{-89}$, that is, $\frac{L}{d_H} \approx 10^{29} \approx e^{68}$; such large numbers will later be explained in a successful inflationary model.

Now, let us try to link this issue with the big numbers encountered in (31) and (32). To do that, next we compute the entropy per horizon, S_H, using (24), finding

$$S_H = \frac{4}{3}\, b\, (d_H\, T)^3 \ , \tag{38}$$

now using (34), (24), and (9) for the matter dominated era and (36), (24), and (23) for the radiation era, we obtain the following results:

$$S_H = S \left(\frac{T_0}{T}\right)^{3/2} \qquad \text{for} \qquad t_{\rm eq.} \leq t \leq t_0 \tag{39}$$

and

$$S_H = S \left(\frac{d_{H_{\rm eq.}}}{L_{\rm eq.}}\right)^3 \left(\frac{T_{\rm eq.}}{T}\right)^3 = S \times 10^{-6} \left(\frac{T_{\rm eq.}}{T}\right)^3 \qquad \text{for} \qquad t \leq t_{\rm eq.} \ , \tag{40}$$

where $S = 10^{88}$ and should be a constant of motion; see (18). From these equations one obtains at $t = t_{\rm ls}$ that $S_{H_{\rm ls}} = 10^{83}$. At a typical time during the nucleosynthesis one finds $S_{H_{\rm ns}} = 10^{63}$, and so on, until the Planck time, where $S_{H_{Pl}} \approx 1$. That is to say that the horizon problem is related to the increase of the horizon entropy as the Universe expands: this increase should be such that currently $S_{H_0} \gtrsim 10^{88}$ can explain the Universe's age, cf. (32). The evolution of horizon entropy in the standard Big Bang model is depicted in Fig. 3. Within the context of the *anthropic principle*, the existence of such big numbers invites us to reflect on our own existence; why are they so big (or so small)? The anthropic principle states that only in this way can life exist to account for it!

Flatness

Why is our Universe today nearly flat, and why was it almost identically flat at the very beginning? [30]. From (4) and (8) one finds that

$$\begin{aligned} \Omega(t) - 1 = \frac{\rho - \rho_c}{\rho_c} &= \frac{k}{a^2 H^2} \\ &= \frac{k}{\frac{8\pi G}{3} M_\omega a^{-(1+3\omega)} - k} \ , \end{aligned} \tag{41}$$

where the density parameter is defined as $\Omega(t) \equiv \rho(t)/\rho_c(t)$ and the critical density as $\rho_c(t) \equiv 3H^2(t)/8\pi G$. From (41) one can see that $k = 0$, $\Omega = 1$

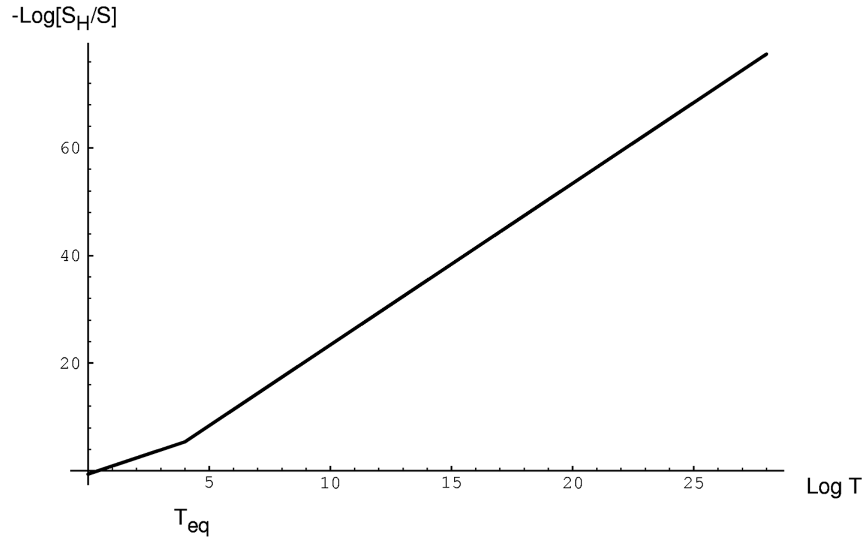

Fig. 3. The entropy per horizon is shown as the Universe cools. For the matter era the solution is given by (39) and for the radiation era by (40). The entropy per horizon presently is $S_{H_0} = S \sim 10^{88}$ at $T = 2.7\ ^\circ K$.

is an unstable point. Consider the limit $a \to 0$, then $\Omega \to 1$ for $\omega > -1/3$. Now, if $k = -1$, as $a \to \infty$ then $\Omega \to 0$; while for $k = +1$, as $a \to a_{\rm max.}$ then $\Omega \to \infty$. That is, unless $k = 0$ and exactly $\Omega = 1$, the spatially flat Universe is unstable [72]; see Fig. 4.

Let us analyze in greater detail the first limit taken above. In order to compare the presently observed $\Omega_0 = \mathcal{O}(1)$ with that in the past, we first consider the evolution during the matter dominated era, given by (9) with $\omega = 0$. It implies that

$$\Omega(t) - 1 = k \left(\frac{H_0^{-1}}{a_0}\right)^2 \left(\frac{t}{t_0}\right)^{2/3} \qquad \text{for} \qquad t_{\rm eq.} \leq t \leq t_0 \tag{42}$$

which at $t = t_0$ implies $\Omega - 1 \approx k$, but at $t = t_{\rm ls}$, $\Omega - 1 = k\ 10^{-4}$. Therefore, in order to explain the present $\Omega_0 = \mathcal{O}(1)$ one has to fine tune the density value at $t = t_{\rm ls}$ to be very similar to the critical value, the difference being of the order of only one part in ten thousand. This is the flatness problem.

For $t < t_{\rm eq.}$, we use the radiation solution, given by (9) with $\omega = 1/3$, to have

$$\Omega(t) - 1 = k \left(\frac{H_{\rm eq.}^{-1}}{a_{\rm eq.}}\right)^2 \frac{t}{t_{\rm eq.}} \qquad \text{for} \qquad t \leq t_{\rm eq.}\ , \tag{43}$$

at $t = t_{Pl}$, $\Omega - 1 = k\ 10^{-59}$! Thus, considering the entire evolution of the Universe beginning with Planckian initial conditions, one needs again to fine–tune the initial density value to be $\rho = (1 \pm 10^{-59})\rho_c$ in order to explain the

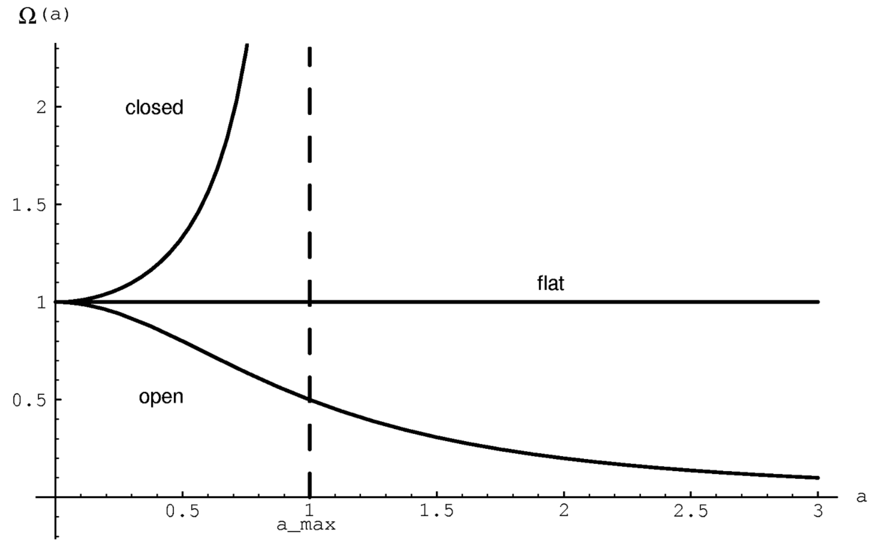

Fig. 4. The parameter Ω as a function of the scale factor, a, in a radiation dominated Universe. For closed models, with $k = +1$, Ω diverges as the scale factor approaches its maximum value, whereas for open models, $k = -1$, Ω asymptotically approaches to zero as the Universe expands. Finally, for a flat metric, $k = 0$, Ω is always equal to one. The behavior for a dust model is similar.

currently observed energy content of the Universe, i.e. to explain our own existence. The anthropic principle would just restate that the Universe has chosen those initial conditions necessary for us to be here! Nevertheless, this is no explanation but more a philosophical posture.

Let us try again to relate this issue to the aforementioned big numbers, (31) and (32). To do that, we express the above quantities in terms of the entropy within the horizon, (39) and (40). Since (24) is always valid, one obtains for both eras that

$$\Omega(t) - 1 = k \left(\frac{S_H}{S} \right)^{2/3} \qquad \text{for all times;} \tag{44}$$

again, at $t = t_0$, $\Omega(t) - 1 \approx k$. At $t = t_{\rm ls}$, $S_{H_{\rm ls}} = 10^{83}$ implies that $\Omega(t) - 1 \approx k10^{-4}$, whereas at the Planck time $S_H \approx 1$, one once again obtains that $\Omega(t) - 1 \approx k10^{-59}$. Very similar to the horizon problem, here one finds that the very small numbers come from the vast entropy increase within the horizon, which is the entropy necessary to fit the Universe's age, cf. (32).

Thus, the last two puzzles can be restated as: Why was the horizon entropy at the Planck time $S_{H_{Pl}} \approx 1$, but now $S_{H_0} \approx 10^{88}$?

Baryon Asymmetry

We observe that our Universe is apparently made of matter but not of antimatter. Why is this? Furthermore, the present different types of matter (fermions, bosons) are not in equal proportion. As we have already mentioned, nucleosynthesis restricts the value of η to be[5] $\eta \approx (3-4) \times 10^{-10}$, cf. (12); this fact tells us that the Universe is far more filled with photons than with baryons and if the baryon number is conserved η must also be conserved since the beginning of nucleosynthesis. In the standard model of cosmology one has to assume this as a given input. Let us explain this in more detail:

As far as observations show, within the solar system and our galaxy there is no evidence of primordial anti–baryons; if there were, some amount of gamma rays would be detected because of their annihilation with their baryon counterparts, something which has been not observed [89]. In going beyond galaxy scales, antimatter in galaxy clusters is ruled out by simple arguments that in fact are related to the horizon problem: one can imagine a baryon symmetric early Universe, whereby baryons and anti–baryons coexist in equilibrium. Their particle numbers in a co–moving volume should remain constant only until they become non–relativistic, when (25) begin to be valid; after that their particle abundances decrease exponentially. The particles remain in equilibrium annihilation until the temperature $T \approx 22$ MeV, when the annihilation rate, Γ, falls below the expansion rate. Then the ratio n_B/s is fixed to be $n_b/s = n_{\bar{b}}/s = 7 \times 10^{-20}$ [89, 60], nine orders of magnitude smaller than the currently observed ratio n_B/s! In order to avoid this annihilation catastrophe one can try in some manner to stop the annihilation mechanism some time before, at about $T \approx 38$ MeV when $n_b/s = n_{\bar{b}}/s = 10^{-10} - 10^{-11}$, by separating baryons from anti–baryons. Even so the horizon at that time contained the following amount of matter:

$$M_d = \rho d_H^3 = (\rho a^3)_0 \left(\frac{d_H}{a_0}\right)^3 = M_0 \times 10^{-6} \left(\frac{T_{\rm eq.}}{T}\right)^3 \quad , \tag{45}$$

where we have used (37); at $T = 38$ MeV, $M_d = 5 \times 10^{26}{\rm gr} = 2.7 \times 10^{-7} M_\odot$, which is clearly very much smaller compared to galaxy cluster mass scale; again, very fine tuning must be done. Instead of appealing to rare initial conditions, an alternative is to explain the baryon asymmetry by means of the Universe's evolution. Accordingly, at some high temperature $T \gtrsim 1$ GeV,

---

[5] In fact, η cannot be directly determined, nor can n_B/s. They are fitted to the currently observed values of light element composition in the Universe, i.e. $0.22 \lesssim Y_{^4{\rm He}} \lesssim 0.26$, ${\rm D/H} \gtrsim (1-2) \times 10^{-5}$, $({\rm D} + {}^3{\rm He})/{\rm H} \lesssim 10^{-4}$ and $({}^7{\rm Li/H}) \lesssim 2 \times 10^{-10}$. In this way, one can relate η with the baryon content of the Universe. Accordingly, for baryons $n_B = \rho_B/m_B = 1.13 \times 10^{-5} \Omega_B h^2/cm^3$ and $n_\gamma = \frac{2\zeta(3)}{\pi^2} T^3 \approx \frac{422}{cm^3} T_{2.75}^3$, therefore, $\Omega_B = 3.6 \times 10^7 \eta T_{2.75}^3/h^2$, and from the above mentioned value of η one finds that $0.01 \leq \Omega_B \leq 0.10$; that is, the Universe cannot be closed by baryons alone.

when $n_q \approx n_{\bar{q}} \approx n_\gamma$, a tiny asymmetry was already present and it prevented the total annihilation of quarks (q) and anti–quarks ($\bar{q}$), in such a manner that $\frac{n_q - n_{\bar{q}}}{n_q} = 3 \times 10^{-8}$ in order to explain the n_B/s needed for successful nucleosynthesis to take place [60].

The particle physics implementation in the early Universe by which that tiny asymmetry could be solved is called *baryogenesis*. The first attempt to address this problem was made by Sakharov [84], who pointed out three ingredients necessary to attain a baryon asymmetry in the Big Bang model. Let us review these: (*i*) baryon number violation, otherwise the baryon asymmetry can only reflect asymmetric initial conditions; (*ii*) violation of charge conjugation (C) and charge conjugation combined with parity (CP) are necessary to achieve different production rates for baryon and anti–baryons, otherwise a net zero baryon number is maintained; and (*iii*) non–equilibrium conditions, otherwise the same Fermi distribution of baryons and anti–baryons would guarantee the same phase space for them, i.e. $n_b = n_{\bar{b}}$.

It is curious that these three conditions were pointed out before there was a theory which could accomplish them. Indeed, first GUT appeared in the 70s and when one realized them in an early Universe scenario, they met the three ingredients: the first two are fulfilled because, by construction, strong and electroweak interactions are unified; this implies that quarks and leptons are members of a common irreducible representation of the GUT gauge group. In that way, gauge bosons mediate interactions in which baryons can decay into leptons, or the inverse, giving rise to a baryon number violation. C is violated by weak interactions and CP violation is observed in Kaon K^0 (meson) interaction. Thus, one also expects that the massive X-bosons decay into quarks/leptons, with a branching ratio of, say, r, and $\bar{X}$ with $\bar{\text{r}}$, such that $\text{r} \neq \bar{\text{r}}$. The third condition is attained due to Universe expansion, which evolves as $H \sim T^2/M_{Pl}$ in the radiation dominated era. For that to happen, one takes the reaction rates (decay, annihilation, and inverse processes) $\Gamma_X > H$. Then, through the *out-of-equilibrium decay* mechanism the $X-$bosons have a long enough lifetime so that their inverse decays go out of equilibrium as they are still abundant. In this way, the baryon number is produced by the X free decay, whereas the inverse rates are turned off.

Nevertheless, GUT have their own problems. For instance, precisely because of the first two ingredients above, the proton should decay too; in the minimal SU(5) GUT its lifetime is $\tau_p \approx 10^{29\pm1}$ years, but the experimental limit is greater, $\tau_p \approx 10^{31-32}$ years. Thus, something is wrong with this theory.

Another problem of GUT is that unless the model is B-L conserving, any net baryon number generated might be brought to zero by efficient anomalous electroweak processes, at temperatures of about $T \sim 100$ GeV. Though this seems not possible within the standard model electroweak baryogenesis, there are model extensions where this can be possible; see contribution of Piccinelly and Ayala in this book. This possibility represents a serious problem to GUT, and it opens new windows for low energy physics. Here, we depict briefly the idea of how electroweak baryogenesis works[31]: the vacuum manifold of the electroweak model, the so-called $\theta-$vacuum, has degenerated minima separated by energy barriers in the field configuration space as a result of non–trivial vacuum gauge configurations ($A^a{}_\mu \neq 0$) of non–Abelian gauge theories. Different minima have different baryon and lepton numbers, with the net difference between two adjacent minima being given by the number of families. Thus, for the standard model, jumps between these minima imply the creation of three baryons and leptons, hence, there is B-L conservation and B+L violation. At $T = 0$, tunneling (jumps) between two adjacent minima is mediated by *instantons* and the tunneling rate is exponentially suppressed [92] $\Gamma \sim e^{-\frac{1}{\alpha_{EW}}}$, where $\alpha_{EW} = 1/170$ is the electroweak coupling constant; this is why the proton is stable. However, at finite temperature, $T \approx 100$ GeV, one can go over the energy barrier to achieve a baryon number violation, as first described in [61]. The height of the barrier is a solution of an unstable static configuration called *sphaleron*, whose rate is $\Gamma \sim e^{-\frac{E_S}{T}}$, with its associated energy $E_S \approx M_W/\alpha_{EW}$. For temperatures above the critical temperature of electroweak symmetry restoration, the rate is no more strongly suppressed, but $\Gamma \sim (\alpha_{EW} T)^4$, indeed making possible baryon number violation. The other two ingredients to achieve baryogenesis could also be present, but a detailed analysis is in order; for a short review see [32, 44] and for an extended one see [31].

Monopole and Other Relics

Another problem of GUT is the production of magnetic monopoles [91, 77] as a consequence of GUT symmetry–breaking to some semi simple group U(1). In the course of the phase transition, bubbles of the new phase are produced and on scales greater than d_H one expects different Higgs field alignments. Because of this randomness, topological knots are present and they are the magnetic *monopoles*. It has been proved that their number density should be comparable to the baryon density, but their mass is 10^{16} times greater than that of the protons; in this case, the Universe should have recollapsed long before [55, 102, 78].

Additionally, some theories predict primordial cosmological particles (or structures) that could be present currently, also as a result of some spontaneous symmetry–breaking process. Among these cosmological relics are massive neutrinos, gravitinos, domain walls, cosmic strings, axions, etc.

Cosmological Constant

Another problem that arises as a consequence of theories of grand unification (or theories of everything, including gravity) is that the vacuum energy associated with these, $< 0|T_{\mu\nu}|0 >=< \rho > g_{\mu\nu}$, turns out to be very large.

Summing the zero-point energies of all normal modes of some field of mass m, one obtains $<\rho>\approx M^4/(16\pi^2)$, where M represents some cutoff in the integration, $M \gg m$. Then, assuming GR is valid up to the Planck scale, one should take $M \approx 1/\sqrt{8\pi G}$, which gives $<\rho>= 10^{71}$ GeV^4. This term plays the role of an effective cosmological constant of $\Lambda = 8\pi G <\rho>\approx M_{Pl}^2 \sim 10^{38}$ GeV^2 which must be added to the l.h.s. of Einstein equations (2) and yields an inflationary solution (10). However, if the cosmological constant is at present of the order of magnitude of the material content of the Universe, one has that

$$\Lambda \sim 8\pi G\rho_0 = 3H_0^2 \sim 10^{-83}\mathrm{GeV}^2, \tag{46}$$

which is very small compared with the value derived above on dimensional grounds. Thus, the cosmological constraint and theoretical expectations are rather dissimilar, by about 121 orders of magnitude! Even if one considers symmetries at lower energy scales, the theoretical Λ is indeed smaller, but never as small as the cosmological constraint. One finds that $\Lambda_{GUT} \sim 10^{21}$ GeV^2 and $\Lambda_{SU(2)} \sim 10^{-29}$ GeV^2 in contrast to (46). For an analysis of this problem in terms of longitude scales (not of mass square scales), see the contribution by E. Copeland in this book. This problem has been reviewed in [98, 19].

Large–Scale Structure

The problem of explaining structure formation in the Universe is most fascinating. There exist stars, galaxies, clusters of galaxies, superclusters, voids, and a variety of large–scale structures in the currently observed Universe, whose origin one hopes to understand within the framework of Newtonian or GR physics. Such systems represent complicated problems, for which one needs a deep understanding of both the initial conditions of the relevant physical quantities and their evolution: among them are the Universe composition (accounted in the density of the different i-species, Ω_i) and the type of perturbation the Universe experienced, i.e. adiabatic or isocurvature (isothermal).

Imagine an early Universe filled with a radiation fluid (i.e. effective relativistic) and some non–relativistic components. Let us consider the following density contrast:

$$\delta(\boldsymbol{x}) \equiv \frac{\delta\rho(\boldsymbol{x})}{\bar{\rho}} = \frac{\rho(\boldsymbol{x}) - \bar{\rho}}{\bar{\rho}} \ , \tag{47}$$

where $\bar{\rho}$ is the average density of the Universe. If in the past there were small density perturbations that grew as time went on, the formation of some structure will be favored. This density contrast is commonly expanded into a Fourier expansion:

$$\delta(\boldsymbol{x}) = \frac{V}{(2\pi)^3}\int_{\rm vol.} \delta_k e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} d^3k \qquad \text{and}$$
$$\delta_k = \frac{1}{V}\int_{\rm vol.} \delta(\boldsymbol{x}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}} d^3x \ , \tag{48}$$

where for $\boldsymbol{x}$ are chosen co–moving coordinates. A given "perturbation" mode λ is associated with its wave number $k = 2\pi/\lambda$. The physical mode is given by $\lambda_{\rm phys.} = \lambda\frac{a(t)}{a_*}$, then $k_{\rm phys.} = k\frac{a_*}{a(t)}$, when the expansion begins at $t = t_*,\ a(t_*)$. One can also relate to a given λ a mass defined as the rest mass contained in a radius $\lambda_{\rm phys.}/2$, $M \equiv \frac{\pi}{6}\rho_m\lambda^3_{\rm phys.} = 1.5\times10^{11}M_\odot(\Omega_0h^2)\lambda^3_{\rm Mpc}$. For a galaxy, for instance this corresponds to $\lambda_{\rm gal.} \approx \frac{1.9\rm Mpc}{(\Omega_0h^2)^{1/3}}$; this is the physical scale that would contain today a galaxy mass (of approximately $\sim 10^{12}M_\odot$) and after its non–linear regime would give rise to a typical galaxy size of approximately 30 kpc.

The fundamental quantity $|\delta_k|^2$, called the power spectrum, $\mathcal{P}(k)$, determines any statistical quantity for gaussian random fluctuations. In the absence of a fundamental theory of structure origin, one admits a power spectrum of the type

$$\mathcal{P}(k) \equiv |\delta_k|^2 = \text{const. } k^{n_s} \ , \tag{49}$$

where an isotropic wave number $|\boldsymbol{k}| = k$ has been assumed which is allowed in an FRW Universe symmetry; n_s is a constant called the *spectral index*[6]. At first, the Cosmic Background Explorer (COBE) satellite DMR results[86] suggested that $n_s \sim 1$ [11, 12, 100, 7]. Recently, the WMAP satellite measurements of the CMBR concluded more precisely that $n_s = 0.93 \pm 0.03$ [13].

The evolution of the density contrast determines whether and when the perturbation grows to arrive at its non–linear stage, when $|\delta_k|^2 > 1$, it starts to develop structure formation. This comes out by analyzing the Jeans equations in an expanding Universe. One finds that effectively there exists growing modes solutions (for $k < k_J$), which open, in principle, the possibility of describing the presently observed large–scale structure.

A particular perturbation is given through a Fourier component and is characterized by its amplitude and its co–moving wave number, in terms of which one can write the root-mean-square (*rms*) density fluctuation $\delta\rho/\rho$ as,

$$\left(\frac{\delta\rho}{\rho}\right)^2 \equiv \langle\delta(\boldsymbol{x})\delta(\boldsymbol{x})\rangle = \frac{1}{V}\int_0^\infty \frac{k^3|\delta_k|^2}{2\pi^2}\frac{dk}{k} \ , \tag{50}$$

where $\langle\cdots\rangle$ stands for spatial average. An *rms* mass fluctuation, δM, corresponds to a density contrast such that

[6] The spectral index, n_s, is sometimes referred as n, cf. contribution of E. Copeland in this book.

$$\left(\frac{\delta M}{M}\right)^2_\lambda = \frac{1}{(2\pi)^3}\frac{1}{VV_W^2}\int \mathcal{P}(k)|W(k)|^2 d^3k \ , \tag{51}$$

where $W(k)$ is a window function, typically Gaussian, i.e. $W(k) = V_W e^{-k^2 r_o^2/2}$, where r_o is the radius within which the mass M is contained; $V_W = (2\pi)^{3/2} r_o^3$ being its volume. Then, one has for the *rms* mass perturbation at a given λ that

$$\left(\frac{\delta M}{M}\right)^2_\lambda \sim k^3 \mathcal{P}(k) \sim k^{3+n_s} \ , \tag{52}$$

where the overall normalization amplitude, for all λ, has not yet been specified, according to (49). Note that $\left(\frac{\delta\rho}{\rho}\right)_k \approx \left(\frac{\delta M}{M}\right)_\lambda$, where the subindex k in the density contrast means $\frac{\delta\rho}{\rho}$ in a given logarithmic interval $\frac{dk}{k} \sim 1$.

Because of the existence of a causal horizon, there should be some $\lambda-$ modes that were once super-horizon sized and that some time later enter inside the horizon. These modes begin to grow posteriorly at $t > t_{\rm eq.}$. Once the perturbation enters the horizon, the Universe is well described with Newtonian physics and the distinction between adiabatic and isothermal perturbations becomes irrelevant; see Fig. 5.

The density and matter contrasts evolve for superhorizon modes as $\frac{\delta M}{M} \sim a^m(t) \sim t^{\frac{2m}{3(1+\omega)}}$, see (9); during the radiation era $\omega = 1/3$, $m = 2$ and for the matter dominated era $\omega = 0$, $m = 1$. During the time the physical mode is superhorizon sized, it scales as $\lambda_{\rm phys.} = \lambda\, a(t) \sim \lambda\, t^{\frac{2}{3(1+\omega)}}$ and at the moment this mode enters the horizon it is valid that $\lambda_{\rm phys.} = d_H \sim t_H$, therefore, $\lambda^{\frac{3(1+\omega)}{1+3\omega}} \sim k^{-\frac{3(1+\omega)}{1+3\omega}} \sim t_H$; since $\frac{3(1+\omega)}{1+3\omega} > 0$, then the larger the initial perturbation wavelength is, the later it enters the horizon. This means larger perturbation wavelengths begin later to develop to their non–linear regime, thus, one expects large–scale perturbations in the present to be smaller than the small scale perturbations. This is in accordance with the fact that $\delta\rho/\rho \sim 10^{30}$ for stars, 10^5 for galaxies, 10^{1-3} for cluster of galaxies, and $\mathcal{O}(1)$ for superclusters.

One has at the moment of horizon entering for every $\lambda-$scale that

$$\delta_{\rm H}(k) \equiv \frac{\delta M}{M}(k, t_H) \sim t_H^{\frac{2m}{3(1+\omega)}} \cdot \frac{\delta M}{M}(k,t) = k^{-\frac{2m}{1+3\omega}} \cdot k^{\frac{3+n_s}{2}} = k^{\frac{n_s-1}{2}} \ . \tag{53}$$

This is valid for both radiation and matter dominated horizon entering modes, since $\frac{2m}{1+3\omega} = 2$ for both cases. Harrison [49] and Zel'dovich [101] have argued that at the time perturbations enter the horizon they should have equal amplitude, that is, a *scale invariant spectrum*, which is achieved by choosing $n_s = 1$. This value is preferred by observations, as mentioned above. For instance, if $n_s > 1$, the perturbations are too strong and tend to close up and form island Universes; for $n_s < 1$ they are too weak to form galaxies [72]. Furthermore, the amplitude is required to be $\mathcal{O}(10^{-5})$ in magnitude when the perturbations start to grow at $t_{eq.}$.

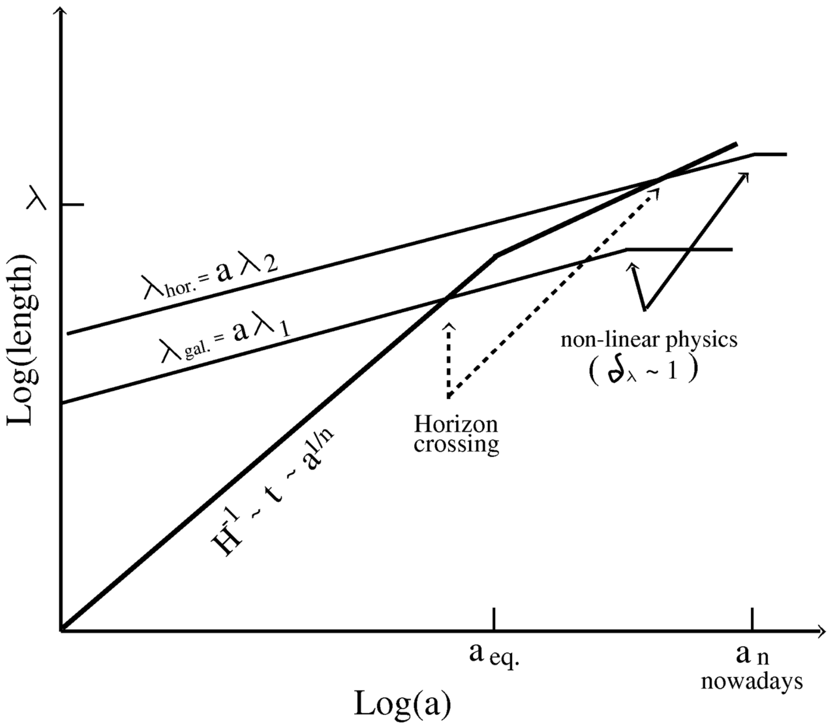

Fig. 5. At some early times there were superhorizon sized perturbations characterized by their wavelength and amplitude. As the Universe evolves these perturbations grow with the scale factor and eventually they cross inside the horizon, H^{-1}. After that, these modes first have a linear regime ($\delta_\lambda < 1$) and afterwards they start to develop the sky structure we see nowadays. In the figure we show two typical modes, one corresponding to galaxy scales and the other to horizon scales. The horizon evolves as $H^{-1} \sim t \sim a^{1/n}$, where here n denotes $n = 3(1+\omega)/2$, from solution (9). (Figure adapted from Kolb and Turner 1990).

During its radiation dominated phase, the Universe will not significantly develop non–linearities inside the horizon. Furthermore, some $\lambda-$scales are forbidden because of damping phenomena due to collisionless phase mixing of relativistic particles, known as Landau damping or *free–streaming* (fs). This is caused because relativistic particles move freely away from overdense regions and their velocity dispersion dissolves the compression regions. In this way, only those wavelengths greater than some free streaming scale [7] (λ_{fs}) are

[7] That is, wavelengths corresponding to scales greater than the horizon size when the particles become non–relativistic. $\lambda_{fs} = \mathcal{O}(1) \times \left(\frac{t}{a}\right)_{nr} \approx 1\mathrm{Mpc}\frac{\mathrm{keV}}{m_X}\frac{\mathrm{T_X}}{T_\gamma}$, where nr refers to when the $X-$component, with mass m_X and temperature T_X, becomes non–relativistic.

allowed to maintain overdense regions. Once a non–relativistic component becomes dominant, the growing modes increase typically as $\delta \sim a(t)$. However, baryons suffer from collisional damping due to photon diffusion, mainly during photon decoupling. This allows growing $\lambda-$modes that are larger than the diffusion scale, also known as the *Silk scale*[8] (λ_S).

The amplitude of density perturbations affects the temperature profile and, at the last scattering surface, the last stage of matter-radiation equilibrium. This can be quantified, since the present–day observed temperature contrast, $\frac{\delta T}{T} \sim 10^{-5}$, represents a "cosmic imprint" of the Universe since decoupling.

The currently measured temperature perturbations are partially caused by fluctuations in the gravitational potential, the so–called the *Sachs-Wolf effect*. This effect is responsible for large angular–scale ($\theta > 1°$) anisotropies[9]. The other known effect, the dipole anisotropy, is presumably due to our galaxy's peculiar velocity with respect to the cosmic rest frame.

One finds in the synchronous gauge for a flat Universe and large angular scales that

$$\begin{aligned} \left.\frac{\delta T(\boldsymbol{x})}{T}\right|_{\rm ls} &= -\frac{a_0^2 H_0^2}{2(2\pi)^3} \int k^{-2} \delta_k e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} d^3 k \\ &= \frac{1}{3} \Delta\phi(\boldsymbol{x}, t_0) \approx \left(\frac{\delta\rho}{\rho}\right)_{\lambda \sim H_0^{-1}} , \end{aligned} \tag{54}$$

where $\Delta\phi(\boldsymbol{x}, t_0)$ is the perturbed Newtonian potential and $\boldsymbol{x}$ points to the last scattering surface and has a length of $2H_0^{-1}$. It is convenient to expand $\delta T/T$ in spherical harmonics,

$$\frac{\delta T(\boldsymbol{x})}{T} = \sum_{l=2}^{\infty} \sum_{m=-l}^{m=+l} \mathbf{a}_{lm} Y_{lm}(\theta, \phi) \ , \tag{55}$$

where θ and ϕ are the spherical angles in the sky ($\theta > 1°$ corresponds to $l < 100$). The coefficients $\mathbf{a}_{lm}$ can be computed for the power spectrum of (49) resulting in

$$\begin{aligned} C_l \equiv \langle |\mathbf{a}_{lm}|^2 \rangle &= \frac{H_0^4}{2\pi V} \int_0^\infty |\delta_k|^2 [j_l(kx)]^2 \frac{dk}{k^2} \\ &\sim H_0^{n_s+3} \approx \left(\frac{\delta\rho}{\rho}\right)^2_{H_0^{-1}} \end{aligned} \tag{56}$$

---

[8] $\lambda_S \approx \left(\frac{\Omega_0}{\Omega_B}\right)^{1/2} \frac{3.5\text{Mpc}}{(\Omega_0 h^2)^{3/4}}$, where Ω_B is the baryonic contribution to the total density, i.e. $\Omega_0 = \Omega_B + \Omega_\gamma + \cdots$.

[9] 1° because the horizon scale at decoupling subtended approximately that angle. Then, $\theta > 1°$ corresponds to a $\lambda-$perturbation that was super-horizon sized. The existence of such super-horizon scales represents an initial density spectrum beyond causality!

where j_l is the spherical Bessel function of order l. Hence, $\delta T/T$ depends upon the present horizon scale, H_0^{-1}. For a Harrison-Zel'dovich form, $n_s = 1$, the spectrum of density fluctuations with a $\lambda-$scale greater than $\lambda_{\rm eq.} \approx 13(\Omega_0 h^2)^{-1}$ Mpc should conserve its initial form, and is related to an *rms* mass perturbation as follows:

$$C_l \approx (H_0 L)^4 \left(\frac{\delta M}{M}\right)^2_L . \tag{57}$$

Evaluating it in $L = 30h^{-1}$ Mpc and $(\delta M/M)_L = 1/4$, as inferred from measurements, one gets $C_l^{1/2} \approx 2 \times 10^{-5}$ which is experimentally confirmed from the CMBR anisotropies measurements; a more general discussion is found, e.g., in [60].

Note that, on the one hand, one has that $\left(\frac{\delta\rho}{\rho}\right)_{\rm ls} = {\rm const.} \left(\frac{\delta T}{T}\right)_{\rm ls}$ and, on the other hand[10], $\delta\rho/\rho \sim a(t) \sim 1/(1+z)$; thus, the maximal amplitude grow from last scattering until now is given by $1 + z_{\rm ls} \equiv \frac{a_0}{a_{\rm ls}} = 1200$ times 10^{-5}, which is not sufficient to form the non–linear structures we observe today. Therefore, other (dark) components must have played a role in the growing of perturbations. It seems that considering dark matter together with a quintessence field the correct grow can be achieved, as is suggested in the forthcoming lectures in this book.

One more intriguing aspect of the large–scale structure arises by observing the autocorrelation function of galaxies, clusters, etc.; $\xi(r) \equiv \langle \delta n(\boldsymbol{x}+\boldsymbol{r})\delta(\boldsymbol{x})\rangle$, which is proportional to the probability of finding an emitting object at a distance r from a given object, i.e. $\delta P = n\delta V[1 + \xi(r)]$; for instance, in a totally uniform distribution $\xi(r) = 0$, whereas $\xi(r) > 0$ indicates an enhancement of density near a given object. This has been measured for galaxies, clusters, and superclusters, showing that approximately [5, 9, 74]

$$\begin{aligned} \xi_g &\approx 20\, r^{-\gamma} && \text{for} \quad 2 \lesssim r_{\rm Mpc} \lesssim 10 \ , \\ \xi_c &\approx 360\, r^{-\gamma} && \text{for} \quad 15 \lesssim r_{\rm Mpc} \lesssim 100 \ , \\ \xi_{sc} &\approx 1500\, r^{-\gamma} && \text{for} \quad 100 \lesssim r_{\rm Mpc} \lesssim 200 \ , \end{aligned} \tag{58}$$

where $\gamma \approx 1.7 - 1.8$ and the distance $r_{\rm Mpc}$ is given in Mpc. Again, due to the existence of a causal horizon, if one goes back in time one finds out that the co–moving separation between two emitting objects, e.g. cluster of galaxies, is larger than the light cone of causality at $t = t_{\rm eq}$. Therefore, if the initial density perturbations, $\delta(\boldsymbol{x})$, were produced before, or at, $t_{\rm eq}$, the correlations cannot be explained by causal mechanisms. Hence, why do have approximately the same slope, but different magnitudes or correlation lengths, is a mysterious aspect that should be explained by the structure formation theory,

[10] By considering the evolution of the Universe, one often uses, instead of the time parameter, the redshift (z), defined as $1 + z \equiv a_0/a$.

still in development. The different magnitudes of $\xi(r)$s suggest the existence of some dark matter present in greater quantities on bigger scales. This is in agreement with the determination of Ω_0 from dynamical computations, according to which one needs more dark matter as one goes from galaxy to cluster scales, and so forth.

In order to explain these issues, investigators have developed some numerical codes in which they include Hot Dark Matter (HDM) like light neutrinos in order to achieve a *top–down* scenario, first performing large structures that should fragment to give rise to smaller ones. One finds that typical large–scale filamentary structures and voids are well reproduced, but smaller scales are underweighted and, for instance, galaxy formation should have taken place rather recently, at $z \lesssim 1$, in order to obtain the observed galaxy correlation function mentioned above. Besides, the predicted limits for $\Delta T/T$ are near the upper limit measured. Numerical simulations also include Cold Dark Matter (CDM) with WIMPs (weakly interacting massive particles) to have now a *bottom–up* scenario, where first the smaller structures are formed and, later on, the larger ones. These simulations are in good agreement with galaxy correlation functions for acceptable redshifts, say $z \sim$ few. Nevertheless, some negative correlation function is expected for galaxies that have not been observed. Furthermore, the cluster correlation function is predicted to be about three times less than the measured value, that is, large scales are underweighted; the temperature contrast in this model can be as much as one order of magnitude below that observed. One notes that HDM and CDM play opposite roles for structure formation, because of their different free streaming ranges. Therefore, one considers mixed models (MDM) which include both CDM and HDM, and additionally some smooth components as a cosmological constant term (ΛDM). Up until now, it is not known for sure which particles have participated as the relevant building blocks that eventually brought about the formation of the large–scale structure with the right spectrum of anisotropies we observe today; computations hint there may be mixture of them: cold, hot, and a $\Lambda-$term; see the contribution of Cabral–Rosetti et al in this book. A review of these topics can be found in [4].

Summarizing, the problem of density perturbations lies in the understanding of its growth during its linear era in such away that on large scales $\delta\rho/\rho \sim \delta T/T \sim 10^{-5}$, but on small scales $\delta\rho/\rho > 1$ in order to reproduce the structures we see in the sky nowadays.

We have mentioned above some important aspects to be considered in order to achieve a better understanding of our present Universe. One should mention that these problems do not imply logical inconsistencies with the SBB. Nevertheless, for their explanation one is forced to appeal to very special initial conditions, a thing that a physicist would hardly accept. Moreover, solutions like the anthropic principle result to be dissatisfying. Therefore, extensions of the SBB are required, because by nature some of the presented problems, but not all, already come from its extensions, as by incorporating high energy ($\gg$ MeV) physics; ergo, the solutions to the puzzles should hopefully come from correct implementations of such extensions. Some proposed scenarios are already known, such as the inflationary one, which is the best candidate and is the subject of the next section. In order to achieve a further understanding of the early Universe, it is tempting to move from energies in MeVs to the Planck energy scale, that is, 22 orders of magnitude greater!

2 Beyond the Standard Big Bang Model: Inflation

In this section we show a way to solve the problems of the SBB presented in Sect. 1. We explain the general inflationary scenario that took place in the very early Universe and that gave rise to a successful cosmological model; that is, a Universe with its right causal size, age, temperature, and the perturbations spectrum that originated structure formation. At the end of this section we point out some remarks on inflationary models.

Inflation was accomplished as a natural extension of the "new" physics of the 70s being incorporated into the SBB. With the advent of GUT it was natural to study their cosmological consequences. In the late 70s and the beginning of the 80s some publications appeared about effects of GUT phase transitions in the very early Universe; even in this respect some authors considered exponential solutions, see [65, 72]. But the cornerstone paper was that of A. Guth [46], where he stressed that due to these phase transitions, the Universe could have experienced an exponential expansion of approximately e^{65} foldings, and in this way one could solve the horizon, flatness, and monopole problems, all at once. This was the first model of inflation. Although the original model suffered from some problems, this has shown that it is, in principle, possible to tackle the problems of the SBB by considering some vacuum energy or scalar fields to be present at the very beginning of time. Next, we explain how inflation addresses this.

2.1 Inflation: The General Idea

As we mentioned in Sect. 1, the FRW cosmological (4)-(6) admit very rapid expanding solutions for the scale factor. This is achieved when the inflation pressure, $\rho + 3p$, is negative, i.e. when the equation of state admits negative pressure such that $\omega < -1/3$, to have $\ddot{a} > 0$. For instance, if $\omega = -2/3$, one has that $a \sim t^2$ and $\rho \sim 1/a$, that is, the source of rapid expansion decreases inversely proportional to the expansion. Of special interest is the case when $\omega = -1,\ \rho = \text{const.}$, because this guarantees that the expansion rate will not diminish. Thus, if $\rho = \text{const.}$ is valid for a period of time, τ, then the Universe will experience an expansion of $N = \tau H$ $e-$foldings, given by $a = a_* e^N$, (10).

This is the well known de Sitter cosmological solution[11] [27], achieved here only for a $\tau-$stage in an FRW model.[12]

We shall now see how an inflationary stage helps to solve the problems of the SBB. First consider the particle (causal) horizon, given by (33), during inflation; with $k=0$, again one obtains

$$d_H = H^{-1}(e^{Ht} - 1) \ ; \tag{59}$$

the causal horizon grows exponentially, whereas H^{-1} remains constant. Since $d_H \neq H^{-1}$, we call H^{-1} the *Hubble horizon* to distinguish it from the causal horizon. We again compare, in analogy to (34) and (36), the horizon distance with that of any physical length scale, $L(t) = L_* \frac{a(t)}{a_*} = L_* e^{Ht}$, to get

$$\frac{d_H}{L} = \frac{H^{-1}(e^{Ht} - 1)}{L_* e^{Ht}} \gtrsim 1 - e^{-Ht} \ , \tag{60}$$

for initial length scales $L_* \lesssim H^{-1}$. After a few e-fold times the causal horizon is as big as any length scale which initially was subhorizon sized. Therefore, if the original patch before inflation is causally connected, and presumably in equilibrium[13], after inflation this region of causality is exponentially bigger than it was; thus all the present observed (apparent) Universe can stem from it. Therefore, at some later time, say, at the time of last scattering (photon decoupling) the Universe has all the mentioned 10^5 regions (and more than that!) causally connected, then solving the horizon problem. In fact, if the inflation stage is sufficiently long, there can presently exist regions which are still so distant from each other that they are no longer in contact, even though originally they come from the same causal patch in existence before inflation; they will be in contact again when light reaches these distant points.

From (59) one can observe that if the initial physical length scale is greater than the Hubble distance, $L_* > H_*^{-1}$, then $L > d_H$ during inflation. Events initially outside the Hubble horizon remain acausal. This is better observed by considering the *event horizon*, d_e, that determines the region of space

[11] The de Sitter model contained no matter, $p = \rho = 0$. Instead de Sitter considered a cosmological constant such $H^2 = \Lambda/3$. In this sense, this was an anti–Machian solution, since matter was not needed to produce inertia. Alternatively, by choosing $p = -\rho$, $\Lambda = 0$ the dynamic is the same and the solution is Machian, but the price paid is that of having such exotic matter, $p = -\rho$.

[12] From now on, *inflation* will refer to the period of exponential expansion. However, an *inflationary scenario* also implies some other physical processes such generation of density fluctuations, reheating, etc.

[13] One can imagine the initial stage of the Universe to possess some inhomogeneities, anisotropies, and a rather chaotic distribution of particles. At some time after the Planck scale, $10^3 t_{Pl}$, one expects the dampening of the anisotropies and inhomogeneities in the metric, and due to statistical processes, the Universe should thermalize in some local scale ($< d_H$), which we now call the *original patch* [46].

which will remain in causal contact after some time; that is, it delimits the region from which one can ever receive (up to some time $t_{\rm max}$) information about events taking place now (at the time t):

$$d_e(t) = a(t) \int_t^{t_{\rm max}} \frac{dt'}{a(t')} \ . \tag{61}$$

For a flat model during its matter dominated era ($a \sim t^{2/3}$), like our Universe in the present, $d_e \to \infty$ as $t_{\rm max} \to \infty$. However, during inflation one finds that

$$d_e = H^{-1}(1 - e^{-(t_{\rm max}-t)H}) \approx H^{-1}, \tag{62}$$

which implies that any observer sees only those events that take place within a distance $\leq H^{-1}$. In this respect, there is an analogy with black holes, from whose surface no information can escape. Here, in an exponential expanding Universe, observers encounter themselves in a region which apparently was surrounded by a black hole [43, 65], since they receive no information located further than H^{-1}.

So far we have seen how the exponential solution can offer means to understand the horizon problem of the SBB. Now, we analyze it quantitatively. First, consider the evolution of a co–moving length since the very beginning. Before inflation starts the Universe is characterized by some initial values of temperature, Hubble horizon, etc., related by

$$\begin{aligned} \rho_* &= b\, T_*^4 \ , \\ H_*^2 &= \frac{8\pi G}{3}\, b\, T_*^4 \ , \\ S_* &= \frac{4}{3}\, b\, (a_* T_*)^3 \ . \end{aligned} \tag{63}$$

If inflation does not take place, an initial horizon–sized length scale, $L_* \leq H_*^{-1}$, grows only proportionally to $t^{1/2}$ (radiation era) and, later, to $t^{2/3}$ (matter era). Therefore its size would currently be much smaller than our apparent Universe; in others words, this is the horizon problem. Hence, suppose that inflation indeed takes place during a time period $\tau = NH_*^{-1}$ and, afterwards, a usual Friedmann Universe follows. Then, a typical co–moving scale evolves as

$$L = L_* e^N \left(\frac{t_{\rm eq}}{t_f}\right)^{1/2} \left(\frac{t}{t_{\rm eq}}\right)^{2/3} \ . \tag{64}$$

The first term accounts for the inflationary stage (until the final time $t_f \approx \tau = NH_*^{-1}$), the second for the radiation era that lasts until the time of equal densities ($\rho_r = \rho_m$ at $t = t_{\rm eq}$), and the third term for the matter dominated era; all of them being solutions of the FRW cosmological equations characterized by different equations of state, cf. (9) and (10). Consequently, the number of e-folds (N) of growth in the initial co–moving scale L_*, to achieve at present (t_0) a size L_0, is given by

$$N = \frac{1}{\log\, e}\left(\log\frac{L_0}{L_*} - \frac{1}{2}\log\frac{t_{\rm eq}}{t_f} - \frac{2}{3}\log\frac{t_0}{t_{\rm eq}}\right) \equiv N_{\rm min1}\ . \qquad (65)$$

In principle, one could substitute the value $t_f \approx NH_*^{-1}$ into (65) and try to solve it for N, but one finds no analytical solution. As a matter of fact, one typically obtains $t_f \approx 10^2 H_*^{-1}$ and this turns out to be a very good approximation. For definiteness, let us assume this and, furthermore, that the initial co–moving scale is the horizon at the beginning, i.e. our original patch; then from (63), $L_* = H_*^{-1} \approx 10^{-1} M_{Pl}/T_*^2$. T_* is the temperature that characterizes some phase transition and is typically $T_* = 10^{14}$ GeV. Then, $H_*^{-1} = 10^{-24}$ cm, $t_* = 10^{-35}$ s, $\frac{t_{\rm eq}}{t_f} = \frac{10^{12}\rm s}{10^{-33}\rm s} = 10^{45}$, $\frac{t_0}{t_{\rm eq}} = \frac{10^{17}\rm s}{10^{12}\rm s} = 10^5$, and $\frac{L_0}{L_*} = \frac{10^{28}\rm cm}{10^{-24}\rm cm} = 10^{52}$ to yield $N = 60.2$. In fact, this value represents the minimal number of e-folds of inflation necessary for an initial horizon sized co–moving length scale to grow as big as our presently observed Universe, $L_0 = H_0^{-1} \approx 10^{28}$ cm. If the original patch is horizon sized, then during inflation it remains within the causal horizon, d_H, according to (59). After inflation the causal horizon grows as $d_H \sim t$, whereas the co–moving scale expands only as the scale factor does, in conformity with (9). Therefore, the co–moving length scale, L, remains always within the causal horizon.

For the inflationary Universe the currently apparent horizon comes from a region delimited by the original patch H^{-1}, which during inflation remains almost constant and, after, evolves as $H^{-1} \sim t$. At the end of inflation $a(t) \gg H^{-1}(t)$. Subsequently, the scale factor expands only with the power law solution $t^{1/2}$ (or $t^{2/3}$), whereas the Hubble horizon evolves faster, $H^{-1} \sim t$. Then, at some later time the Hubble horizon is as large as the scale factor, $H^{-1} \sim a(t)$. Accordingly, the value $N_{\rm min1}$ defines the minimal number of e-folds of inflation necessarily to have this equality at present; that is, the original patch grown until now is as big as our apparent Hubble horizon. Hence, some time ago, say, at the last scattering surface (photon decoupling), the Universe consisted of 10^5 Hubble horizon regions, yet all these regions stem from one original patch of size H_*^{-1} just at the start of inflation. Recall that the causal horizon, d_H, is always bigger than H^{-1}, except at the outset of inflation. Naturally, our original patch could have experienced a longer period of inflation. In this case, $N > N_{\rm min1}$ and the Universe is bigger than we observe it to be today: our event horizon will enable us to explore the Universe beyond.

Let us see how the flatness comes out of inflation. Consider originally a Hubble patch, H^{-1}, that might even possess some curvature different from zero. If there were the above conditions for such a gigantic expansion to take place, then in a short time the original Hubble patch will become very flat, since $H = \sqrt{\frac{8\pi G}{3}\rho}$ stays constant during that τ-stage. On the other hand, a typical scale $L_* \leq H^{-1}$ will exponentially increase in size as $L(t) = L_*\frac{a(t)}{a_*} = L_* e^N$. That is, all physical inhomogeneities, anisotropies and/or 'perturbations' of any kind (including particles!) will be diluted away. Their density becomes insignificant, thus solving the monopole (and other relics) problem.

When we discussed the flatness problem it was pointed out that Ω closely approaches unity as one goes back in time (see 42 and 43), in a way such that one must choose very special initial density values for explaining our flatness today, i.e. $\Omega_0 \approx \mathcal{O}(1)$. Imagine the Universe with initial conditions taken above, $a_* \sim H_*^{-1}$, then $\Omega_* - 1 \approx k$. Now, if the exponential expansion occurs, $\Omega(t = \tau)$ evolves to

$$\Omega(\tau) - 1 = \frac{\rho - \rho_c}{\rho_c} = \frac{k}{a^2 H^2} = k e^{-2N} \quad . \tag{66}$$

If N is sufficiently large, which will be case since typically $N > N_{\min 1}$, after a de Sitter stage the Universe looks like an almost perfectly flat model. Therefore, the initial density plays almost no role; if the exponential expansion occurs the Universe becomes effectively flat; see Fig. 6. In this way, instead of appealing to very special initial conditions, one starts with a Universe with more normal -that is, not very fine tuned- conditions which permit the Universe to evolve to an inflationary stage, after which it looks like it would have very special conditions, i.e. with $\Omega \approx 1$ with exponential accuracy. Thereupon, the flatness problem is no longer present.

One can also observe this by geometric means. The *radius of curvature* of the Universe is defined as

$$R_{\text{curv}} \equiv \frac{a(t)}{|k|^{1/2}} = \frac{H^{-1}}{|\Omega - 1|^{1/2}} \quad , \tag{67}$$

where we have used (4). Initially one may find that $\Omega \approx \mathcal{O}(1)$, implying that $R_{\text{curv}} \sim H^{-1}$, but after inflation Ω is very close to unity; thus the radius of curvature is exponentially larger than the Hubble distance, making the latter look very flat.

We have argued that the Universe becomes very flat. However, after inflation the r.h.s. of (66) starts to grow linearly with time during the radiation dominated era, as was derived in (43), and during the matter era it grows as $\sim t^{2/3}$, cf. (42). Therefore, the time evolution of $\Omega(t)$ is given by

$$\Omega(t) - 1 = \frac{k}{a^2 H^2} = k\ e^{-2N} \frac{t_{\text{eq}}}{t_f} \left(\frac{t}{t_{\text{eq}}}\right)^{2/3} \quad \text{for} \quad t_{\text{eq}} \le t \quad . \tag{68}$$

If the number of e-folds of inflation is not sufficient, at some t, Ω will be very different from one. Accordingly, from (68) one can compute the minimal number of e-folds such that currently $\Omega_0 \approx \mathcal{O}(1)$, as suggested from observations. One gets,

$$N > \frac{1}{2\ \log\ e} \left(\log \frac{t_{\text{eq}}}{t_f} + \frac{2}{3} \log \frac{t_0}{t_{\text{eq}}}\right) \equiv N_{\min 2} \quad , \tag{69}$$

which with the values used above implies that $N > 56 = N_{\min 2}$. However, the value of N chosen greater than (65), $N > N_{\min 1}$, already fulfills our new requirement, since $N_{\min 1} \gtrsim N_{\min 2}$. Therefore, in most models of inflation $N \gg N_{\min 1}$ to predict Ω_0 to be unity to high accuracy, in accordance with the recent WMAP measurements [13] of the CMBR. However, if Ω is close, but different than the unity, then there will come a future time when the value of Ω is arbitrarily close to zero (for $k = -1$) or arbitrarily large (for $k = 1$), the same as happens in the SBB; compare Figs. 4 and 6.

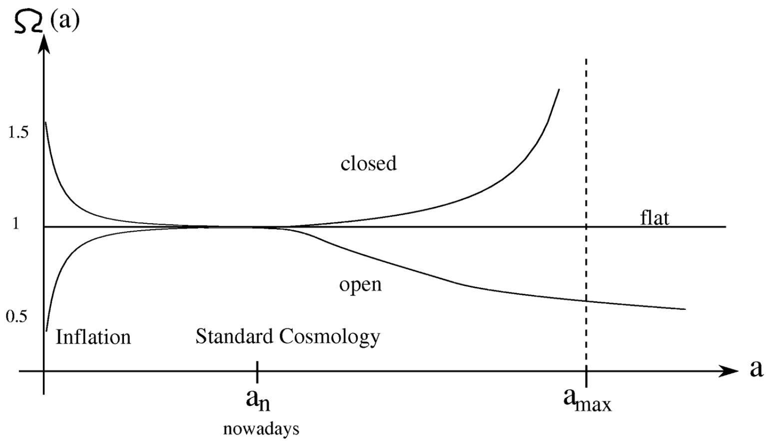

Fig. 6. The parameter Ω as a function of the scale factor, a, during inflation and thereafter in a radiation dominated Universe. Inflation makes the space seem almost flat. Having enough e-folds of inflation to solve the horizon problem implies that the Universe still looks currently very flat. Later on, the behavior is as in Fig. 4.

2.2 Transition to the Physical Universe

Provided that the Universe underwent a period of $N(\geq N_{\min 1} \approx 60)$ e-folds of inflation, it seems that the horizon and flatness problems are no longer present: all physical events are, or were, in causal contact. However, as we shall see now, this is only a necessary, but not a sufficient, condition to assure that the original patch, H_*^{-1}, contains the properties of our Universe today.

In Sect. 1 we saw that both problems are related to the increase of entropy per horizon and that the age of the Universe is related to the total entropy, cf. (32): the Universe is too old because the entropy is too large. Then, within the inflationary Universe there arises the question of how the $N_{\min's}$ are related to the entropy enhancement.

In this section we have remarked that with the aid of inflation any co–moving length scale remains within causal horizon. Therefore, in the present case the entropy per causal horizon remains constant if the Universe evolves adiabatically ($T \sim 1/a$), and therefore it is no longer necessary to distinguish between horizon entropy and the total entropy. Thus, it suffices to compute the entropy at any time to know how large it is. Accordingly, at the very beginning $S_* = \frac{4}{3}b(a_* T_*)^3 \sim 10^{13}$ with initial conditions taken from above. Clearly, this value is much smaller than the observed photon entropy today, $S_0 \sim 10^{88}$. Once the entropy augments to 10^{88}, the flat patch can explain the currently observed Universe with a big mass and age; see (31) and (32). In this way, to fully solve the horizon and flatness problems, one has to find an entropy production mechanism such that the increase factor grows over 10^{75} orders of magnitude! Would this entropy enhancement not exist, the original patch must contain a huge entropy (10^{88}), and therefore, the Universe would consist at the very beginning of too many disconnected causal horizons; ergo, the horizon problem would stay unsolved. The natural solution is to obtain after inflation a mechanism by which the entropy increases from some initial value to $S_0 \sim 10^{88}$. To see how this happens, consider first a Universe model filled with some relativistic components with an energy density given by

$$\rho = bT^4 + V(0) \ , \tag{70}$$

where $V(0) \equiv M^4$ is a constant associated with the vacuum energy density of some GUT; M is some mass term. As the Universe cools the energy density diminishes until certain time, say, $t = t_c$, at which the constant dominates the dynamics over the radiative components. At that moment the entropy within the horizon is $S_{d_H} = \frac{4}{3}b(a_c T_c)^3 = \text{const.}$, assuming adiabatic processes. At the moment when the constant $V(0)$ begins to dominate the dynamics, the solution is given by (10). Then, the original patch, $a_c \sim H_c^{-1}$, expands exponentially and the Universe supercools $T = T_c e^{-H\tau}$, since $aT = \text{const.}$ during the τ-stage of inflation. Note that the entrance to an inflationary era is natural as a consequence of the Universe cooling and, of course, of the presence of the constant $V(0)$. Typically, $M \sim T_c \sim 10^{14}$ GeV is related to the critical temperature of a spontaneous symmetry–breaking process, whereas $H_c^{-1} \approx 10^{-1} M_{Pl}/M^2 \sim 10^{-34}$s, that is, the values we have chosen above to yield $S_{d_H} \sim 10^{13} \ll 10^{88}$. Furthermore, after inflation the Universe contains a very low particle density and is very cold, even as cold as it is today! The transition to a radiation–(or matter–)dominated era with sufficient entropy and particle content comes from the 'decaying' or transformation of the energy source of inflation, $\rho = V(0)$, into heat, a process called *reheating* (RH). In his original model of inflation A. Guth [46] showed that if the Universe super–cools sufficiently its temperature is $T_s \sim e^{-N} T_c$ and a phase transition proceeds releasing latent heat of characteristic temperature $\sim T_c$. Then, the Universe is reheated to some temperature (T_{RH}) of the order of T_c. Through this mechanism the entropy increases from the initial value $S_* = \frac{4}{3}b(a_c T_c)^3$ to the final, after reheating, $S_f = \frac{4}{3}b(a_{RH} T_{RH})^3 \approx \frac{4}{3}b(e^N a_c T_c)^3$, that is, by a factor of $(T_{RH}/T_s)^3 \sim e^{3N}$, achieving the desired numbers.

2.3 Density Perturbations

The vacuum energy density responsible for inflation is associated with some scalar field that experiences quantum fluctuations around some vacuum expectation value. The theory of quantum fluctuations in a de Sitter space was developed by Bunch and Davies [17] and was applied by several authors [50, 47, 88, 8] to the inflationary universe in order to compute its contribution to $\delta\rho/\rho$.

It is our intention in this subsection to depict qualitatively how these fluctuations are responsible for a density perturbation spectrum, as required to understand structure formation; see the problem of structure formation above. For a detailed treatment of this topic see the lecture by R. Brandenberger in this book and his review article [69] .

We begin by noting that the event horizon during a de Sitter stage is $d_e \approx H^{-1}$, cf. (62). This means that microphysics can only operate coherently within distances at most as big as the Hubble horizon, H^{-1}. Recall that the causal horizon, d_H, expands exponentially and it is very large compared to the almost constant H^{-1} during inflation, see (59). Hence, during the de Sitter stage the generation of perturbations, which is a causal microphysical process, is localized in regions of the order of H^{-1}. That is, in all regions of size H^{-1} that comprise the Universe during inflation there should be such generation of perturbations.

Furthermore, it has been shown that the amplitude of inhomogeneities produced corresponds to the Hawking temperature in the de Sitter space, $T_H = H/(2\pi)$. In turn, this means that perturbations with a fixed physical wavelength of size H^{-1} are produced throughout the inflationary era. Accordingly, a physical scale associated with a quantum fluctuation, $\lambda_{\rm phys} = \lambda a(t)$, expands exponentially and once it leaves the (Hubble) horizon it behaves as a metric perturbation; its description is then classical, general relativistic. If inflation lasts for enough time, the physical scale can grow as much as a galaxy– or horizon–sized perturbation. The field fluctuation always expands with the scale factor and after inflation it evolves according to t^n ($n = 1/2$ radiation or $n = 2/3$ matter). On the other hand, the Hubble horizon evolves after inflation as $H^{-1} \sim t$. This means that there will come a time at which field fluctuations cross inside the Hubble horizon and re-enter as density fluctuations. Thus, inflation produces a large spectrum of perturbations, of which the largest originated at the start of inflation with a size H_i^{-1}, and the smallest with H_f^{-1} at the end of inflation; see Fig. 7.

Finally, analogous to the density perturbations spectrum created by the scalar field during inflation, any massless (or very light $H \gg M$) field is excited in the de Sitter space. Once the excited modes re–enter the Hubble horizon they will propagate as particles [94]. In this way, gravitational wave (GW) perturbations re–enter the Hubble horizon during the radiation do-

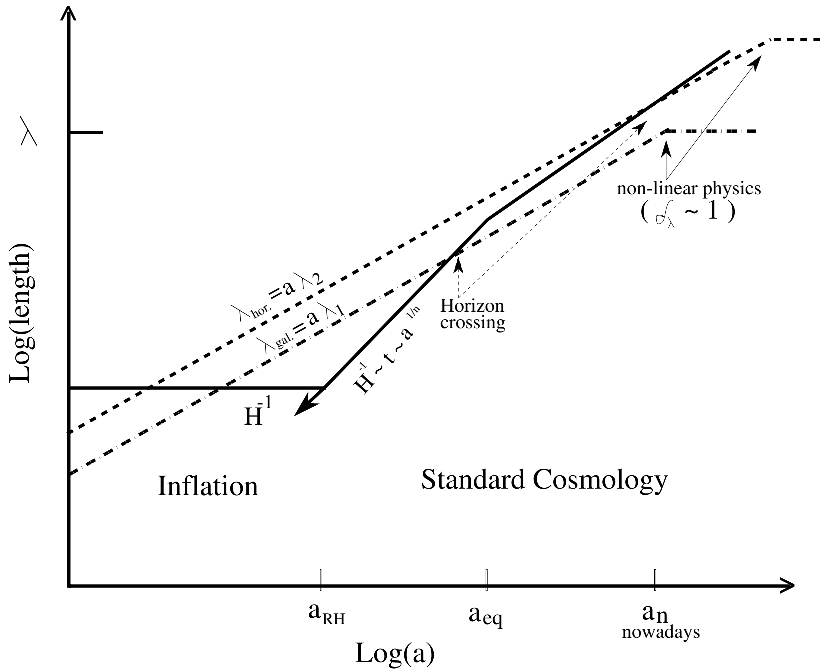

Fig. 7. Quantum perturbations were initially subhorizon–sized. During inflation they grow exponentially ($\lambda_{\rm phys.} = \lambda a(t)$), whereas the Hubble horizon remains almost constant. Then they eventually cross outside H^{-1} and evolve as classical perturbations. Later on, they re-enter the Hubble horizon to produce an almost scale invariant, Harrison-Zel'dovich density perturbation spectrum; in this way, its origin is no longer a mystery. In the figure there are two physical perturbations scales depicted, galaxy and horizon sized. (Figure adapted from Kolb and Turner 1990).

minated era [45, 87, 83, 35]. The amplitude of these perturbations must be $\leq 10^{-5}$ in order to be consistent with the isotropy levels of the CMBR.

2.4 Final Remarks on Inflationary Models

The descriptions presented above generically describe the inflationary scenario without referring to the many specific models that have been proposed in the course of more than twenty years of inflationary cosmology. Below we mention the primary features of the first inflationary models that permitted the understanding of the physical properties of a general inflationary scenario.

The first inflationary model was due to A. Guth, who in 1981 published [46] a model which later became known as *old inflation*. This model proposed the formation of bubbles of scalar fields obeying a first–order GUT phase transition in the early Universe; see Fig. 8. However, the nucleation of the bubbles and Hubble expansion rates could not encompass the right numbers of e-folds required ($\geq N_{\rm min}$) and, at the same time, achieve a homogeneous thermalized model. The model also had difficulties in ending the inflationary period, a problem known as *graceful exit*.

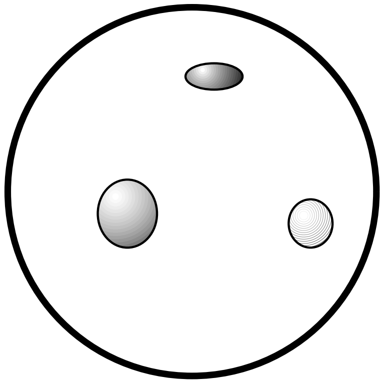

Fig. 8. An artistic picture of how the formation of bubbles after inflation should take place. The big circle represents the universe filled with a background of false vacuum, $\phi = 0$. Thus, some bubbles are forming in a sea of false vacuum. The different bubble ϕ-field values are represented by different gray tones.

Shortly afterwards, a model called *new inflation* was proposed [63, 2], in which the scalar field experiences a second–order phase transition. In this model the whole Universe stems from an initially uniform single bubble or fluctuation region ($\leq d_H$), which after its exponential expansion can be as large as our apparent horizon. In this way the Universe does not possess the problem of bubble nucleation, nor that of the presence of unwanted relics, because they are produced at bubble boundaries, which in this case are exponentially far away from our apparent horizon. The source of exponential expansion is achieved by permitting the scalar field to slowly evolve from its symmetric state ($\phi = 0$) to its ground state, $\phi = v$, of a typical potential $V(\phi) = \lambda(\phi^2 - v^2)^2$. During this time the potential associated with the scalar field is almost constant, thus providing an effective cosmological constant to the FRW equations; see Fig. 9. The process of *slow rolling down* of ϕ along its potential curvature is the main new ingredient of this model of achieving an inflationary stage.

After inflation the ϕ-field begins to oscillate around its stable minimum, $\phi = v$. The energy stored in $V(0)$ decays to reheat the Universe to have the required entropy. One must point out that over the course of the years, important steps to consolidate the theory of reheating have been made; see for example [93]. But, qualitative new ideas have been introduced only since the

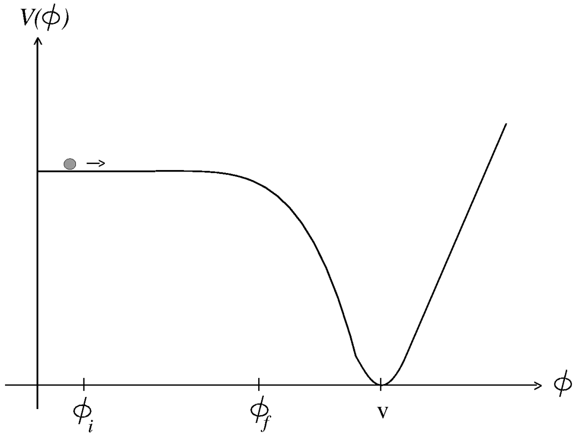

Fig. 9. A sketch of the new inflationary potential is shown. The potential curvature is very flat in order to permit the field to slow roll down the hill to yield enough e-folds of inflation during that time. Inflation begins at some ϕ_i and ends at ϕ_f when the field begins to evolve rapidly to its stable symmetry–breaking state $\phi = v$, around which the field oscillates until reheating.

mid 90s [57, 58]. Accordingly, the process of reheating should consist of three different stages. At the first phase, the ϕ-field decays into massive bosons (fermions) due to parametric resonance given through a Mathieu equation that determines the regions of stability and instability (particle production) in the quantum fluctuations of the created particles. These can be ϕ-particles or other bosons (fermions) coupled to the ϕ-field. This process is very efficient, even explosive, and many bosons can be created in this stage. Note that the original theory is based upon the decay of the ϕ-particles, whereas in the present theory the ϕ-field decays into ϕ-particles, and perhaps others, and only after this process does the decay of ϕ-particles proceed. Then, to distinguish this explosive process from the normal stage of particle decay, the authors of [57] have called it *preheating*. Bosons produced at this stage are far from thermal equilibrium and have very big occupational numbers. The second stage of this scenario describes the decay of the already produced particles. This phase is described as in the original theory. Thus, the methods developed for the original theory are now applied to the product particles, but not to the decay of the ϕ-field itself. The third stage is the thermalization by which the system reaches equilibrium; for review of this topic see [59].

A very important result found in the context of the "new" inflationary model is that perturbations of the scalar field can explain the required initial conditions of structure formation, i.e. an almost scale invariant, Harrison-Zel'dovich spectrum. This spectrum results from the original quantum fluctuations of the ϕ-field. These field fluctuations cross outside the Hubble horizon during inflation, evolve classically and, eventually, return back to re–enter the Hubble horizon as density perturbations with scales of galaxy or present–horizon size, see Sect. 2.3. But its correct accomplishment demands, according with the COBE and WMAP measurements, a magnitude of $\delta\rho/\rho \sim 10^{-5}$ when perturbations re–enter the Hubble horizon. This fact demands the above potential, or the generic toy *chaotic model* potential $V = \lambda\phi^n$ [64], to have an extremely small value for λ which in turn implies very particular choices of particle physics models (fine–tuning), or even wrong models. In this way, some inconsistencies appear. Inflation was thought to circumvent the choosing of special initial conditions of the SBB, but now we see that it encounters its own. In spite of this and other difficulties, the new and the chaotic inflationary models served to show how the very idea of field slow rollover dynamics can be implemented in many particle physics and/or gravity theories with general success.

Over the course of two decades of the inflationary theory, many related scenarios and models have been proposed with concrete physical mechanisms to achieve inflation, reheating, baryogenesis and a causal perturbation spectrum, among others. It has been found that many models generally suffer from "unnatural" fine–tuning of parameters. Nevertheless, some of these models have interesting properties, and the relation among theoretical and observational cosmology and particle physics has become tighter than ever. For instance, typical unification theories have different scalar fields which have been used to have one or more inflationary stages. Additionally, they have been used to produce the correct density perturbation spectrum together with a sufficient reheating temperature. That is, it is tempting to use the various fields for achieving different cosmological tasks, as in the case of *hybrid inflation*. Then, to distinguish among the different fields, the field responsible for the period of exponential expansion is generically called the *inflaton*. The modern view is that this inflaton is a primary ingredient in offering a solution to the above–mentioned problems. A general description of the scalar field (inflaton) dynamics, as well as some of its quantitative parameters, are found in the contributions of E. Copeland and C. A. Terrero–Escalante in this book.

3 Overview

Finally, we are going to review the topics that were covered in this contribution. We have presented a general view of the SBB, its problems, and the main ideas involved in the inflationary Universe. We have shown how inflation achieves an explanation of the horizon and flatness problems. In doing so, a period exponential expansion of about 60 e-folds in the scale factor is necessary. However, this condition is not sufficient at all to yield a Universe like the one we live in: after inflation the region of the Universe which will give rise to our present apparent horizon is almost devoid of particles and, because of the adiabatic exponential expansion, it is also very cold. Therefore, a process of reheating is mandatory. GUT offer through phase transition phenomena an appealing scenario for creating both a constant energy density in the very early Universe, to have inflation, and its subsequent preheating and reheating. After inflation Baryogenesis can take place in the context of GUT, a scenario which can attain the Sakharov required conditions.

We have also shown that the inflationary theory provides means to solve the monopole and other relics problems within the new inflationary scenario. Furthermore, there are inflationary models with no singularity, because they begin with finite initial conditions, for instance, those attached to the potential energy. In this way, one excludes singularities by appealing to the physical limitations of the classical theory; this is the case for chaotic inflation. Other inflationary models without initial singularity have also been proposed [95].

The homogeneity and isotropy of the large–scale structure of the Universe are also a consequence of a long period of exponential expansion. Recall that all inhomogeneities are shifted away from the Hubble horizon. Thus, inflation makes the space very homogeneous and isotropic. This assertion is known as the *cosmic no hair* conjecture [43, 15]. The question remains whether the initial patch was sufficiently smooth to enable inflation to start. That is, in some sense, the Universe should be, at some level homogeneous and isotropic to consider it as an original patch with which to begin inflation. It turns out that scalar fields, initially without considering potential terms, can bring an initially homogeneous, anisotropic patch to an almost FRW symmetry [20, 67]. That is, an anisotropic Universe can begin with such initial conditions that the potential term is yet not the primary contribution to the cosmological field equations and, after some time, it can become nearly isotropic. Then, inflation occurs when the potential term begins to dominate [21, 22, 23].

The cosmological constant, $\Lambda = V(0)$, provides the energy density to have an exponential expansion. However, in the inflationary theory this constant must be assumed, and inflation provides no explanation of this. From the particle physics point of view, it is also intriguing to consider why this constant should or should not exist; there is no known principle that demands it to vanish. Thus, for convenience one usually assumes $V(0) \neq 0$ in order to be able to achieve inflation and to have today $V(v) = 0$ with $\phi = v$. In this way, one avoids choosing the tuned value $V(v) \lesssim \rho_{c_0} \approx 10^{-47}$ GeV^4 today. However, this last possibility is very interesting in the context of the present huge expansion rate; see below.

It is worthwhile to notice that because inflation predicts $\Omega_0 \approx 1$ and since our observed baryonic Universe only contributes $\Omega_B \sim 0.05$ to the total energy density, then inflation predicts some amount of dark components, namely, $\Omega_{\rm dark} \sim 0.95$! The motivation for one of the dark candidates, Λ, came from different cosmological measurements which without a cosmological constant (or function) are rather difficult to explain. For instance, W. Priester et al. [51, 52, 14] pointed out that with a present cosmological constant one can explain the absorption lines of quasars, the so called Ly α-forest spectrum, assuming a Hubble constant of $H_0 = 90$ km s^{-1} Mpc^{-1}. Further, measurements of the Hubble constant employing a variety of techniques suggested a rather high value for it: $H_0 = 80 \pm 5$ km s^{-1} Mpc^{-1} [39]. Also, the *Hubble Space Telescope* (HST) measurements of Cepheid variable stars in the Virgo cluster evidence such high values as $H_0 = 80 \pm 17$ km s^{-1} Mpc^{-1} [36, 71]. If this evidence is correct, it turns out that the age of a flat Big Bang Universe (suggested by inflation) is too small, $t_0 = \frac{2}{3} H_0^{-1} = 8 - 10$ Gyr, to explain the oldest globular clusters estimated to be 16 ± 3 Gyr [85] or 9.3 ± 2 Gyr [99]. This means that some other contribution to the FRW cosmological equations should be present to let the Universe be older. This effect is carried out precisely by a cosmological constant, or function term. This is so because Λ corresponds to a negative pressure (repulsive force) so that the expansion rate first decreases more slowly (than if $\Lambda = 0$) and eventually decreases faster, yielding a larger expansion age.

Astonishingly, recent, independent observational data measured in the CMBR on various angular scales [26, 13], in type Ia supernovae[14] [79, 76, 80], as well as in the 2dF Galaxy Redshift survey [75, 34], suggest that $\Omega = \Omega_\Lambda + \Omega_\mathrm{m} \approx 1$, or $\Omega_\Lambda \approx 0.7$ and $\Omega_\mathrm{m} \approx 0.3$, implying the existence of dark energy and dark matter, respectively. One particular candidate for dark energy is a scalar field usually called quintessence [24]. Naturally, particular inflationary scenarios motivated from different particle physics theories have their own dark matter candidates, as such the Axion, neutralino, Higgs particle, etc, and additionally a quintessence field; see the contributions of E. Copeland and A. de la Macorra in this book.

Additionally, three–dimensional numerical simulations of structure formation have incorporated cold, warm, or hot particles into their analyses and, up to now, the best fittings with sky surveys turn out to be a mixture of the different dark matter ingredients with $\Omega_\mathrm{matter} \sim 0.3$, also including a cosmological constant with $\Omega_\Lambda \sim 0.7$ [4].

Next we comment on the first two problems listed in Sect. 1: the dimensionality and euclidicity problems. They seem to go beyond the scope of the inflationary theory. They touch the foundations of a theory of everything, including gravity. However, some cosmological solutions in Kaluza Klein theories have been found that are related to modern particle physics and gravity theories [37]. With the advent of fundamental string theory, some cosmological solutions have been found that compactify the $D - 4$ dimensions to four [28], and there are inflationary solutions stemming from effective string theories in which various fields exist; one of them, the dilaton, plays the role of the inflaton [41, 42, 16]. The issue of string cosmology has become of much interest in recent years, and modern implementations are accomplished within braneworld scenarios; this topic is extensively explained in the contributions of K. Maeda and J. Lidsey in this book.

[14] See the contribution of A. Filippenko in this book.

Inflation turns out to be a possible, natural, cosmological realization of high energy physics with qualitatively outstanding results. However, a better implementation of it must be achieved, perhaps within new theories or as extensions of the known ones, such as the ones presented in the forthcoming chapters of this book.

Acknowledgements

This work has been supported by CONACYT grant number 33278-E.

References

1. Abbott L F, Farhi E and Wise M B 1982 *Phys. Lett. B* **117** 29.
2. Albrecht A and Steinhardt P J 1982 *Phys. Rev. Lett.* **48** 1220.
3. Albrecht A, Steinhardt P J, Turner M S and Wilczek F 1982 *Phys. Rev. Lett.* **48** 1437.
4. Varios Contributions in 2003 *Rev Mex A A (SC)*, **17**, eds. Avila-Reese V Firmani C, Frenk C S and Allen C.
5. Bahcall N A 1988 *Ann. Rev. Astron. Astrophys.* **26** 631.
6. Balbi A et al. 2000, *Astrophys. J.* **545** L1.
7. Banday A J *et al* 1994 *Astrophys. J.* **436** L99.
8. Bardeen J M, Steinhardt P J and Turner M S 1983 *Phys. Rev. D* **28** 679.
9. Barrow J D 1993 *Phys. Rev. D* **47** 5329.
10. Barrow J D and Sonoda D H 1986 *Phys. Reports* **139** 1.
11. Bennett C L *et al* 1994 *Astrophys. J.* **436** 423.
12. Bennett C L *et al* 1996 *Astrophys. J.* **464** L1.
13. Bennett C L *et al* 2003 *Astrophys. J. Suppl.* **148** 1.
14. Blome H J and Priester W 1991 *Astron. Astrophys.* **250** 43.
15. Boucher W and Gibbons G W 1983 *The very early universe* Editors: G.W. Gibbons and S.W. Hawking (Cambridge Univ. Press) 273.
16. Brustein R, Gasperini M, Giovannini M, Mukhanov V F and Veneziano G 1995 *Phys. Rev. D* **51** 6744.
17. Bunch T S and Davies P C W 1978 *Proc. Roy. Soc. London A* **360** 117.
18. Cahill K and Podolský J 1994 *J. Phys. G* **20** 571.
19. Carrol S, Press W and Turner E 1992 *Ann. Rev. Astron. Astrophys.* **30** 499.
20. Chauvet P and Cervantes-Cota J L 1995 *Phys. Rev. D* **52** 3416.
21. Cervantes-Cota J L 1996 *Ph.D. thesis: Induced gravity and Cosmology* (University of Konstanz, Hartung Verlag)
22. Cervantes-Cota J L and Chauvet P 1999 *Phys. Rev. D* **59** 043501-1.
23. Cervantes-Cota J L 1999 *Class. Quant. Grav.* **16** 3903.
24. Caldwell R R Dave R and Steinhard P J 1998 *Phys. Rev. Lett.*, **80** 1582.
25. Collins C B and Hawking S W 1973 *Astrophys. J.* **180** 317.
26. de Bernardis P. et al, 2000, Nat, 404, 995.
27. de Sitter W 1917 *Proc. Kon. Ned. Akad. Wet.* **19** 1217; ibid **20** 229; 1916 *Mon. Not. R. Astron. Soc.* **76** 699; ibid **77** 155; 1917 ibid **78** 3.

28. de Vega H J and Sánchez N 1995 *Current Topics in Astrofundamental Physics: The Early Universe*. Editors: N. Sánchez and A. Zichichi (Kluwer Acad. Pu.) 99.
29. Dicke R H, Peebles P J E, Roll P G and Willkinson D T 1965 *Astrophys. J.* **142** 414.
30. Dicke R H and Peebles P J E 1979 *General Relativity: An Einstein centenary survey* Editors: S.W. Hawking and W. Israel (Cambridge Univ. Press) 504.
31. Dolgov A D 1992 *Phys. Reports* **222** 309.
32. Dolgov A D 1994 *Nucl. Phys. B* (Proc. Suppl.) **35** 28.
33. Dolgov A D and Linde A D 1982 *Phys. Lett. B* **116** 329.
34. Efstathiou G 2002 *MNRAS* **330** L29.
35. Fabri R and Pollock M D 1983 *Phys. Lett. B* **125** 445.
36. Freedman W L *et al* 1994 *Nature* **371** 757.
37. Freund P G O 1982 *Nucl. Phys. B***209** 146.
38. Friedmann A 1922 Zeit. f. Phys. **10** 377; 1924 ibid **21** 326.
39. Fukugita M, Hogan C J and Peebles P J E 1993 *Nature* **366** 309.
40. Gamov G 1946 *Phys. Rev.* **70** 572; 1948 ibib **74** 505.
41. Gasperini M and Veneziano G 1992 Phys. Lett. B **277** 256;
42. Gasperini M and Veneziano G 1994 *Phys. Rev. D* **50** 2519.
43. Gibbons G W and Hawking S W 1977 *Phys. Rev. D* **15** 2738.
44. Gleiser M, Lect. Notes Phys. 455, 387 (1995). Preprint: hep-ph/9407369.
45. Grishchuk L P 1975 *Sov. Phys. JETP* **40** 409.
46. Guth A H 1981 *Phys. Rev. D* **23** 347.
47. Guth A H and Pi S Y 1982 *Phys. Rev. Lett.* **49** 1110.
48. Hanany S et al. 2000, ApJ, **545**, L5.
49. Harrison E R 1970 *Phys. Rev. D* **1** 2726.
50. Hawking S W 1982 *Phys. Lett. B* **115** 295.
51. Hoell J and Priester W 1991 *Comments Astrophys.* **15** 127.
52. Hoell J and Priester W 1991 *Astron. Astrophys.* **251** L23.
53. Hubble E P 1929 *Proc. Nat. Acad. Sci.* **15** 168.
54. Kaluza T 1921 *Sitzungsber. Preus. Akad. Wiss. Berlin* **1** 966.
55. Kibble T W B 1976 *J. Phys. A* **9** 1387.
56. Klein O 1926 Zeit. Phys. **37** 895; 1926 *Nature* **118** 516.
57. Kofman L, Linde A and Starobinsky A A 1994 *Phys. Rev. Lett.* **73** 3195.
58. Kofman L, Linde A and Starobinsky A A 1996 *Phys. Rev. Lett.* **76** 1011.
59. Kofman L, Linde A and Starobinsky A A 1997 *Phys. Rev. D* **56** 3258.
60. Kolb E W and Turner M S 1990 *The Early Universe* "Frontiers in Physics" no. 69 (Addison-Wesley).
61. Kuzmin V A, Rubakov V A and Shaposhnikov M E 1985 *Phys. Lett. B* **155** 36.
62. Liddle A R and Lyth D H; 2000 *Cosmological Inflation and Large-Scale Structure*, Cambridge U.P., Cambridge, UK.
63. Linde A D 1982 *Phys. Lett. B* **108** 389.
64. Linde A D 1983 *Phys. Lett. B* **129** 177.
65. Linde A D 1990 *Particle Physics and Inflationary Cosmology*, (Harwood Ac.).
66. Lüst D and Theisen S 1989 *Lectures on string theory* (Springer-Verlag).
67. Mimoso J P and Wands D 1995 *Phys. Rev. D* **52** 5612.
68. Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (Freeman and Company).

69. Mukhanov V F, H.A. Feldman H A and Brandenberger R H 1992 *Phys. Reports* **215** 203.
70. Narlikar J V 1993 *Introduction to Cosmology* (Cambridge University press).
71. Pierce M *et al* 1994 *Nature* **371** 385.
72. Olive K A 1990 *Phys. Reports* **190** 307.
73. Penzias A A and Wilson R W 1965 *Astrophys. J.* **142** 419.
74. Peacock J A 1999 *Cosmological Physics* (Cambridge University press).
75. Peacock J A et al 2002 in *A New Era in Cosmology* (ASP Conference Proceedings), eds Shanks T and Metcalfe N. Preprint astro-ph/0204239.
76. Perlmutter S et al. 1999 *Astrophys. J.* **517** 565.
77. Polyakov A M 1974 *Sov. Phys. JETP* Lett. **20** 194.
78. Preskill J 1984 *Ann. Rev. Nucl. Part. Sci.* **34** 461.
79. Riess A G et al 1998 *Astronom. J.* **116** 1009.
80. Riess A G et al. 2001 *Astrophys. J.* **560** 49.
81. Robertson H P 1935 *Astrophys. J.* **82** 284; 1936 ibid **83** 187, 257.
82. Rindler W 1956 *Mon. Not. Roy. Astron. Soc.* **116** 663.
83. Rubakov V A, Sazhin M V and Veryaskin A V 1982 *Phys. Lett. B* **115** 189.
84. Sakharov A D 1967 *Sov. Phys. JEPT Lett.* **5** 24.
85. Sarajedini A and King C R 1989 *Astron. J.* **98** 1624.
86. Smoot G et al. 1992 *Astrophys. J.* **396** L1.
87. Starobinsky A A 1979 *Sov. Phys. JETP* Lett. **30** 682.
88. Starobinsky A A 1982 *Phys. Lett. B* **117** 175.
89. Steigman G 1976 *Ann. Rev. Astr. Astrophys.* **14** 339.
90. Steigman G 1979 *Ann. Rev. Nucl Part. Sci.* **29** 313.
91. 't Hooft G 1974 *Nucl. Phys. B***79** 276.
92. 't Hooft G 1976 Phys. Rev. Lett **37** 8; 1976 *Phys. Rev. D* **14** 3432.
93. Traschen J H and Brandenberger R H 1990 *Phys. Rev. D* **42** 2491.
94. Turner M S and Widrow L M 1988 *Phys. Rev. D* **37** 3428.
95. Vilenkin A 1982 *Phys. Lett. B* **117** 25.
96. Walker A G 1937 *Proc. Lond. Math. Soc. (2)* **42** 90.
97. Weinberg S 1972 *Gravitation and Cosmology: principles and applications of the general theory of relativity* (John Wiley & Sons).
98. Weinberg S 1989 *Rev. Mod. Phys.* **61** 1.
99. Winget D E *et al* 1987 *Astrophys. J.* **315** L77.
100. Wright E l *et al* 1994 *Astrophys. J.* **436** 443;
101. Zel'dovich Y B 1972 *Mon. Not. R. Astron. Soc.* **160** 1p.
102. Zel'dovich Y B and Khlopov M Y 1978 *Phys. Lett. B* **79** 239.