Cosmology

J. García-Bellido
Instituto de Física Teórica IFT-UAM/CSIC, Cantoblanco 28049 Madrid, Spain

Abstract
In these lectures I review the present status of the so-called Standard Cosmological Model, based on the hot Big Bang Theory and the Inflationary Paradigm. I will make special emphasis on the recent developments in observational cosmology, mainly the acceleration of the universe, the precise measurements of the microwave background anisotropies, and the formation of structure like galaxies and clusters of galaxies from tiny primordial fluctuations generated during inflation.

1 Introduction

The last ten years have seen the coming of age of Modern Cosmology, a mature branch of science based on the hot Big Bang theory and the Inflationary Paradigm. In particular, we can now define rather precisely a Standard Model of Cosmology, where the basic parameters are determined within small uncertainties, of just a few percent, thanks to a host of experiments and observations. This precision era of cosmology has become possible thanks to important experimental developments in all fronts, from measurements of supernovae at high redshifts to the microwave background anisotropies, as well as to the distribution of matter in galaxies and clusters of galaxies.

In these lecture notes I will first introduce the basic concepts and equations associated with hot Big Bang cosmology, defining the main cosmological parameters and their corresponding relationships. Then I will address in detail the three fundamental observations that have shaped our present knowledge: the recent acceleration of the universe, the distribution of matter on large scales and the anisotropies in the microwave background. Together these observations allow the precise determination of a handful of cosmological parameters, in the context of the inflationary plus cold dark matter paradigm.

2 Big Bang Cosmology

Our present understanding of the universe is based upon the successful hot Big Bang theory, which explains its evolution from the first fraction of a second to our present age, around 13.6 billion years later. This theory rests upon four robust pillars, a theoretical framework based on general relativity, as put forward by Albert Einstein [1] and Alexander A. Friedmann [2] in the 1920s, and three basic observational facts: First, the expansion of the universe, discovered by Edwin P. Hubble [3] in the 1930s, as a recession of galaxies at a speed proportional to their distance from us. Second, the relative abundance of light elements, explained by George Gamow [4] in the 1940s, mainly that of helium, deuterium and lithium, which were cooked from the nuclear reactions that took place at around a second to a few minutes after the Big Bang, when the universe was a few times hotter than the core of the sun. Third, the cosmic microwave background (CMB), the afterglow of the Big Bang, discovered in 1965 by Arno A. Penzias and Robert W. Wilson [5] as a very isotropic blackbody radiation at a temperature of about 3 degrees Kelvin, emitted when the universe was cold enough to form neutral atoms, and photons decoupled from matter, approximately 380,000 years after the Big Bang. Today, these observations are confirmed to within a few percent accuracy, and have helped establish the hot Big Bang as the preferred model of the universe.

Modern Cosmology begun as a quantitative science with the advent of Einstein's general relativity and the realization that the geometry of space-time, and thus the general attraction of matter, is

determined by the energy content of the universe [6]

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

These non-linear equations are simply too difficult to solve without invoking some symmetries of the problem at hand: the universe itself.

We live on Earth, just 8 light-minutes away from our star, the Sun, which is orbiting at 8.5 kpc from the center of our galaxy,[1] the Milky Way, an ordinary galaxy within the Virgo cluster, of size a few Mpc, itself part of a supercluster of size a few 100 Mpc, within the visible universe, approximately 10,000 Mpc in size. Although at small scales the universe looks very inhomogeneous and anisotropic, the deepest galaxy catalogs like 2dF GRS and SDSS suggest that the universe on large scales (beyond the supercluster scales) is very homogeneous and isotropic. Moreover, the cosmic microwave background, which contains information about the early universe, indicates that the deviations from homogeneity and isotropy were just a few parts per million at the time of photon decoupling. Therefore, we can safely impose those symmetries to the univerge at large and determine the corresponding evolution equations. The most general metric satisfying homogeneity and isotropy is the Friedmann-Robertson-Walker (FRW) metric, written here in terms of the invariant geodesic distance $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ in four dimensions [6] $\mu = 0, 1, 2, 3,$[2]

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

characterized by just two quantities, a *scale factor* $a(t)$, which determines the physical size of the universe, and a constant K, which characterizes the *spatial* curvature of the universe,

$$^{(3)}R = \frac{6K}{a^2(t)} \qquad \begin{cases} K = -1 & \text{OPEN} \\ K = 0 & \text{FLAT} \\ K = +1 & \text{CLOSED} \end{cases} \tag{3}$$

Spatially open, flat and closed universes have different three-geometries. Light geodesics on these universes behave differently, and thus could in principle be distinguished observationally, as we shall discuss later. Apart from the three-dimensional spatial curvature, we can also compute a four-dimensional *space-time* curvature,

$$^{(4)}R = 6\frac{\ddot{a}}{a} + 6\left(\frac{\dot{a}}{a}\right)^2 + 6\frac{K}{a^2}\,. \tag{4}$$

Depending on the dynamics (and thus on the matter/energy content) of the universe, we will have different possible outcomes of its evolution. The universe may expand for ever, recollapse in the future or approach an asymptotic state in between.

2.1 The matter and energy content of the universe

The most general matter fluid consistent with the assumption of homogeneity and isotropy is a perfect fluid, one in which an observer *comoving with the fluid* would see the universe around it as isotropic. The energy momentum tensor associated with such a fluid can be written as [6]

$$T^{\mu\nu} = p\,g^{\mu\nu} + (p + \rho)\,U^\mu U^\nu\,, \tag{5}$$

where $p(t)$ and $\rho(t)$ are the pressure and energy density of the fluid at a given time in the expansion, as measured by this comoving observer, and U^μ is the comoving four-velocity, satisfying $U^\mu U_\mu = -1$. For such a comoving observer, the matter content looks isotropic (in its rest frame),

$$T^\mu_{\ \nu} = \text{diag}(-\rho(t),\, p(t),\, p(t),\, p(t))\,. \tag{6}$$

---

[1]One parallax second (1 pc), *parsec* for short, corresponds to a distance of about 3.26 light-years or 3.09×10^{18} cm.

[2]I am using $c = 1$ everywhere, unless specified, and a metric signature $(-, +, +, +)$.

The conservation of energy ($T^{\mu\nu}_{\;;\nu} = 0$), a direct consequence of the general covariance of the theory ($G^{\mu\nu}_{\;;\nu} = 0$), can be written in terms of the FRW metric and the perfect fluid tensor (5) as

$$\dot{\rho} + 3\frac{\dot{a}}{a}(p + \rho) = 0\,. \tag{7}$$

In order to find explicit solutions, one has to supplement the conservation equation with an *equation of state* relating the pressure and the density of the fluid, $p = p(\rho)$. The most relevant fluids in cosmology are barotropic, i.e. fluids whose pressure is linearly proportional to the density, $p = w\,\rho$, and therefore the speed of sound is constant in those fluids.

We will restrict ourselves in these lectures to three main types of barotropic fluids:

- *Radiation*, with equation of state $p_R = \rho_R/3$, associated with relativistic degrees of freedom (i.e. particles with temperatures much greater than their mass). In this case, the energy density of radiation decays as $\rho_R \sim a^{-4}$ with the expansion of the universe.
- *Matter*, with equation of state $p_M \simeq 0$, associated with nonrelativistic degrees of freedom (i.e. particles with temperatures much smaller than their mass). In this case, the energy density of matter decays as $\rho_M \sim a^{-3}$ with the expansion of the universe.
- *Vacuum energy*, with equation of state $p_V = -\rho_V$, associated with quantum vacuum fluctuations. In this case, the vacuum energy density remains constant with the expansion of the universe.

This is all we need in order to solve the Einstein equations. Let us now write the equations of motion of observers comoving with such a fluid in an expanding universe. According to general relativity, these equations can be deduced from the Einstein equations (1), by substituting the FRW metric (2) and the perfect fluid tensor (5). The $\mu = i$, $\nu = j$ component of the Einstein equations, together with the $\mu = 0$, $\nu = 0$ component constitute the so-called Friedmann equations,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{K}{a^2}\,, \tag{8}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}\,. \tag{9}$$

These equations contain all the relevant dynamics, since the energy conservation equation (7) can be obtained from these.

2.2 The Cosmological Parameters

I will now define the most important cosmological parameters. Perhaps the best known is the *Hubble parameter* or rate of expansion today, $H_0 = \dot{a}/a(t_0)$. We can write the Hubble parameter in units of 100 $\mathrm{km\,s^{-1}Mpc^{-1}}$, which can be used to estimate the order of magnitude for the present size and age of the universe,

$$H_0 \equiv 100\,h \;\mathrm{km\,s^{-1}Mpc^{-1}}\,, \tag{10}$$

$$c\,H_0^{-1} = 3000\,h^{-1}\;\mathrm{Mpc}\,, \tag{11}$$

$$H_0^{-1} = 9.773\,h^{-1}\;\mathrm{Gyr}\,. \tag{12}$$

The parameter h was measured to be in the range $0.4 < h < 1$ for decades, and only in the last few years has it been found to lie within 4% of $h = 0.70$. I will discuss those recent measurements in the next Section.

Using the present rate of expansion, one can define a *critical* density ρ_c, that which corresponds to a flat universe,

$$\rho_c \equiv \frac{3H_0^2}{8\pi G} = 1.88\,h^2\,10^{-29}\;\mathrm{g/cm^3} \tag{13}$$

$$= 2.77 \, h^{-1} \, 10^{11} \, M_\odot/(h^{-1} \, \text{Mpc})^3 \tag{14}$$

$$= 11.26 \, h^2 \, \text{protons/m}^3 \,, \tag{15}$$

where $M_\odot = 1.989 \times 10^{33}$ g is a solar mass unit. The critical density ρ_c corresponds to approximately 6 protons per cubic meter, certainly a very dilute fluid!

In terms of the critical density it is possible to define the density parameter

$$\Omega_0 \equiv \frac{8\pi G}{3H_0^2} \, \rho(t_0) = \frac{\rho}{\rho_c}(t_0) \,, \tag{16}$$

whose sign can be used to determine the spatial (three-)curvature. Closed universes ($K = +1$) have $\Omega_0 > 1$, flat universes ($K = 0$) have $\Omega_0 = 1$, and open universes ($K = -1$) have $\Omega_0 < 1$, no matter what are the individual components that sum up to the density parameter.

In particular, we can define the individual ratios $\Omega_i \equiv \rho_i/\rho_c$, for matter, radiation, cosmological constant and even curvature, today,

$$\Omega_M = \frac{8\pi G \, \rho_M}{3H_0^2} \qquad\qquad \Omega_R = \frac{8\pi G \, \rho_R}{3H_0^2} \tag{17}$$

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2} \qquad\qquad \Omega_K = -\frac{K}{a_0^2 H_0^2} \,. \tag{18}$$

For instance, we can evaluate today the radiation component Ω_R, corresponding to relativistic particles, from the density of microwave background photons, $\rho_{\text{CMB}} = \pi^2 k^4 T_{\text{CMB}}^4/(15\hbar^3 c^3) = 4.5 \times 10^{-34}$ g/cm^3, which gives $\Omega_{\text{CMB}} = 2.4 \times 10^{-5} \, h^{-2}$. Three approximately massless neutrinos would contribute a similar amount. Therefore, we can safely neglect the contribution of relativistic particles to the total density of the universe today, which is dominated either by non-relativistic particles (baryons, dark matter or massive neutrinos) or by a cosmological constant, and write the rate of expansion in terms of its value today, as

$$H^2(a) = H_0^2 \left(\Omega_R \frac{a_0^4}{a^4} + \Omega_M \frac{a_0^3}{a^3} + \Omega_\Lambda + \Omega_K \frac{a_0^2}{a^2} \right) \,. \tag{19}$$

An interesting consequence of these definitions is that one can now write the Friedmann equation today, $a = a_0$, as a *cosmic sum rule*,

$$1 = \Omega_M + \Omega_\Lambda + \Omega_K \,, \tag{20}$$

where we have neglected Ω_R today. That is, in the context of a FRW universe, the total fraction of matter density, cosmological constant and spatial curvature today must add up to one. For instance, if we measure one of the three components, say the spatial curvature, we can deduce the sum of the other two.

Looking now at the second Friedmann equation (9), we can define another basic parameter, the *deceleration parameter*,

$$q_0 = -\frac{a \, \ddot{a}}{\dot{a}^2}(t_0) = \frac{4\pi G}{3H_0^2} \left[\rho(t_0) + 3p(t_0) \right] \,, \tag{21}$$

defined so that it is positive for ordinary matter and radiation, expressing the fact that the universe expansion should slow down due to the gravitational attraction of matter. We can write this parameter using the definitions of the density parameter for known and unknown fluids (with density Ω_x and arbitrary equation of state $w(z)$) as

$$q_0 = \Omega_R + \frac{1}{2}\Omega_M - \Omega_\Lambda + \frac{1}{2} \sum_x (1 + 3w) \, \Omega_x \,. \tag{22}$$

Uniform expansion corresponds to $q_0 = 0$ and requires a cancellation between the matter and vacuum energies. For matter domination, $q_0 > 0$, while for vacuum domination, $q_0 < 0$. As we will see in a
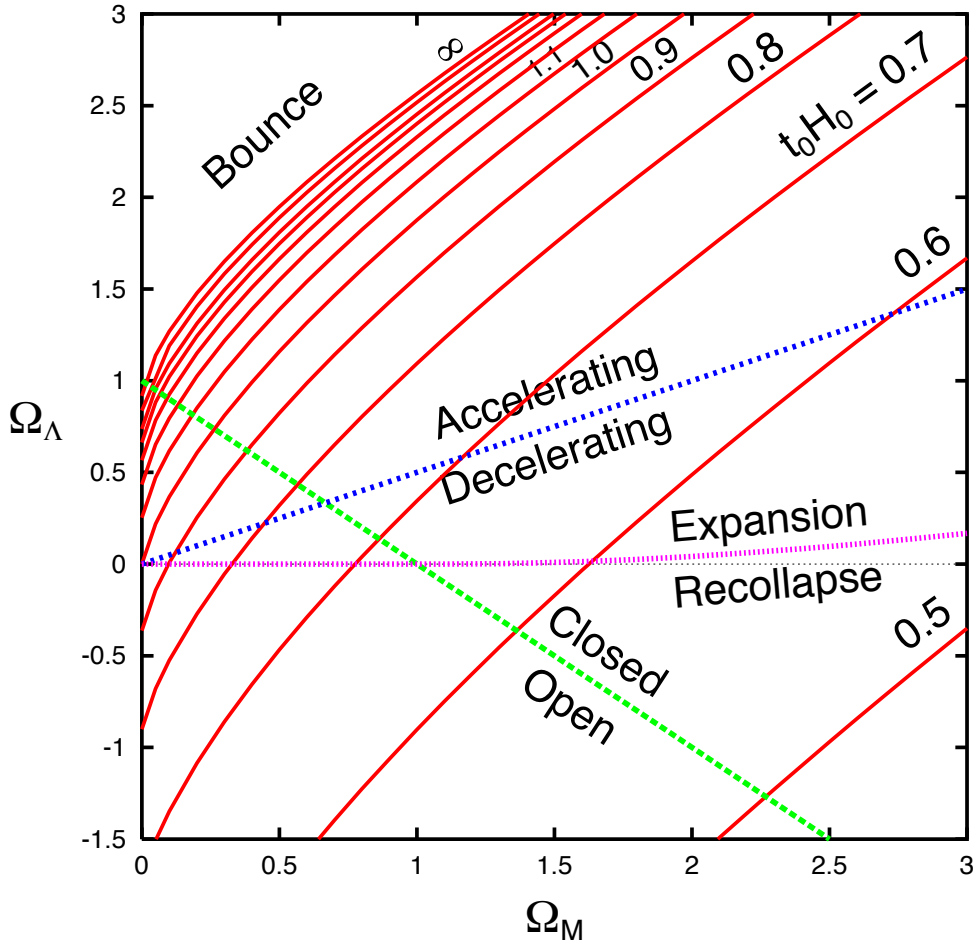
Fig. 1: Parameter space $(\Omega_M, \Omega_\Lambda)$. The green (dashed) line $\Omega_\Lambda = 1 - \Omega_M$ corresponds to a flat universe, $\Omega_K = 0$, separating open from closed universes. The blue (dotted) line $\Omega_\Lambda = \Omega_M/2$ corresponds to uniform expansion, $q_0 = 0$, separating accelerating from decelerating universes. The violet (dot-dashed) line corresponds to critical universes, separating eternal expansion from recollapse in the future. Finally, the red (continuous) lines correspond to $t_0 H_0 = 0.5, 0.6, \ldots, \infty$, beyond which the universe has a bounce.

moment, we are at present probing the time dependence of the deceleration parameter and can determine with some accuracy the moment at which the universe went from a decelerating phase, dominated by dark matter, into an acceleration phase at present, which seems to indicate the dominance of some kind of vacuum energy.

2.3 The $(\Omega_M, \Omega_\Lambda)$ plane

Now that we know that the universe is accelerating, one can parametrize the matter/energy content of the universe with just two components: the matter, characterized by Ω_M, and the vacuum energy Ω_Λ. Different values of these two parameters completely specify the universe evolution. It is thus natural to plot the results of observations in the plane $(\Omega_M, \Omega_\Lambda)$, in order to check whether we arrive at a consistent picture of the present universe from several different angles (different sets of cosmological observations).

Moreover, different regions of this plane specify different behaviors of the universe. The boundaries between regions are well defined curves that can be computed for a given model. I will now describe the various regions and boundaries.

- *Uniform expansion* ($q_0 = 0$). Corresponds to the line $\Omega_\Lambda = \Omega_M/2$. Points above this line correspond to universes that are accelerating today, while those below correspond to decelerating universes, in particular the old cosmological model of Einstein-de Sitter (EdS), with $\Omega_\Lambda = 0$, $\Omega_M = 1$. Since 1998, all the data from Supernovae of type Ia appear above this line, many standard deviations away from EdS universes.
- *Flat universe* ($\Omega_K = 0$). Corresponds to the line $\Omega_\Lambda = 1 - \Omega_M$. Points to the right of this line correspond to closed universes, while those to the left correspond to open ones. In the last few years we have mounting evidence that the universe is spatially flat (in fact Euclidean).
- *Bounce* ($t_0 H_0 = \infty$). Corresponds to a complicated function of $\Omega_\Lambda(\Omega_M)$, normally expressed as an integral equation, where

$$t_0 H_0 = \int_0^1 da \, [1 + \Omega_M(1/a - 1) + \Omega_\Lambda(a^2 - 1)]^{-1/2}$$

is the product of the age of the universe and the present rate of expansion. Points above this line correspond to universes that have contracted in the past and have later rebounced. At present, these universes are ruled out by observations of galaxies and quasars at high redshift (up to $z = 10$).
- *Critical Universe* ($H = \dot{H} = 0$, $\ddot{H} > 0$). Corresponds to the boundary between eternal expansion in the future and recollapse. For $\Omega_M \leq 1$, it is simply the line $\Omega_\Lambda = 0$, but for $\Omega_M > 1$, it is a more complicated curve,

$$\Omega_\Lambda = 4\Omega_M \sin^3 \left[\frac{1}{3} \arcsin \left(\frac{\Omega_M - 1}{\Omega_M} \right) \right] \simeq \frac{4}{27} \frac{(\Omega_M - 1)^3}{\Omega_M^2}.$$

These critical solutions are asymptotic to the EdS model.

These boundaries, and the regions they delimit, can be seen in Fig. 1, together with the lines of equal $t_0 H_0$ values.

In summary, the basic cosmological parameters that are now been hunted by a host of cosmological observations are the following: the present rate of expansion H_0; the age of the universe t_0; the deceleration parameter q_0; the spatial curvature Ω_K; the matter content Ω_M; the vacuum energy Ω_Λ; the baryon density Ω_B; the neutrino density Ω_ν, and many other that characterize the perturbations responsible for the large scale structure (LSS) and the CMB anisotropies.

2.4 The accelerating universe

Let us first describe the effect that the expansion of the universe has on the objects that live in it. In the absence of other forces but those of gravity, the trajectory of a particle is given by general relativity in terms of the geodesic equation

$$\frac{du^\mu}{ds} + \Gamma^\mu_{\nu\lambda} u^\nu u^\lambda = 0 \,, \tag{23}$$

where $u^\mu = (\gamma, \gamma v^i)$, with $\gamma^2 = 1 - v^2$ and v^i is the peculiar velocity. Here $\Gamma^\mu_{\nu\lambda}$ is the Christoffel connection [6], whose only non-zero component is $\Gamma^0_{ij} = (\dot{a}/a) \, g_{ij}$; substituting into the geodesic equation, we obtain $|\vec{u}| \propto 1/a$, and thus the particle's momentum decays with the expansion like $p \propto 1/a$. In the case of a photon, satisfying the de Broglie relation $p = h/\lambda$, one obtains the well known *photon redshift*

$$\frac{\lambda_1}{\lambda_0} = \frac{a(t_1)}{a(t_0)} \quad \Rightarrow \quad z \equiv \frac{\lambda_0 - \lambda_1}{\lambda_1} = \frac{a_0}{a_1} - 1 \,, \tag{24}$$

where λ_0 is the wavelength measured by an observer at time t_0, while λ_1 is the wavelength emitted when the universe was younger ($t_1 < t_0$). Normally we measure light from stars in distant galaxies and compare their observed spectra with our laboratory (restframe) spectra. The fraction (24) then gives the

redshift z of the object. We are assuming, of course, that both the emitted and the restframe spectra are identical, so that we can actually measure the effect of the intervening expansion, i.e. the growth of the scale factor from t_1 to t_0, when we compare the two spectra. Note that if the emitting galaxy and our own participated in the expansion, i.e. if our measuring rods (our rulers) also expanded with the universe, we would see no effect! The reason we can measure the redshift of light from a distant galaxy is because our galaxy is a gravitationally bounded object that has decoupled from the expansion of the universe. It is the distance between galaxies that changes with time, not the sizes of galaxies, nor the local measuring rods.

We can now evaluate the relationship between physical distance and redshift as a function of the rate of expansion of the universe. Because of homogeneity we can always choose our position to be at the origin $r = 0$ of our spatial section. Imagine an object (a star) emitting light at time t_1, at coordinate distance r_1 from the origin. Because of isotropy we can ignore the angular coordinates (θ, ϕ). Then the physical distance, to first order, will be $d = a_0 r_1$. Since light travels along null geodesics [6], we can write $0 = -dt^2 + a^2(t) \, dr^2 / (1 - Kr^2)$, and therefore,

$$\int_{t_1}^{t_0} \frac{dt}{a(t)} = \int_0^{r_1} \frac{dr}{\sqrt{1 - Kr^2}} \equiv f(r_1) = \begin{cases} \arcsin r_1 & K = 1 \\ r_1 & K = 0 \\ \text{arcsinh} \, r_1 & K = -1 \end{cases} \tag{25}$$

If we now Taylor expand the scale factor to first order,

$$\frac{1}{1+z} = \frac{a(t)}{a_0} = 1 + H_0(t - t_0) + \mathcal{O}(t - t_0)^2 \,, \tag{26}$$

we find, to first approximation,

$$r_1 \approx f(r_1) = \frac{1}{a_0}(t_0 - t_1) + \cdots = \frac{z}{a_0 H_0} + \dots$$

Putting all together we find the famous Hubble law

$$H_0 \, d = a_0 H_0 r_1 = z \simeq vc \,, \tag{27}$$

which is just a kinematical effect (we have not included yet any dynamics, i.e. the matter content of the universe). Note that at low redshift ($z \ll 1$), one is tempted to associate the observed change in wavelength with a Doppler effect due to a hypothetical recession velocity of the distant galaxy. This is only an approximation. In fact, the redshift cannot be ascribed to the relative velocity of the distant galaxy because in general relativity (i.e. in curved spacetimes) one cannot compare velocities through parallel transport, since the value depends on the path! If the distance to the galaxy is small, i.e. $z \ll 1$, the physical spacetime is not very different from Minkowsky and such a comparison is approximately valid. As z becomes of order one, such a relation is manifestly false: galaxies cannot travel at speeds greater than the speed of light; it is the stretching of spacetime which is responsible for the observed redshift.

Hubble's law has been confirmed by observations ever since the 1920s, with increasing precision, which have allowed cosmologists to determine the Hubble parameter H_0 with less and less systematic errors. Nowadays, the best determination of the Hubble parameter was made by the Hubble Space Telescope Key Project [8], $H_0 = 72 \pm 8$ km/s/Mpc. This determination is based on objects at distances up to 500 Mpc, corresponding to redshifts $z \leq 0.1$.

Nowadays, we are beginning to probe much greater distances, corresponding to $z \simeq 1$, thanks to type Ia supernovae. These are white dwarf stars at the end of their life cycle that accrete matter from a companion until they become unstable and violently explode in a natural thermonuclear explosion that out-shines their progenitor galaxy. The intensity of the distant flash varies in time, it takes about
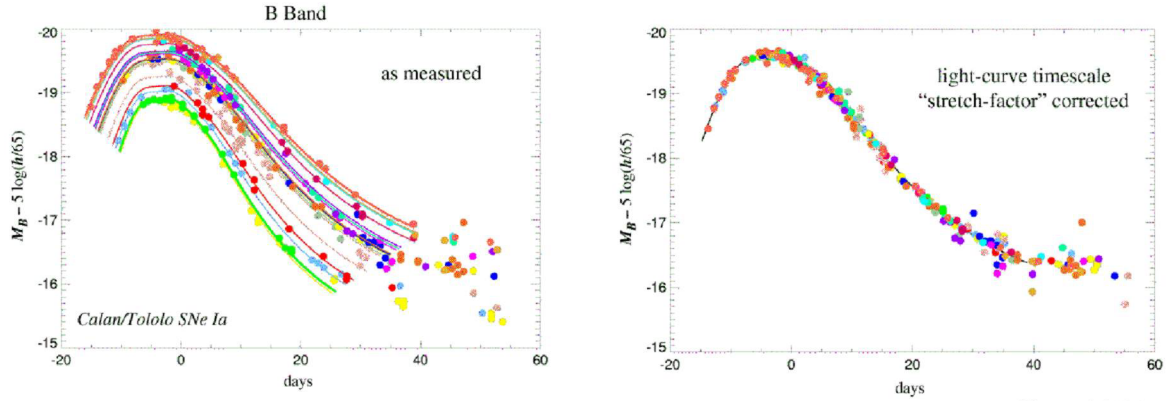
Fig. 2: The Type Ia supernovae observed nearby show a relationship between their absolute luminosity and the timescale of their light curve: the brighter supernovae are slower and the fainter ones are faster. A simple linear relation between the absolute magnitude and a "stretch factor" multiplying the light curve timescale fits the data quite well. From Ref. [7].

three weeks to reach its maximum brightness and then it declines over a period of months. Although the maximum luminosity varies from one supernova to another, depending on their original mass, their environment, etc., there is a pattern: brighter explosions last longer than fainter ones. By studying the characteristic light curves, see Fig. 2, of a reasonably large statistical sample, cosmologists from the Supernova Cosmology Project [7] and the High-redshift Supernova Project [9], are now quite confident that they can use this type of supernova as a standard candle. Since the light coming from some of these rare explosions has travelled a large fraction of the size of the universe, one expects to be able to infer from their distribution the spatial curvature and the rate of expansion of the universe.

The connection between observations of high redshift supernovae and cosmological parameters is done via the luminosity distance, defined as the distance d_L at which a source of absolute luminosity (energy emitted per unit time) $\mathcal{L}$ gives a flux (measured energy per unit time and unit area of the detector) $\mathcal{F} = \mathcal{L}/4\pi d_L^2$. One can then evaluate, within a given cosmological model, the expression for d_L as a function of redshift [10],

$$H_0 \, d_L(z) = \frac{(1+z)}{|\Omega_K|^{1/2}} \, \mathrm{sinn} \left[\int_0^z \frac{|\Omega_K|^{1/2} \, dz'}{\sqrt{(1+z')^2(1+z'\Omega_M) - z'(2+z')\Omega_\Lambda}} \right] , \qquad (28)$$

where $\mathrm{sinn}(x) = x$ if $K = 0$; $\sin(x)$ if $K = +1$ and $\sinh(x)$ if $K = -1$, and we have used the cosmic sum rule (20).

Astronomers measure the relative luminosity of a distant object in terms of what they call the effective magnitude, which has a peculiar relation with distance,

$$m(z) \equiv M + 5 \log_{10} \left[\frac{d_L(z)}{\mathrm{Mpc}} \right] + 25 = \bar{M} + 5 \log_{10}[H_0 \, d_L(z)] . \qquad (29)$$

Since 1998, several groups have obtained serious evidence that high redshift supernovae appear fainter than expected for either an open ($\Omega_M < 1$) or a flat ($\Omega_M = 1$) universe, see Fig. 3. In fact, the universe appears to be accelerating instead of decelerating, as was expected from the general attraction of matter, see Eq. (22); something seems to be acting as a repulsive force on very large scales. The most natural explanation for this is the presence of a cosmological constant, a diffuse vacuum energy that permeates all space and, as explained above, gives the universe an acceleration that tends to separate gravitationally bound systems from each other. The best-fit results from the Supernova Cosmology Project [11] give a
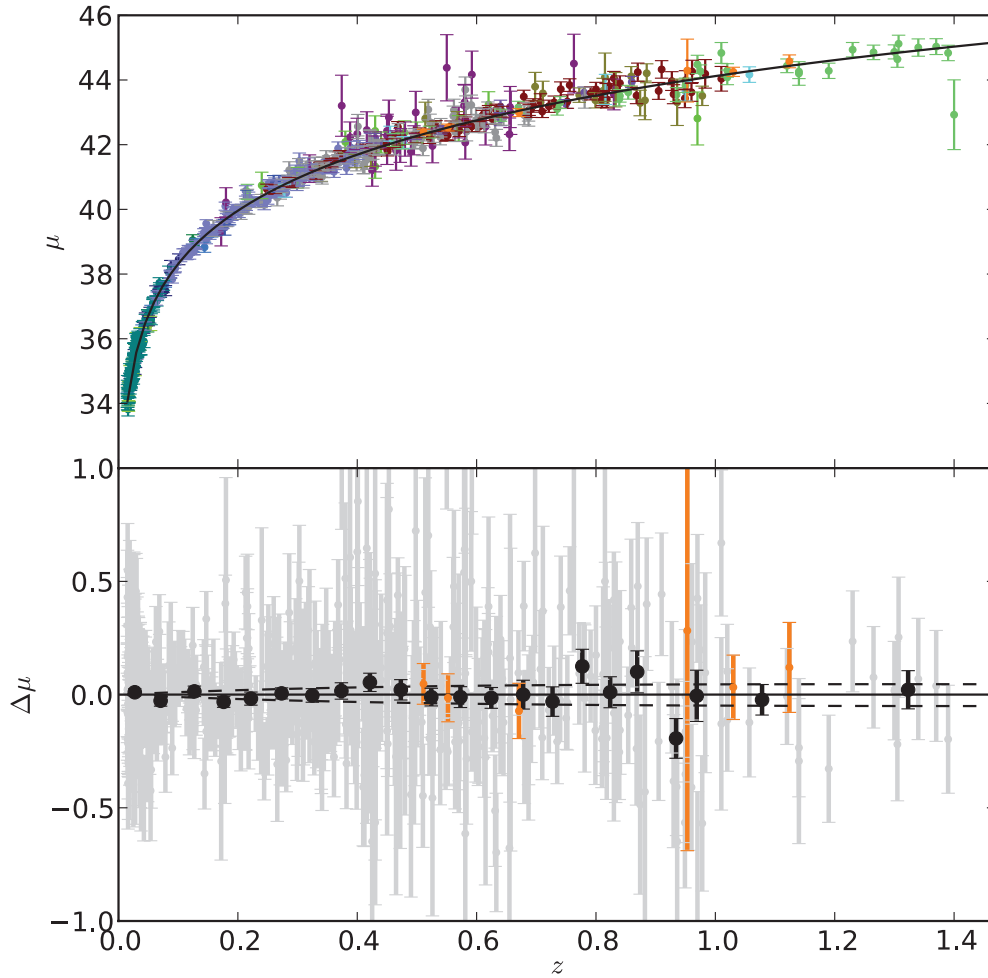
Fig. 3: The Hubble diagram in linear redshift scale. Supernovae with $\Delta z < 0.05$ of eachother have been weighted-averaged binned. The solid curve represents the best-fit flat universe model, ($\Omega_M = 0.29$, $\Omega_\Lambda = 0.71$). Lower panel: Residuals of the averaged data relative to a ΛCDM universe with fiducial cosmology. Dashed lines correspond to $w = -1 \pm 0.1$. From Ref. [13].

linear combination

$$0.8\,\Omega_M - 0.6\,\Omega_\Lambda = -0.16 \pm 0.05 \ \ (1\sigma),$$

which is now many sigma away from an EdS model with $\Lambda = 0$. In particular, for a flat universe this gives

$$\Omega_\Lambda = 0.71 \pm 0.05 \qquad \text{and} \qquad \Omega_M = 0.29 \pm 0.05 \ \ (1\sigma).$$

Surprising as it may seem, arguments for a significant dark energy component of the universe where proposed long before these observations, in order to accommodate the ages of globular clusters, as well as a flat universe with a matter content below critical, which was needed in order to explain the observed distribution of galaxies, clusters and voids.
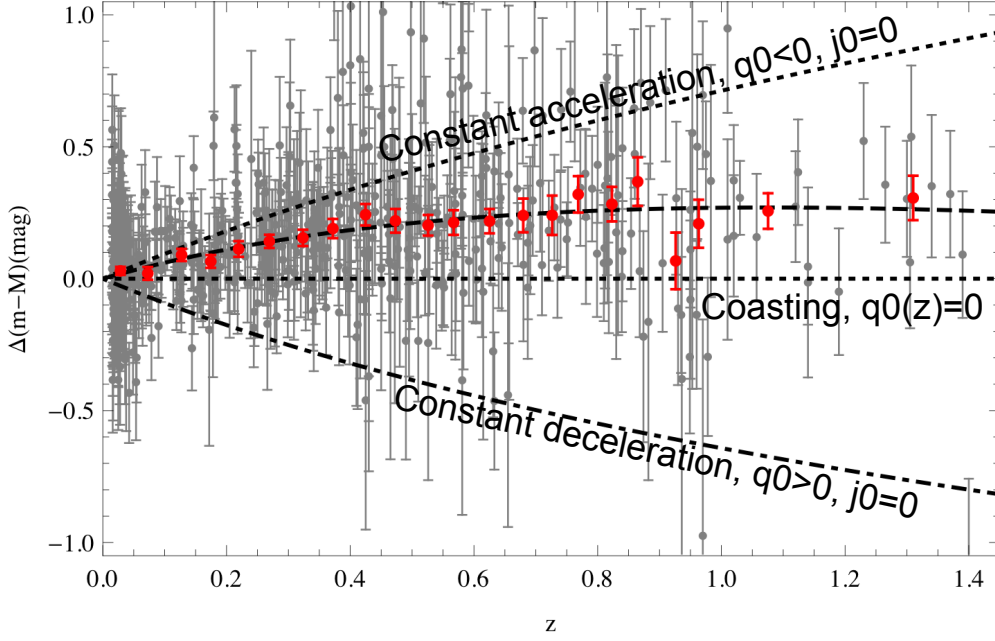
Fig. 4: The Supernovae Ia residual Hubble diagram. Upper panel: Ground-based discoveries are represented by diamonds, HST-discovered SNe Ia are shown as filled circles. Lower panel: The same but with weighted averaged in fixed redshift bins. Kinematic models of the expansion history are shown relative to an eternally coasting model $q(z) = 0$. Adapted from Ref. [13].

Taylor expanding the scale factor to third order,

$$\frac{a(t)}{a_0} = 1 + H_0(t - t_0) - \frac{q_0}{2!} H_0^2 (t - t_0)^2 + \frac{j_0}{3!} H_0^3 (t - t_0)^3 + \mathcal{O}(t - t_0)^4 \,, \tag{30}$$

where

$$q_0 = -\frac{\ddot{a}}{aH^2}(t_0) = \frac{1}{2} \sum_i (1 + 3w_i)\Omega_i = \frac{1}{2}\Omega_M - \Omega_\Lambda \,, \tag{31}$$

$$j_0 = +\frac{\dddot{a}}{aH^3}(t_0) = \frac{1}{2} \sum_i (1 + 3w_i)(2 + 3w_i)\Omega_i = \Omega_M + \Omega_\Lambda \,, \tag{32}$$

are the deceleration and "jerk" parameters. Substituting into Eq. (28) we find

$$H_0 \, d_L(z) = z + \frac{1}{2}(1 - q_0)\, z^2 - \frac{1}{6}(1 - q_0 - 3q_0^2 + j_0)\, z^3 + \mathcal{O}(z^4) \,. \tag{33}$$

This expression goes beyond the leading linear term, corresponding to the Hubble law, into the second and third order terms, which are sensitive to the cosmological parameters Ω_M and Ω_Λ. It is only recently that cosmological observations have gone far enough back into the early universe that we can begin to probe these terms, see Fig. 4.

This extra component of the critical density would have to resist gravitational collapse, otherwise it would have been detected already as part of the energy in the halos of galaxies. However, if most of the energy of the universe resists gravitational collapse, it is impossible for structure in the universe to grow. This dilemma can be resolved if the hypothetical dark energy was negligible in the past and only recently became the dominant component. According to general relativity, this requires that the dark energy have

negative pressure, since the ratio of dark energy to matter density goes like $a(t)^{-3p/\rho}$. This argument would rule out almost all of the usual suspects, such as cold dark matter, neutrinos, radiation, and kinetic energy, since they all have zero or positive pressure. Thus, we expect something like a cosmological constant, with a negative pressure, $p \approx -\rho$, to account for the missing energy.

However, if the universe was dominated by dark matter in the past, in order to form structure, and only recently became dominated by dark energy, we must be able to see the effects of the transition from the deceleration into the acceleration phase in the luminosity of distant type Ia supernovae. This has been searched for since 1998, when the first convincing results on the present acceleration appeared. However, only recently [12] do we have clear evidence of this transition point in the evolution of the universe. This *coasting point* is defined as the time, or redshift, at which the deceleration parameter vanishes,

$$q(z) = -1 + \frac{d \ln H(z)}{d \ln(1+z)} = 0 \,, \tag{34}$$

where

$$H(z) = H_0 \Big[\Omega_M (1+z)^3 + \Omega_x \, e^{3 \int_0^z (1+w(z')) \frac{dz'}{1+z'}} + \Omega_K (1+z)^2 \Big]^{1/2} \,, \tag{35}$$

and we have assumed that the dark energy is parametrized by a density Ω_x today, with a redshift-dependent equation of state, $w(z)$, not necessarily equal to -1. Of course, in the case of a true cosmological constant, this reduces to the usual expression.

Let us suppose for a moment that the barotropic parameter w is constant, then the coasting redshift can be determined from

$$q(z) = \frac{1}{2} \Big[\frac{\Omega_M + (1+3w) \, \Omega_x \, (1+z)^{3w}}{\Omega_M + \Omega_x \, (1+z)^{3w} + \Omega_K (1+z)^{-1}} \Big] = 0 \,, \tag{36}$$

$$\Rightarrow z_c = \Big(\frac{(3|w|-1)\Omega_x}{\Omega_M} \Big)^{\frac{1}{3|w|}} - 1 \,, \tag{37}$$

which, in the case of a true cosmological constant, reduces to

$$z_c = \Big(\frac{2\Omega_\Lambda}{\Omega_M} \Big)^{1/3} - 1 \,. \tag{38}$$

When substituting $\Omega_\Lambda \simeq 0.71$ and $\Omega_M \simeq 0.29$, one obtains $z_c \simeq 0.7$, in excellent agreement with recent observations [13].

Now, if we have to live with this vacuum energy, we might as well try to understand its origin. For the moment it is a complete mystery, perhaps the biggest mystery we have in physics today [14]. We measure its value but we don't understand why it has the value it has. In fact, if we naively predict it using the rules of quantum mechanics, we find a number that is many (many!) orders of magnitude off the mark. Let us describe this calculation in some detail. In non-gravitational physics, the zero-point energy of the system is irrelevant because forces arise from gradients of potential energies. However, we know from general relativity that even a constant energy density gravitates. Let us write down the most general energy momentum tensor compatible with the symmetries of the metric and that is covariantly conserved. This is precisely of the form $T_{\mu\nu}^{(vac)} = p_V \, g_{\mu\nu} = -\rho_V \, g_{\mu\nu}$, see Fig. 5. Substituting into the Einstein equations (1), we see that the cosmological constant and the vacuum energy are completely equivalent, $\Lambda = 8\pi G \, \rho_V$, so we can measure the vacuum energy with the observations of the acceleration of the universe, which tells us that $\Omega_\Lambda \simeq 0.7$.

On the other hand, we can estimate the contribution to the vacuum energy coming from the quantum mechanical zero-point energy of the quantum oscillators associated with the fluctuations of all quantum fields,

$$\rho_V^{th} = \sum_i \int_0^{\Lambda_{UV}} \frac{d^3 k}{(2\pi)^3} \frac{1}{2} \hbar \omega_i(k) = \frac{\hbar \Lambda_{UV}^4}{16\pi^2} \sum_i (-1)^F N_i + \mathcal{O}(m_i^2 \Lambda_{UV}^2) \,, \tag{39}$$

where Λ_{UV} is the ultraviolet cutoff signaling the scale of new physics. Taking the scale of quantum gravity, $\Lambda_{UV} = M_{Pl}$, as the cutoff, and barring any fortuituous cancellations, then the theoretical expectation (39) appears to be 120 orders of magnitude larger than the observed vacuum energy associated with the acceleration of the universe,

$$\rho_V^{th} \simeq 1.4 \times 10^{74} \text{ GeV}^4 = 3.2 \times 10^{91} \text{ g/cm}^3 \,, \tag{40}$$

$$\rho_V^{obs} \simeq 0.7 \, \rho_c = 0.66 \times 10^{-29} \text{ g/cm}^3 = 2.9 \times 10^{-11} \text{ eV}^4 \,. \tag{41}$$

Even if we assumed that the ultraviolet cutoff associated with quantum gravity was as low as the electroweak scale, the theoretical expectation would still be 60 orders of magnitude too big. This is by far the worst mismatch between theory and observations in all of science. There must be something seriously wrong in our present understanding of gravity at the most fundamental level. Perhaps we don't understand the vacuum and its energy does not gravitate after all, or perhaps we need to impose a new principle (or a symmetry) at the quantum gravity level to accommodate such a flagrant mismatch.
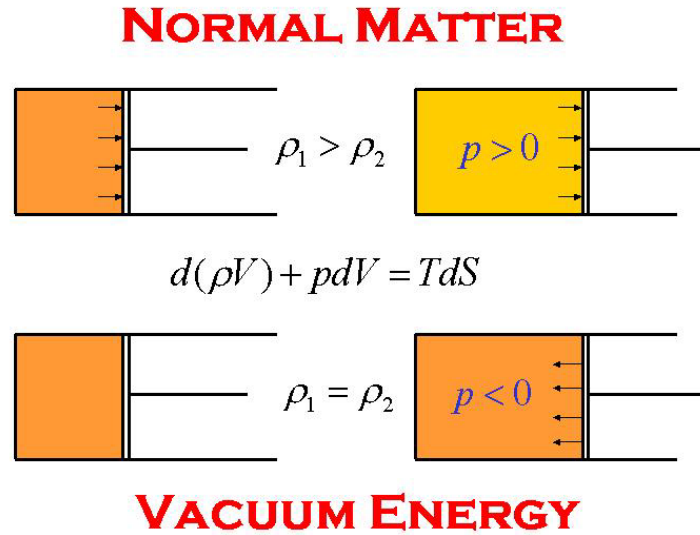
Fig. 5: Ordinary matter dilutes as it expands. According to the second law of Thermodynamics, its pressure on the walls should be positive, which excerts a force, and energy is lost in the expansion. On the other hand, vacuum energy is always the same, independent of the volume of the region, and thus, according to the second law, its pressure must be negative and of the same magnitude as the energy density. This negative pressure means that the volume tends to increase more and more rapidly, which explains the exponential expansion of the universe dominated by a cosmological constant.

In the meantime, one can at least parametrize our ignorance by making variations on the idea of a *constant* vacuum energy. Let us assume that it actually evolves slowly with time. In that case, we do not expect the equation of state $p = -\rho$ to remain true, but instead we expect the barotropic parameter $w(z)$ to depend on redshift. Such phenomenological models have been proposed, and until recently produced results that were compatible with $w = -1$ today, but with enough uncertainty to speculate on alternatives to a truly constant vacuum energy. However, with the recent supernovae results [12], there seems to be little space for variations, and models of a time-dependent vacuum energy are less and less favoured. In the near future, the SNAP satellite [15] will measure several thousand supernovae at high redshift and therefore map the redshift dependence of both the dark energy density and its equation of state with great precision. This will allow a much better determination of the cosmological parameters Ω_M and Ω_Λ.

2.5 Thermodynamics of an expanding plasma

In this section I will describe the main concepts associated with ensembles of particles in thermal equilibrium and the brief periods in which the universe fell out of equilibrium. To begin with, let me make contact between the covariant energy conservation law (7) and the second law of thermodynamics,

$$T \, dS = dU + p \, dV \,, \tag{42}$$

where $U = \rho \, V$ is the total energy of the fluid, and $p = w \, \rho$ is its barotropic pressure. Taking a comoving volume for the universe, $V = a^3$, we find

$$T \frac{dS}{dt} = \frac{d}{dt}(\rho \, a^3) + p \frac{d}{dt}(a^3) = 0 \,, \tag{43}$$

where we have used (7). Therefore, entropy is conserved during the expansion of the universe, $dS = 0$; i.e., the expansion is adiabatic even in those epochs in which the equation of state changes, like in the matter-radiation transition (not a proper phase transition). Using (7), we can write

$$\frac{d}{dt} \ln(\rho \, a^3) = -3H \, w \,. \tag{44}$$

Thus, our universe expands like a gaseous fluid in thermal equilibrium at a temperature T. This temperature decreases like that of any expanding fluid, in a way that is inversely proportional to the cubic root of the volume. This implies that in the past the universe was necessarily denser and hotter. As we go back in time we reach higher and higher temperatures, which implies that the mean energy of plasma particles is larger and thus certain fundamental reactions are now possible and even common, giving rise to processes that today we can only attain in particle physics accelerators. That is the reason why it is so important, for the study of early universe, to know the nature of the fundamental interactions at high energies, and the basic connection between cosmology and high energy particle physics. However, I should clarify a misleading statement that is often used: "high energy particle physics colliders reproduce the early universe" by inducing collisions among relativistic particles. Although the energies of some of the interactions at those collisions reach similar values as those attained in the early universe, the physical conditions are rather different. The interactions within the detectors of the great particle physics accelerators occur typically in the perturbative regime, locally, and very far from thermal equilibrium, lasting a minute fraction of a second; on the other hand, the same interactions occurred within a hot plasma in equilibrium in the early universe while it was expanding adiabatically and its duration could be significantly larger, with a distribution in energy that has nothing to do with those associated with particle accelerators. What is true, of course, is that the fundamental parameters corresponding to those interactions − masses and couplings − are assumed to be the same, and therefore present terrestrial experiments can help us imagine what it could have been like in the early universe, and make predictions about the evolution of the universe, in the context of an expanding plasma a high temperatures and high densities, and in thermal equilibrium.

2.5.1 Fluids in thermal equilibrium

In order to understand the thermodynamical behaviour of a plasma of different species of particles at high temperatures we will consider a gas of particles with g internal degrees of freedom weakly interacting. The degrees of freedom corresponding to the different particles can be seen in Table 1. For example, leptons and quarks have 4 degrees of freedom since they correspond to the two helicities for both particle and antiparticle. However, the nature of neutrinos is still unknown. If they happen to be Majorana fermions, then they would be their own antiparticle and the number of degrees of freedom would reduce to 2. For photons and gravitons (without mass) their 2 d.o.f. correspond to their states of polarization. The 8 gluons (also without mass) are the gauge bosons responsible for the strong interaction betwen

Particle	Spin	Degrees of freedom (g)	Nature
Higgs	0	1	Massive scalar
photon	1	2	Massless vector
graviton	2	2	Massless tensor
gluon	1	2	Massless vector
W y Z	1	3	Massive vector
leptons & quarks	1/2	4	Dirac Fermion
neutrinos	1/2	4 (2)	Dirac (Majorana) Fermion

Table 1: The internal degrees of freedom of various fundamental particles.

quarks, and also have 2 d.o.f. each. The vector bosons $W^{\pm}$ and Z^0 are massive and thus, apart from the transverse components of the polarization, they also have longitudinal components.

For each of these particles we can compute the number density n, the energy density ρ and the pressure p, in thermal equilibrium at a given temperature T,

$$n = g \int \frac{d^3\mathbf{p}}{(2\pi)^3} f(\mathbf{p}) \,, \tag{45}$$

$$\rho = g \int \frac{d^3\mathbf{p}}{(2\pi)^3} E(\mathbf{p}) f(\mathbf{p}) \,, \tag{46}$$

$$p = g \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{|\mathbf{p}|^2}{3E} f(\mathbf{p}) \,, \tag{47}$$

where the energy is given by $E^2 = |\mathbf{p}|^2 + m^2$ and the momentum distribution in thermal (kinetic) equilibrium is

$$f(\mathbf{p}) = \frac{1}{e^{(E-\mu)/T} \pm 1} \qquad \begin{cases} -1 & \text{Bose} - \text{Einstein} \\ +1 & \text{Fermi} - \text{Dirac} \end{cases} \tag{48}$$

The chemical potential μ is conserved in these reactions if they are in thermal equilibrium. For example, for reactions of the type $i + j \longleftrightarrow k + l$, we have $\mu_i + \mu_j = \mu_k + \mu_l$. For example, the chemical potencial of the photon vanishes $\mu_\gamma = 0$, and thus particles and antiparticles have opposite chemical potentials.

From the equilibrium distributions one can obtain the number density n, the energy ρ and the pressure p, of a particle of mass m with chemical potential μ at the temperature T,

$$n = \frac{g}{2\pi^2} \int_m^\infty dE \, \frac{E(E^2 - m^2)^{1/2}}{e^{(E-\mu)/T} \pm 1} \,, \tag{49}$$

$$\rho = \frac{g}{2\pi^2} \int_m^\infty dE \, \frac{E^2(E^2 - m^2)^{1/2}}{e^{(E-\mu)/T} \pm 1} \,, \tag{50}$$

$$p = \frac{g}{6\pi^2} \int_m^\infty dE \, \frac{(E^2 - m^2)^{3/2}}{e^{(E-\mu)/T} \pm 1} \,. \tag{51}$$

For a non-degenerate ($\mu \ll T$) relativistic gas ($m \ll T$), we find

$$n = \frac{g}{2\pi^2} \int_0^\infty \frac{E^2 \, dE}{e^{E/T} \pm 1} = \begin{cases} \dfrac{\zeta(3)}{\pi^2} \, g \, T^3 & \text{Bosons} \\[2mm] \dfrac{3}{4} \dfrac{\zeta(3)}{\pi^2} \, g \, T^3 & \text{Fermions} \end{cases} \,, \tag{52}$$

$$\rho = \frac{g}{2\pi^2} \int_0^\infty \frac{E^3 \, dE}{e^{E/T} \pm 1} = \begin{cases} \dfrac{\pi^2}{30} g \, T^4 & \text{Bosons} \\[2ex] \dfrac{7}{8} \dfrac{\pi^2}{30} g \, T^4 & \text{Fermions} \end{cases}, \tag{53}$$

$$p = \frac{1}{3}\rho, \tag{54}$$

where $\zeta(3) = 1.20206\ldots$ is the Riemann Zeta function. For relativistic fluids, the energy density per particle is

$$\langle E \rangle \equiv \frac{\rho}{n} = \begin{cases} \dfrac{\pi^4}{30\zeta(3)} T \simeq 2.701 \, T & \text{Bosons} \\[2ex] \dfrac{7\pi^4}{180\zeta(3)} T \simeq 3.151 \, T & \text{Fermions} \end{cases} \tag{55}$$

For relativistic bosons or fermions with $\mu < 0$ and $|\mu| < T$, we have

$$n = \frac{g}{\pi^2} T^3 \, e^{\mu/T}, \tag{56}$$

$$\rho = \frac{3g}{\pi^2} T^4 \, e^{\mu/T}, \tag{57}$$

$$p = \frac{1}{3}\rho. \tag{58}$$

For a bosonic particle, a positive chemical potential, $\mu > 0$, indicates the presence of a Bose-Einstein condensate, and should be treated separately from the rest of the modes.

On the other hand, for a non-relativistic gas ($m \gg T$), with arbitrary chemical potential μ, we find

$$n = g \left(\frac{mT}{2\pi}\right)^{3/2} e^{-(m-\mu)/T}, \tag{59}$$

$$\rho = m \, n, \tag{60}$$

$$p = n \, T \ll \rho. \tag{61}$$

The average energy density per particle is

$$\langle E \rangle \equiv \frac{\rho}{n} = m + \frac{3}{2} T. \tag{62}$$

Note that, at any given temperature T, the contribution to the energy density of the universe coming from non-relativistic particles in thermal equilibrium is exponentially suppressed with respect to that of relativistic particles, therefore we can write

$$\rho_{\rm R} = \frac{\pi^2}{30} g_* \, T^4, \qquad p_{\rm R} = \frac{1}{3}\rho_{\rm R}, \tag{63}$$

$$g_*(T) = \sum_{\text{bosons}} g_i \left(\frac{T_i}{T}\right)^4 + \frac{7}{8} \sum_{\text{fermions}} g_i \left(\frac{T_i}{T}\right)^4, \tag{64}$$

where the factor $7/8$ takes into account the difference between the Fermi and Bose statistics; g_* is the total number of light d.o.f. ($m \ll T$), and we have also considered the possibility that particle species i (bosons or fermions) have an equilibrium distribution at a temperature T_i different from that of photons, as happens for example when a given relativistic species decouples from the thermal bath, as we will discuss later. This number, g_*, strongly depends on the temperature of the universe, since as it expands

and cools, different particles go out of equilibrium or become non-relativistic ($m \gg T$) and thus become exponentially suppressed from that moment on. A plot of the time evolution of $g_*(T)$ can be seen in Fig. 6. For example, for $T \ll 1$ MeV, i.e. after the time of primordial Big Bang Nucleosynthesis (BBN) and neutrino decoupling, the only relativistic species are the 3 light neutrinos and the photons; since the temperature of the neutrinos is $T_\nu = (4/11)^{1/3} T_\gamma = 1.90$ K, see below, we have $g_* = 2 + 3 \times \frac{7}{4} \times \left(\frac{4}{11}\right)^{4/3} = 3.36$, while $g_{*S} = 2 + 3 \times \frac{7}{4} \times \left(\frac{4}{11}\right) = 3.91$.
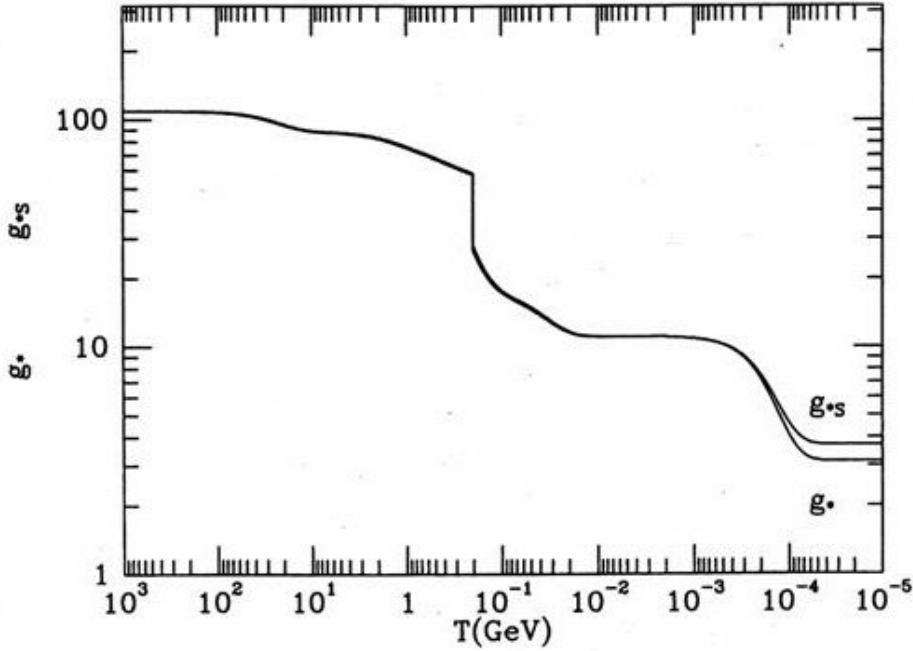
Fig. 6: The light degrees of freedom g_* and g_{*S} as a function of the temperature of the universe. From Ref. [16].

For 1 MeV $\ll T \ll$ 100 MeV, i.e. between BBN and the phase transition from a quark-gluon plasma to hadrons and mesons, we have, as relativistic species, apart from neutrinos and photons, also the electrons and positrons, so $g_* = 2 + 3 \times \frac{7}{4} + 2 \times \frac{7}{4} = 10.75$.

For $T \gg 250$ GeV, i.e. above the electroweak (EW) symmetry breaking scale, we have one photon (2 polarizations), 8 gluons (massless), the $W^\pm$ and Z^0 (massive), 3 families of quarks & leptons, a Higgs (still undiscovered), with which one finds $g_* = \frac{427}{4} = 106.75$.

At temperatures well above the electroweak transition we ignore the number of d.o.f. of particles, since we have never explored those energies in particle physics accelerators. Perhaps in the near future, with the results of the Large Hadron Collider (LHC) at CERN, we may may predict the behaviour of the universe at those energy scales. For the moment we even ignore whether the universe was in thermal equilibrium at those temperatures. The highest energy scale at which we can safely say the universe was in thermal equilibrium is that of BBN, i.e. 1 MeV, due to the fact that we observe the present relative abundances of the light element produced at that time. For instance, we can't even claim that the universe went through the quark-gluon phase transition, at ~ 200 MeV, since we have not observed yet any signature of such an event, not to mention the electroweak phase transition, at ~ 1 TeV.

Let us now use the relation between the rate of expansion and the temperature of relativistic particles to obtain the time scale of the universe as a function of its temperature,

$$H = 1.66 \, g_*^{1/2} \frac{T^2}{M_P} = \frac{1}{2t} \qquad \Longrightarrow \qquad t = 0.301 \, g_*^{-1/2} \frac{M_P}{T^2} \sim \left(\frac{T}{\text{MeV}}\right)^{-2} \text{s} , \qquad (65)$$

thus, e.g. at the EW scale (100 GeV) the universe was just 10^{-10} s old, while during the primordial BBN $(1 - 0.1 \text{ MeV})$, it was 1 s to 3 min old.

2.5.2 *The entropy of the universe*

During most of the history of the universe, the rates of reaction, Γ_{int}, of particles in the thermal bath are much bigger than the rate of expansion of the universe, H, so that local thermal equilibrium was mantained. In this case, the entropy per comoving volume remained constant. In an expanding universe, the second law of thermodynamics, applied to the element of comoving volume, of unit coordinate volume and physical volume $V = a^3$, can be written as, see (42),

$$T\,dS = d(\rho\,V) + p\,dV = d[(\rho + p)V] - V\,dp\,. \tag{66}$$

Using the Maxwell condition of integrability, $\dfrac{\partial^2 S}{\partial T\partial V} = \dfrac{\partial^2 S}{\partial V\partial T}$, we find that $dp = (\rho + p)dT/T$, so that

$$dS = d\left[(\rho + p)\frac{V}{T} + \text{const.}\right], \tag{67}$$

i.e. the entropy in a comoving volume is $S = (\rho + p)V/T$, except for a constant. Using now the first law, the covariant conservation of energy $T^{\mu\nu}_{\ ;\nu} = 0$, we have

$$d\Big[(\rho + p)a^3\Big] = a^3\,dp \qquad \Longrightarrow \qquad d\Big((\rho + p)\frac{a^3}{T}\Big) = 0\,, \tag{68}$$

and thus, in thermal equilibrium, the total entropy in a comoving volume, $S = a^3(\rho + p)/T$, is conserved. During most of the evolution of the universe, this entropy was dominated by the contribution from relativistic particles,

$$S = \frac{2\pi^2}{45}\,g_{*S}\,(aT)^3 = \text{const.}\,, \tag{69}$$

$$g_{*S}(T) = \sum_{\text{bosons}} g_i\left(\frac{T_i}{T}\right)^3 + \frac{7}{8}\sum_{\text{fermions}} g_i\left(\frac{T_i}{T}\right)^3\,, \tag{70}$$

where g_{*S} is the number of "entropic" degrees of freedom, as we can see in Fig. 8. Above the electron-positron annihilation, all relativistic particles had the same temperature and thus $g_{*S} = g_*$. It may be also useful to realize that the entropy density, $s = S/a^3$, is propotional to the number density of relativistic particles, and in particular to the number density of photons, $s = 1.80\,g_{*S}\,n_\gamma$; today, $s = 7.04\,n_\gamma$. However, since g_{*S} in general is a function of temperature, we can't always interchange s and n_γ.

The conservation of S implies that the entropy density satisfies $s \propto a^{-3}$, and thus the physical size of the comoving volume is $a^3 \propto s^{-1}$; therefore, the number of particles of a given species in a comoving volume, $N = a^3 n$, is proportional to the number density of that species over the entropy density s,

$$N \sim \frac{n}{s} = \begin{cases} \dfrac{45\zeta(3)\,g}{2\pi^4\,g_{*S}} & T \gg m,\,\mu \\[4mm] \dfrac{45\,g}{4\pi^5\sqrt{2}\,g_{*S}}\left(\dfrac{m}{T}\right)^{3/2} e^{-\frac{m-\mu}{T}} & T \ll m \end{cases} \tag{71}$$

If this number does not change, i.e. if those particles are neither created nor destroyed, then n/s remains constant. As a useful example, we will consider the barionic number in a comoving volume,

$$\frac{n_B}{s} \equiv \frac{n_b - n_{\bar{b}}}{s}\,. \tag{72}$$

As long as the interactions that violate barion number occur sufficiently slowly, the barionic number per comoving volume, n_B/s, will remain constant. Although

$$\eta \equiv \frac{n_B}{n_\gamma} = 1.80\, g_{*S}\, \frac{n_B}{s}\,, \tag{73}$$

the ratio between barion and photon numbers it does not remain constant during the whole evolution of the universe since g_{*S} varies; e.g. during the annihilation of electrons and positrons, the number of photons per comoving volume, $N_\gamma = a^3\, n_\gamma$, grows a factor $11/4$, and η decreases by the same factor. After this epoch, however, g_* is constant so that $\eta \simeq 7 n_B/s$ and n_B/s can be used indistinctly.

Another consequence of Eq, (69) is that $S = \text{const.}$ implies that the temperature of the universe evolves as

$$T \propto g_{*S}^{-1/3}\, a^{-1}\,. \tag{74}$$

As long as g_{*S} remains constant, we recover the well known result that the universe cools as it expands according to $T \propto 1/a$. The factor $g_{*S}^{-1/3}$ appears because when a species becomes non-relativistic (when $T \leq m$), and effectively disappears from the energy density of the universe, its entropy is transferred to the rest of the relativistic particles in the plasma, making T decrease not as quickly as $1/a$, until g_{*S} again becomes constant.

From the observational fact that the universe expands today one can deduce that in the past it must have been hotter and denser, and that in the future it will be colder and more dilute. Since the ratio of scale factors is determined by the redshift parameter z, we can obtain (to very good approximation) the temperature of the universe in the past with

$$T = T_0\,(1+z)\,. \tag{75}$$

This expression has been spectacularly confirmed thanks to the absorption spectra of distant quasars [17]. These spectra suggest that the radiation background was acting as a thermal bath for the molecules in the interstellar medium with a temperature of 9 K at a redshift $z \sim 2$, and thus that in the past the photon background was hotter than today. Furthermore, observations of the anisotropies in the microwave background confirm that the universe at a redshift $z = 1089$ had a temperature of 0.3 eV, in agreement with Eq. (75).

2.6 The thermal evolution of the universe

In a strict mathematical sense, it is impossible for the universe to have been always in thermal equilibrium since the FRW model does not have a timelike Killing vector. In practice, however, we can say that the universe has been most of its history very close to thermal equilibrium. Of course, those periods in which there were deviations from thermal equilibrium have been crucial for its evolution thereafter (e.g. baryogenesis, QCD transition, primordial nucleosynthesis, recombination, etc.); without these the universe today would be very different and probably we would not be here to tell the story.

The key to understand the thermal history of the universe is the comparison between the rates of interaction between particles (microphysics) and the rate of expansion of the universe (macrophysics). Ignoring for the moment the dependence of g_* on temperature, the rate of change of T is given directly by the rate of expansion, $\dot{T}/T = -H$. As long as the local interactions − necessary in order that the particle distribution function adjusts *adiabatically* to the change of temperature − are sufficiently fast compared with the rate of expansion of the universe, the latter will evolve as a succession of states very close to thermal equilibrium, with a temperature proportional to a^{-1}. If we evaluate the interaction rates as

$$\Gamma_{\text{int}} \equiv \langle n\,\sigma\,|v| \rangle\,, \tag{76}$$

where $n(t)$ is the number density of target particles, σ is the cross section on the interaction and v is the relative velocity of the reaction, all averaged on a thermal distribution; then a rule of thumb for ensuring

that thermal equilibrium is maintained is

$$\Gamma_{\text{int}} \gtrsim H \,. \tag{77}$$

This criterium is understandable. Suppose, as often occurs, that the interaction rate in thermal equilibrium is $\Gamma_{\text{int}} \propto T^n$, with $n > 2$; then, the number of interactions of a particle after time t is

$$N_{\text{int}} = \int_t^\infty \Gamma_{\text{int}}(t')dt' = \frac{1}{n-2} \frac{\Gamma_{\text{int}}}{H}(t)\,, \tag{78}$$

therefore the particle interacts less than once from the moment in which $\Gamma_{\text{int}} \approx H$. If $\Gamma_{\text{int}} \gtrsim H$, the species remains coupled to the thermal plasma. This doesn't mean that, necessarily, the particle is out of local thermal equilibrium, since we have seen already that relativistic particles that have decoupled retain their equilibrium distribution, only at a different temperature from that of the rest of the plasma.

In order to obtain an approximate description of the decoupling of a particle species in an expanding universe, let us consider two types of interaction:

i) interactions mediated by massless gauge bosons, like for example the photon. In this case, the cross section for particles with significant momentum transfer can be written as $\sigma \sim \alpha^2/T^2$, with $\alpha = g^2/4\pi$ the coupling constant of the interaction. Assuming local thermal equilibrium, $n(t) \sim T^3$ and thus the interaction rate becomes $\Gamma \sim n\,\sigma\,|v| \sim \alpha^2\,T$. Therefore,

$$\frac{\Gamma}{H} \sim \alpha^2 \frac{M_P}{T}\,, \tag{79}$$

so that for temperatures of the universe $T \lesssim \alpha^2\,M_P \sim 10^{16}$ GeV, the reactions are fast enough and the plasma is in equilibrium, while for $T \gtrsim 10^{16}$ GeV, reactions are too slow to maintain equilibrium and it is said that they are "frozen-out". An important consequence of this result is that the universe could never have been in thermal equilibrium above the grand unification (GUT) scale.

ii) interactions mediated by massive gauge bosons, e.g. like the $W^\pm$ and Z^0, or those responsible for the GUT interactions, X and Y. We will generically call them X bosons. The cross section depends rather strongly on the temperature of the plasma,

$$\sigma \sim \begin{cases} G_X^2 T^2 & T \ll M_X \\[2mm] \dfrac{\alpha^2}{T^2} & T \gg M_X \end{cases} \tag{80}$$

where $G_X \sim \alpha/M_X^2$ is the effective coupling constant of the interaction at energies well below the mass of the vector boson, analogous to the Fermi constant of the electroweak interaction, $G_F = g^2/(4\sqrt{2}M_W^2)$ at tree level. Note that for $T \gg M_X$ we recover the result for massless bosons, so we will concentrate here on the other case. For $T \leq M_X$, the rate of thermal interactions is $\Gamma \sim n\,\sigma\,|v| \sim G_X^2\,T^5$. Therefore,

$$\frac{\Gamma}{H} \sim G_X^2\,M_P\,T^3\,, \tag{81}$$

such that at temperatures in the range

$$M_X \gtrsim T \gtrsim G_X^{-2/3} M_P^{-1/3} \sim \left(\frac{M_X}{100\text{ GeV}}\right)^{4/3} \text{MeV}\,, \tag{82}$$

reactions occur so fast that the plasma is in thermal equilibrium, while for $T \lesssim (M_X/100\text{ GeV})^{4/3}$ MeV, those reactions are too slow for maintaining equilibrium and they effective freeze-out, see Eq. (78).

2.6.1 The decoupling of relativistic particles

Those relativistic particles that have decoupled from the thermal bath do not participate in the transfer of entropy when the temperature of the universe falls below the mass thershold of a given species $T \simeq m$; in fact, the temperature of the decoupled relativistic species falls as $T \propto 1/a$, as we will now show. Suppose that a relativistic particle is initially in local thermal equilibrium, and that it decoples at a temperature T_D and time t_D. The phase space distribution at the time of decoupling is given by the equilibrium distribution,

$$f(\mathbf{p}, t_D) = \frac{1}{e^{E/T_D} \pm 1} . \tag{83}$$

After decoupling, the energy of each massless particle suffers redshift, $E(t) = E_D\,(a_D/a(t))$. The number density of particles also decreases, $n(t) = n_D\,(a_D/a(t))^3$. Thus, the phase space distribution at a time $t > t_D$ is

$$f(\mathbf{p}, t) = \frac{d^3 n}{d^3 \mathbf{p}} = f(\mathbf{p}\frac{a}{a_D}, t_D) = \frac{1}{e^{Ea/a_D T_D} \pm 1} = \frac{1}{e^{E/T} \pm 1} , \tag{84}$$

so that we conclude that the distribution function of a particle that has decoupled while being relativistic remains self-similar as the universe expands, with a temperature that decreases as

$$T = T_D\,\frac{a_D}{a} \propto a^{-1} , \tag{85}$$

and *not* as $g_{*S}^{-1/3}\,a^{-1}$, like the rest of the plasma in equilibrium (74).

2.6.2 The decoupling of non-relativistic particles

Those particles that decoupled from the thermal bath when they were non-relativistic ($m \gg T$) behave differently. Let us study the evolution of the distribution function of a non-relativistic particle that was in local thermal equilibrium at a time t_D, when the universe had a temperature T_D. The moment of each particle suffers redshift as the universe expands, $|\mathbf{p}| = |\mathbf{p}_D|\,(a_D/a)$, see Eq. (24). Therefore, their kinetic energy satisfies $E = E_D\,(a_D/a)^2$. On the other hand, the particle number density also varies, $n(t) = n_D\,(a_D/a(t))^3$, so that a decoupled non-relativistic particle will have an equilibrium distribution function characterized by a temperature

$$T = T_D\,\frac{a_D^2}{a^2} \propto a^{-2} , \tag{86}$$

and a chemical potential

$$\mu(t) = m + (\mu_D - m)\,\frac{T}{T_D} , \tag{87}$$

whose variation is precisely that which is needed for the number density of particle to decrease as a^{-3}.

In summary, a particle species that decouples from the thermal bath follows an equilibrium distribution function with a temperature that decreases like $T_R \propto a^{-1}$ for relativistic particles ($T_D \gg m$) or like $T_{NR} \propto a^{-2}$ for non-relativistic particles ($T_D \ll m$). On the other hand, for semi-relativistic particles ($T_D \sim m$), its phase space distribution *does not maintain* an equilibrium distribution function, and should be computed case by case.

2.6.3 Brief thermal history of the universe

I will briefly summarize here the thermal history of the universe, from the Planck era to the present. As we go back in time, the universe becomes hotter and hotter and thus the amount of energy available for particle interactions increases. As a consequence, the nature of interactions goes from those described at low energy by long range gravitational and electromagnetic physics, to atomic physics, nuclear physics,

all the way to high energy physics at the electroweak scale, gran unification (perhaps), and finally quantum gravity. The last two are still uncertain since we do not have any experimental evidence for those ultra high energy phenomena, and perhaps Nature has followed a different path.

The way we know about the high energy interactions of matter is via particle accelerators, which are unravelling the details of those fundamental interactions as we increase in energy. However, one should bear in mind that the physical conditions that take place in our high energy colliders are very different from those that occurred in the early universe. These machines could never reproduce the conditions of density and pressure in the rapidly expanding thermal plasma of the early universe. Nevertheless, those experiments are crucial in understanding the nature and *rate* of the local fundamental interactions available at those energies. What interests cosmologists is the statistical and thermal properties that such a plasma should have, and the role that causal horizons play in the final outcome of the early universe expansion. For instance, of crucial importance is the time at which certain particles *decoupled* from the plasma, i.e. when their interactions were not quick enough compared with the expansion of the universe, and they were left out of equilibrium with the plasma.

One can trace the evolution of the universe from its origin till today. There is still some speculation about the physics that took place in the universe above the energy scales probed by present colliders. Nevertheless, the overall layout presented here is a plausible and hopefully testable proposal. According to the best accepted view, the universe must have originated at the Planck era (10^{19} GeV, 10^{-43} s) from a quantum gravity fluctuation. Needless to say, we don't have any experimental evidence for such a statement: Quantum gravity phenomena are still in the realm of physical speculation. However, it is plausible that a primordial era of cosmological *inflation* originated then. Its consequences will be discussed below. Soon after, the universe may have reached the Grand Unified Theories (GUT) era (10^{16} GeV, 10^{-35} s). Quantum fluctuations of the inflaton field most probably left their imprint then as tiny perturbations in an otherwise very homogenous patch of the universe. At the end of inflation, the huge energy density of the inflaton field was converted into particles, which soon thermalized and became the origin of the hot Big Bang as we know it. Such a process is called *reheating* of the universe. Since then, the universe became radiation dominated. It is probable (although by no means certain) that the asymmetry between matter and antimatter originated at the same time as the rest of the energy of the universe, from the decay of the inflaton. This process is known under the name of *baryogenesis* since baryons (mostly quarks at that time) must have originated then, from the leftovers of their annihilation with antibaryons. It is a matter of speculation whether baryogenesis could have occurred at energies as low as the electroweak scale (100 GeV, 10^{-10} s). Note that although particle physics experiments have reached energies as high as 100 GeV, we still do not have observational evidence that the universe actually went through the EW phase transition. If confirmed, baryogenesis would constitute another "window" into the early universe. As the universe cooled down, it may have gone through the quark-gluon phase transition (10^2 MeV, 10^{-5} s), when baryons (mainly protons and neutrons) formed from their constituent quarks.

The furthest window we have on the early universe at the moment is that of *primordial nucleosynthesis* ($1 - 0.1$ MeV, 1 s – 3 min), when protons and neutrons were cold enough that bound systems could form, giving rise to the lightest elements, soon after *neutrino decoupling*: It is the realm of nuclear physics. The observed relative abundances of light elements are in agreement with the predictions of the hot Big Bang theory. Immediately afterwards, electron-positron annihilation occurs (0.5 MeV, 1 min) and all their energy goes into photons. Much later, at about (1 eV, $\sim 10^5$ yr), matter and radiation have equal energy densities. Soon after, electrons become bound to nuclei to form atoms (0.3 eV, 3×10^5 yr), in a process known as *recombination*: It is the realm of atomic physics. Immediately after, photons decouple from the plasma, travelling freely since then. Those are the photons we observe as the cosmic microwave background. Much later ($\sim 1 - 10$ Gyr), the small inhomogeneities generated during inflation have grown, via gravitational collapse, to become galaxies, clusters of galaxies, and superclusters, characterizing the epoch of *structure formation*. It is the realm of long range gravitational physics, perhaps

dominated by a vacuum energy in the form of a cosmological constant. Finally (3K, 13 Gyr), the Sun, the Earth, and biological life originated from previous generations of stars, and from a primordial soup of organic compounds, respectively.

I will now review some of the more robust features of the Hot Big Bang theory of which we have precise observational evidence.

2.6.4 *Primordial nucleosynthesis and light element abundance*

In this subsection I will briefly review Big Bang nucleosynthesis and give the present observational constraints on the amount of baryons in the universe. In 1920 Eddington suggested that the sun might derive its energy from the fusion of hydrogen into helium. The detailed reactions by which stars burn hydrogen were first laid out by Hans Bethe in 1939. Soon afterwards, in 1946, George Gamow realized that similar processes might have occurred also in the hot and dense early universe and gave rise to the first light elements [4]. These processes could take place when the universe had a temperature of around $T_{\rm NS} \sim 1 - 0.1$ MeV, which is about 100 times the temperature in the core of the Sun, while the density is $\rho_{\rm NS} = \frac{\pi^2}{30} g_* T_{\rm NS}^4 \sim 82$ g cm^{-3}, about the same density as the core of the Sun. Note, however, that although both processes are driven by identical thermonuclear reactions, the physical conditions in star and Big Bang nucleosynthesis are very different. In the former, gravitational collapse heats up the core of the star and reactions last for billions of years (except in supernova explosions, which last a few minutes and creates all the heavier elements beyond iron), while in the latter the universe expansion cools the hot and dense plasma in just a few minutes. Nevertheless, Gamow reasoned that, although the early period of cosmic expansion was much shorter than the lifetime of a star, there was a large number of free neutrons at that time, so that the lighter elements could be built up quickly by succesive neutron captures, starting with the reaction $n + p \to D + \gamma$. The abundances of the light elements would then be correlated with their neutron capture cross sections, in rough agreement with observations [6, 18].

Nowadays, Big Bang nucleosynthesis (BBN) codes compute a chain of around 30 coupled nuclear reactions [19], to produce all the light elements up to beryllium-7. [3] Only the first four or five elements can be computed with accuracy better than 1% and compared with cosmological observations. These light elements are $H, {}^4He, D, {}^3He, {}^7Li$, and perhaps also ${}^6Li$. Their observed relative abundance to hydrogen is $[1 : 0.25 : 3 \cdot 10^{-5} : 2 \cdot 10^{-5} : 2 \cdot 10^{-10}]$ with various errors, mainly systematic. The BBN codes calculate these abundances using the laboratory measured nuclear reaction rates, the decay rate of the neutron, the number of light neutrinos and the homogeneous FRW expansion of the universe, as a function of *only* one variable, the number density fraction of baryons to photons, $\eta \equiv n_{\rm B}/n_\gamma$. In fact, the present observations are only consistent, see Fig. 9 and Ref. [18–20], with a very narrow range of values of

$$\eta_{10} \equiv 10^{10}\, \eta = 6.2 \pm 0.6 \,. \tag{88}$$

Such a small value of η indicates that there is about one baryon per 10^9 photons in the universe today. Any acceptable theory of baryogenesis should account for such a small number. Furthermore, the present baryon fraction of the critical density can be calculated from η_{10} as

$$\Omega_{\rm B} h^2 = 3.6271 \times 10^{-3}\, \eta_{10} = 0.0224 \pm 0.0022 \quad \text{(95\% c.l.)} \tag{89}$$

Clearly, this number is well below closure density, so baryons cannot account for all the matter in the universe, as I shall discuss below.

2.6.5 *Neutrino decoupling*

Just before the nucleosynthesis of the lightest elements in the early universe, weak interactions were too slow to keep neutrinos in thermal equilibrium with the plasma, so they decoupled. We can estimate the

---

[3]The rest of nuclei, up to iron (Fe), are produced in heavy stars, and beyond Fe in novae and supernovae explosions.
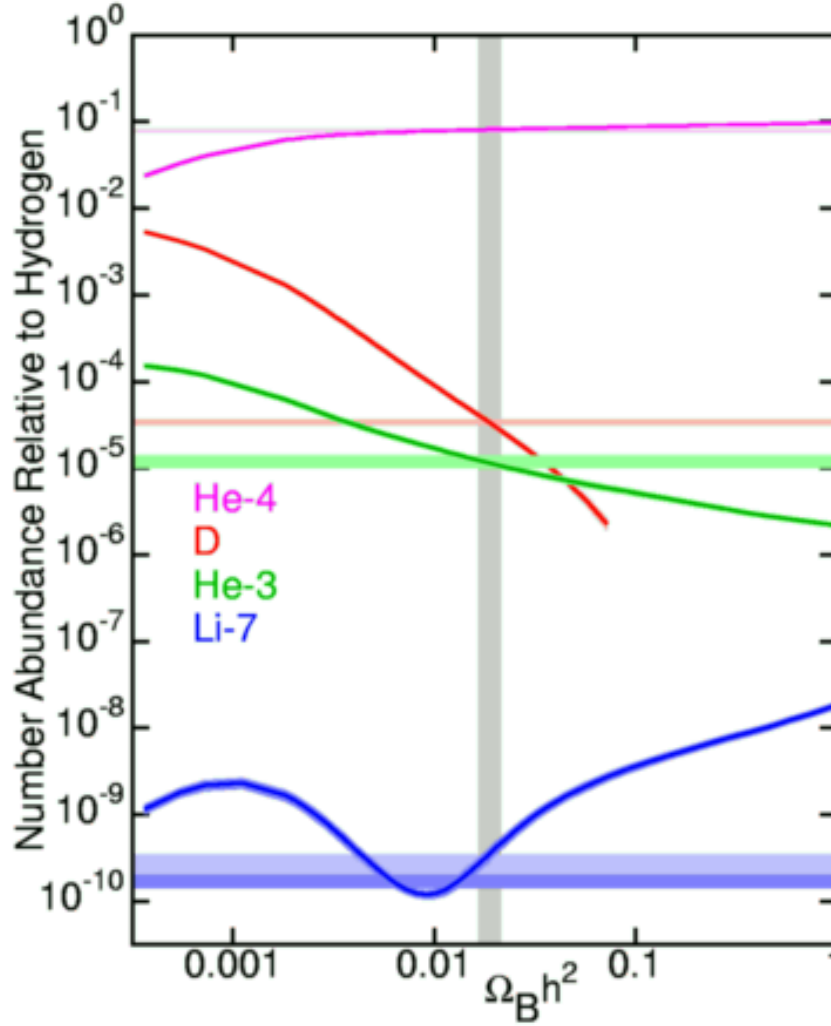
Fig. 7: The relative abundance of light elements to Hidrogen. Note the large range of scales involved. From Ref. [18].

temperature at which decoupling occurred from the weak interaction cross section, $\sigma_{\rm w} \simeq G_F^2 T^2$ at finite temperature T, where $G_F = 1.2 \times 10^{-5}$ GeV^{-2} is the Fermi constant. The neutrino interaction rate, via W boson exchange in $n + \nu \leftrightarrow p + e^-$ and $p + \bar{\nu} \leftrightarrow n + e^+$, can be written as [16]

$$\Gamma_\nu = n_\nu \langle \sigma_{\rm w} |v| \rangle \simeq G_F^2 T^5 \,, \tag{90}$$

while the rate of expansion of the universe at that time ($g_* = 10.75$) was $H \simeq 5.4 \, T^2/M_{\rm P}$, where $M_{\rm P} = 1.22 \times 10^{19}$ GeV is the Planck mass. Neutrinos decouple when their interaction rate is slower than the universe expansion, $\Gamma_\nu \leq H$ or, equivalently, at $T_{\nu-{\rm dec}} \simeq 0.8$ MeV. Below this temperature, neutrinos are no longer in thermal equilibrium with the rest of the plasma, and their temperature continues to decay inversely proportional to the scale factor of the universe. Since neutrinos decoupled before $e^+ e^-$ annihilation, the cosmic background of neutrinos has a temperature today lower than that of the microwave background of photons. Let us compute the difference. At temperatures above the the mass of the electron, $T > m_e = 0.511$ MeV, and below 0.8 MeV, the only particle species contributing to the entropy of the universe are the photons ($g_* = 2$) and the electron-positron pairs ($g_* = 4 \times \frac{7}{8}$); total number of degrees of freedom $g_* = \frac{11}{2}$. At temperatures $T \simeq m_e$, electrons and positrons annihilate into photons, heating up the plasma (but not the neutrinos, which had decoupled already). At temperatures

$T < m_e$, only photons contribute to the entropy of the universe, with $g_* = 2$ degrees of freedom. Therefore, from the conservation of entropy, we find that the ratio of T_γ and T_ν today must be

$$\frac{T_\gamma}{T_\nu} = \left(\frac{11}{4}\right)^{1/3} = 1.401 \quad \Rightarrow \quad T_\nu = 1.945 \text{ K}, \tag{91}$$

where I have used $T_{\text{CMB}} = 2.725 \pm 0.002$ K. We still have not measured such a relic background of neutrinos, and probably will remain undetected for a long time, since they have an average energy of order 10^{-4} eV, much below that required for detection by present experiments (of order GeV), precisely because of the relative weakness of the weak interactions. Nevertheless, it would be fascinating if, in the future, ingenious experiments were devised to detect such a background, since it would confirm one of the most robust features of Big Bang cosmology.

2.6.6 Matter-radiation equality

Relativistic species have energy densities proportional to the quartic power of temperature and therefore scale as $\rho_R \propto a^{-4}$, while non-relativistic particles have essentially zero pressure and scale as $\rho_M \propto a^{-3}$. Therefore, there will be a time in the evolution of the universe in which both energy densities are equal $\rho_R(t_{\text{eq}}) = \rho_M(t_{\text{eq}})$. Since then both decay differently, and thus

$$1 + z_{\text{eq}} = \frac{a_0}{a_{\text{eq}}} = \frac{\Omega_M}{\Omega_R} = 3.1 \times 10^4 \, \Omega_M h^2, \tag{92}$$

where I have used $\Omega_R h^2 = \Omega_{\text{CMB}} h^2 + \Omega_\nu h^2 = 3.24 \times 10^{-5}$ for three massless neutrinos at $T = T_\nu$. As I will show later, the matter content of the universe today is below critical, $\Omega_M \simeq 0.3$, while $h \simeq 0.71$, and therefore $(1 + z_{\text{eq}}) \simeq 3400$, or about $t_{\text{eq}} = 1308 \, (\Omega_M h^2)^{-2}$yr $\simeq 61,000$ years after the origin of the universe. Around the time of matter-radiation equality, the rate of expansion (19) can be written as ($a_0 \equiv 1$)

$$H(a) = H_0 \left(\Omega_R \, a^{-4} + \Omega_M \, a^{-3}\right)^{1/2} = H_0 \, \Omega_M^{1/2} \, a^{-3/2} \left(1 + \frac{a_{\text{eq}}}{a}\right)^{1/2}. \tag{93}$$

The *horizon size* is the coordinate distance travelled by a photon since the beginning of the universe, $d_H \sim H^{-1}$, i.e. the size of causally connected regions in the universe. The *comoving* horizon size is then given by

$$d_H(a) = \int \frac{da}{a^2 H(a)} = \frac{2c \, H_0^{-1}}{\sqrt{\Omega_M(1 + z_{\text{eq}})}} \left(\sqrt{\frac{a}{a_{\text{eq}}} + 1} - 1\right). \tag{94}$$

Thus the horizon size at matter-radiation equality ($a = a_{\text{eq}}$) is

$$d_H(a_{\text{eq}}) \simeq 14 \, (\Omega_M h)^{-1} \text{ Mpc}/h. \tag{95}$$

As we will see later, this scale plays a very important role in theories of structure formation.

2.6.7 Recombination and photon decoupling

As the temperature of the universe decreased, electrons could eventually become bound to protons to form neutral hydrogen. Nevertheless, there is always a non-zero probability that a rare energetic photon ionizes hydrogen and produces a free electron. The *ionization fraction* of electrons in equilibrium with the plasma at a given temperature is given by the Saha equation [16]

$$\frac{1 - X_e^{\text{eq}}}{(X_e^{\text{eq}})^2} = \frac{4\sqrt{2}\zeta(3)}{\sqrt{\pi}} \eta \left(\frac{T}{m_e}\right)^{3/2} e^{E_{\text{ion}}/T}, \tag{96}$$

where $E_{\text{ion}} = 13.6$ eV is the ionization energy of hydrogen, and η is the baryon-to-photon ratio (88). If we now use Eq. (75), we can compute the ionization fraction X_e^{eq} as a function of redshift z. Note that

the huge number of photons with respect to electrons (in the ratio $^4He : H : \gamma \simeq 1 : 4 : 10^{10}$) implies that even at a very low temperature, the photon distribution will contain a sufficiently large number of high-energy photons to ionize a significant fraction of hydrogen. In fact, *defining* recombination as the time at which $X_e^{\text{eq}} \equiv 0.1$, one finds that the recombination temperature is $T_{\text{rec}} = 0.296$ eV $\ll E_{\text{ion}}$, for $\eta_{10} \simeq 6.1$. Comparing with the present temperature of the microwave background, we deduce the corresponding redshift at recombination, $(1 + z_{\text{rec}}) \simeq 1260$.

Photons remain in thermal equilibrium with the plasma of baryons and electrons through elastic Thomson scattering, with cross section

$$\sigma_T = \frac{8\pi\alpha^2}{3m_e^2} = 6.65 \times 10^{-25} \text{ cm}^2 = 0.665 \text{ barn} \,, \tag{97}$$

where $\alpha = 1/137.036$ is the dimensionless electromagnetic coupling constant. The mean free path of photons λ_γ in such a plasma can be estimated from the photon interaction rate, $c\lambda_\gamma^{-1} \simeq \Gamma_\gamma = n_e\sigma_T c$. For temperatures above a few eV, the mean free path is much smaller that the causal horizon at that time and photons suffer multiple scattering: the plasma is like a dense fog. Photons will decouple from the plasma when their interaction rate cannot keep up with the expansion of the universe and the mean free path becomes larger than the horizon size: the universe becomes transparent. We can estimate this moment by evaluating $\Gamma_\gamma = H$ at photon decoupling. Using $n_e = X_e \eta n_\gamma$, one can compute the decoupling temperature as $T_{\text{dec}} = 0.256$ eV, and the corresponding redshift as $1 + z_{\text{dec}} \simeq 1090$. Recently, WMAP and Planck measured this redshift to be $1 + z_{\text{dec}} \simeq 1089 \pm 1$ [21]. This redshift defines the so called *last scattering surface*, when photons last scattered off protons and electrons and travelled freely ever since. This decoupling occurred when the universe was approximately $t_{\text{dec}} = 1.5 \times 10^5 \, (\Omega_M h^2)^{-1/2} \simeq 380,000$ years old.
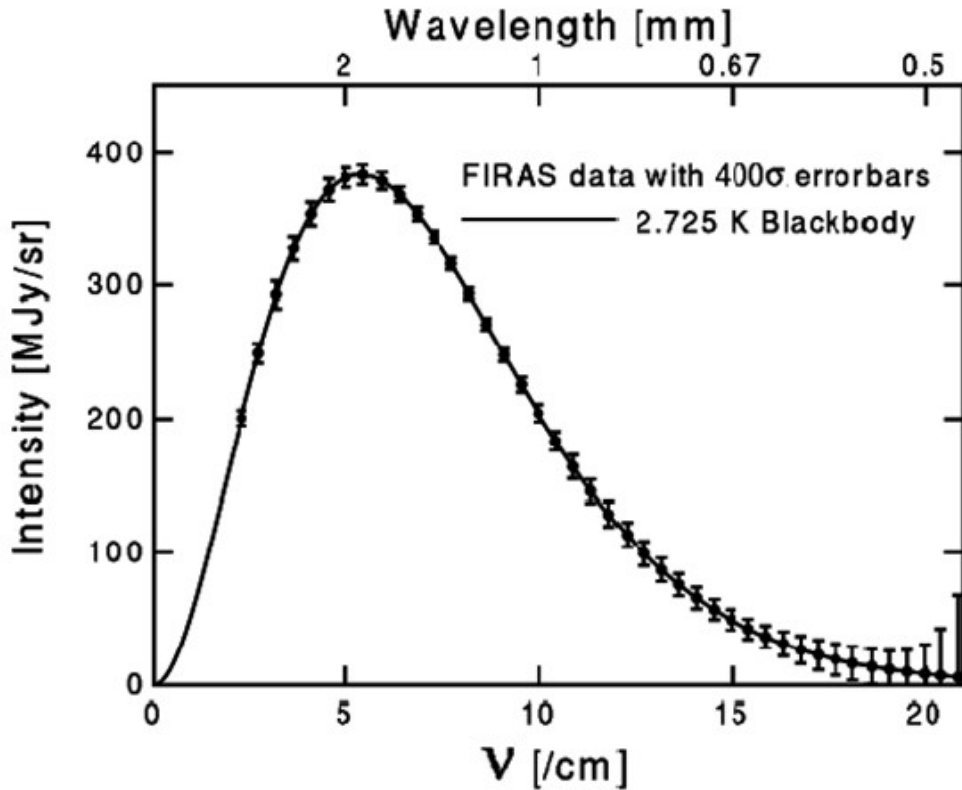
Fig. 8: The Cosmic Microwave Background Spectrum seen by the FIRAS instrument on COBE. The CMB blackbody spectrum has a temperature $T_0 = 2.725 \pm 0.002$ K. From Ref. [22].

2.6.8   The microwave background

One of the most remarkable observations ever made my mankind is the detection of the relic background of photons from the Big Bang. This background was predicted by George Gamow and collaborators in the 1940s, based on the consistency of primordial nucleosynthesis with the observed helium abundance. They estimated a value of about 10 K, although a somewhat more detailed analysis by Alpher and Herman in 1949 predicted $T_\gamma \approx 5$ K. Unfortunately, they had doubts whether the radiation would have survived until the present, and this remarkable prediction slipped into obscurity, until Dicke, Peebles, Roll and Wilkinson [23] studied the problem again in 1965. Before they could measure the photon background, they learned that Penzias and Wilson had observed a weak isotropic background signal at a radio wavelength of 7.35 cm, corresponding to a blackbody temperature of $T_\gamma = 3.5 \pm 1$ K. They published their two papers back to back, with that of Dicke et al. explaining the fundamental significance of their measurement [6].

Since then many different experiments have confirmed the existence of the microwave background. The most outstanding one has been the Cosmic Background Explorer (COBE) satellite, whose FIRAS instrument measured the photon background with great accuracy over a wide range of frequencies ($\nu = 1 - 97$ cm^{-1}), see Ref. [22], with a spectral resolution $\frac{\Delta\nu}{\nu} = 0.0035$. Nowadays, the photon spectrum is confirmed to be a blackbody spectrum with a temperature given by [22]

$$T_{\mathrm{CMB}} = 2.725 \pm 0.002 \text{ K (systematic, 95\% c.l.)} \pm 7\ \mu\text{K (1}\sigma \text{ statistical)} \tag{98}$$

In fact, this is the best blackbody spectrum ever measured, see Fig. 8, with spectral distortions below the level of 10 parts per million (ppm).

Moreover, the differential microwave radiometer (DMR) instrument on COBE, with a resolution of about $7°$ in the sky, has also confirmed that it is an extraordinarily isotropic background. The deviations from isotropy, i.e. differences in the temperature of the blackbody spectrum measured in different directions in the sky, are of the order of $20\ \mu$K on large scales, or one part in 10^5, see Ref. [24]. There is, in fact, a dipole anisotropy of one part in 10^3, $\delta T_1 = 3.372 \pm 0.007$ mK (95% c.l.), in the direction of the Virgo cluster, $(l, b) = (264.14° \pm 0.30, 48.26° \pm 0.30)$ (95% c.l.). Under the assumption that a Doppler effect is responsible for the entire CMB dipole, the velocity of the Sun with respect to the CMB rest frame is $v_\odot = 371 \pm 0.5$ km/s, see Ref. [22].[4] When subtracted, we are left with a whole spectrum of anisotropies in the higher multipoles (quadrupole, octupole, etc.), $\delta T_2 = 18 \pm 2\ \mu$K (95% c.l.), see Ref. [24] and Fig. 9.

Soon after COBE, other groups quickly confirmed the detection of temperature anisotropies at around $30\ \mu$K and above, at higher multipole numbers or smaller angular scales. As I shall discuss below, these anisotropies play a crucial role in the understanding of the origin of structure in the universe.

2.6.9   Large-scale structure formation

Although the isotropic microwave background indicates that the universe in the *past* was extraordinarily homogeneous, we know that the universe *today* is not exactly homogeneous: we observe galaxies, clusters and superclusters on large scales. These structures are expected to arise from very small primordial inhomogeneities that grow in time via gravitational instability, and that may have originated from tiny ripples in the metric, as matter fell into their troughs. Those ripples must have left some trace as temperature anisotropies in the microwave background, and indeed such anisotropies were finally discovered by the COBE satellite in 1992. The reason why they took so long to be discovered was that they appear as perturbations in temperature of only one part in 10^5.

While the predicted anisotropies have finally been seen in the CMB, not all kinds of matter and/or evolution of the universe can give rise to the structure we observe today. If we define the density contrast

---

[4]COBE even determined the annual variation due to the Earth's motion around the Sun – the ultimate proof of Copernicus' hypothesis.
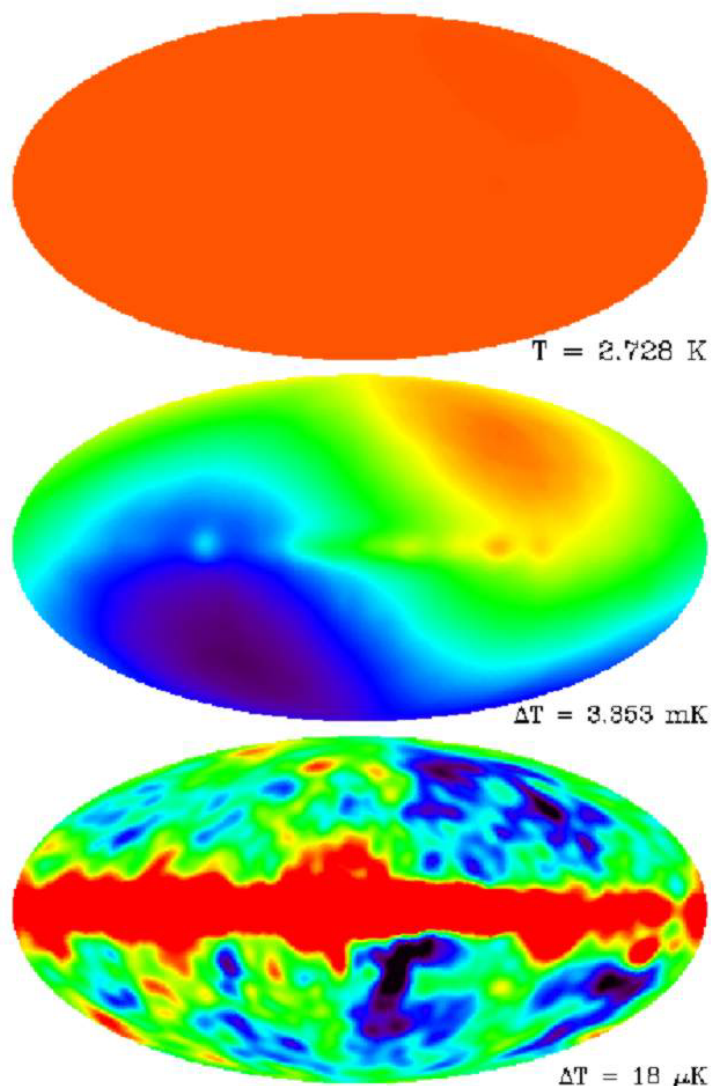
Fig. 9: The Cosmic Microwave Background Spectrum seen by the DMR instrument on COBE. The top figure corresponds to the monopole, $T_0 = 2.725 \pm 0.002$ K. The middle figure shows the dipole, $\delta T_1 = 3.372 \pm 0.014$ mK, and the lower figure shows the quadrupole and higher multipoles, $\delta T_2 = 18 \pm 2$ μK. The central region corresponds to foreground by the galaxy. From Ref. [24].

as [25]

$$\delta(\vec{x}, a) \equiv \frac{\rho(\vec{x}, a) - \bar{\rho}(a)}{\bar{\rho}(a)} = \int d^3\vec{k}\, \delta_k(a)\, e^{i\vec{k}\cdot\vec{x}}, \tag{99}$$

where $\bar{\rho}(a) = \rho_0\, a^{-3}$ is the average cosmic density, we need a theory that will grow a density contrast with amplitude $\delta \sim 10^{-5}$ at the last scattering surface ($z = 1100$) up to density contrasts of the order of $\delta \sim 10^6$ for galaxies at redshifts $z \ll 1$, i.e. today. This is a *necessary* requirement for any consistent theory of structure formation [26].

Furthermore, the anisotropies observed by the Planck satellite correspond to a small-amplitude scale-invariant primordial power spectrum of inhomogeneities

$$P(k) = \langle |\delta_k|^2 \rangle \propto k^n, \quad \text{with} \quad n = 1, \tag{100}$$

where the brackets $\langle \cdot \rangle$ represent integration over an ensemble of different universe realizations. These inhomogeneities are like waves in the space-time metric. When matter fell in the troughs of those waves, it created density perturbations that collapsed gravitationally to form galaxies and clusters of galaxies, with a spectrum that is also scale invariant. Such a type of spectrum was proposed in the early 1970s by Edward R. Harrison, and independently by the Russian cosmologist Yakov B. Zel'dovich, see Ref. [27], to explain the distribution of galaxies and clusters of galaxies on very large scales in our observable universe.
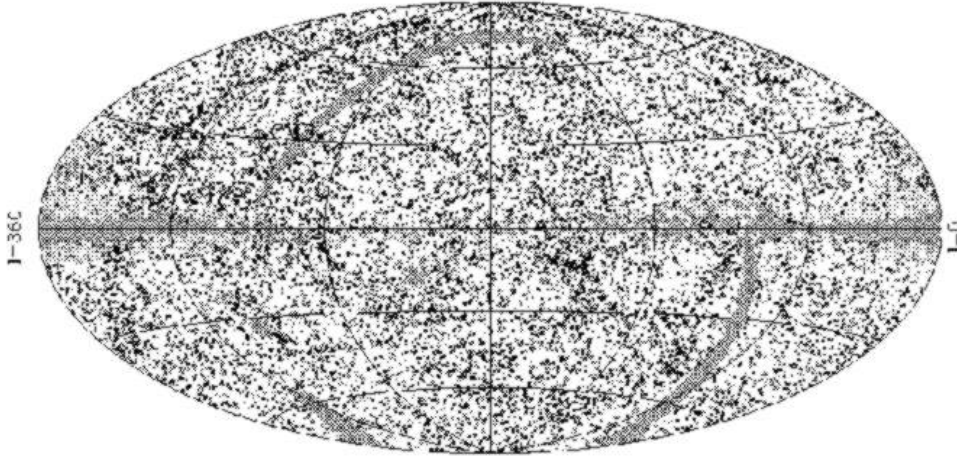
Fig. 10: The IRAS Point Source Catalog redshift survey contains some 15,000 galaxies, covering over 83% of the sky up to redshifts of $z \leq 0.05$. We show here the projection of the galaxy distribution in galactic coordinates. From Ref. [28].

Today various telescopes – like the Hubble Space Telescope, the twin Keck telescopes in Hawaii and the European Southern Observatory telescopes in Chile – are exploring the most distant regions of the universe and discovering the first galaxies at large distances. The furthest galaxies observed so far are at redshifts of $z \simeq 10$ (at a distance of 13.7 billion light years from Earth), whose light was emitted when the universe had only about 3% of its present age. Only a few galaxies are known at those redshifts, but there are at present various catalogs like the CfA and APM galaxy catalogs, and more recently the IRAS Point Source redshift Catalog, see Fig. 10, and Las Campanas redshift surveys, that study the spatial distribution of hundreds of thousands of galaxies up to distances of a billion light years, or $z < 0.1$, or the 2 degree Field Galaxy Redshift Survey (2dFGRS) and the Sloan Digital Sky Survey (SDSS), which reach $z < 0.5$ and study millions of galaxies. These catalogs are telling us about the evolution of clusters and superclusters of galaxies in the universe, and already put constraints on the theory of structure formation. From these observations one can infer that most galaxies formed at redshifts of the order of $2 - 6$; clusters of galaxies formed at redshifts of order 1, and superclusters are forming now. That is, cosmic structure formed from the bottom up: from galaxies to clusters to superclusters, and not the other way around. This fundamental difference is an indication of the type of matter that gave rise to structure.

We know from Big Bang nucleosynthesis that all the baryons in the universe cannot account for the observed amount of matter, so there must be some extra matter (dark since we don't see it) to account for its gravitational pull. Whether it is relativistic (hot) or non-relativistic (cold) could be inferred from observations: relativistic particles tend to diffuse from one concentration of matter to another, thus transferring energy among them and preventing the growth of structure on small scales. This is excluded by observations, so we conclude that most of the matter responsible for structure formation must be cold. How much there is is a matter of debate at the moment. Some recent analyses suggest that there is not enough cold dark matter to reach the critical density required to make the universe flat. If we want to

make sense of the present observations, we must conclude that some other form of energy permeates the universe. In order to resolve this issue, 2dFGRS and SDSS started taking data a few years ago. The first has already been completed, but the second one is still taking data up to redshifts $z \simeq 5$ for quasars, over a large region of the sky. These important observations will help astronomers determine the nature of the dark matter and test the validity of the models of structure formation.

Before COBE discovered the anisotropies of the microwave background there were serious doubts whether gravity alone could be responsible for the formation of the structure we observe in the universe today. It seemed that a new force was required to do the job. Fortunately, the anisotropies were found with the right amplitude for structure to be accounted for by gravitational collapse of primordial inhomogeneities under the attraction of a large component of non-relativistic dark matter. Nowadays, the standard theory of structure formation is a cold dark matter model with a non vanishing cosmological constant in a spatially flat universe. Gravitational collapse amplifies the density contrast initially through linear growth and later on via non-linear collapse. In the process, overdense regions decouple from the Hubble expansion to become bound systems, which start attracting eachother to form larger bound structures. In fact, the largest structures, superclusters, have not yet gone non-linear.

The primordial spectrum (100) is reprocessed by gravitational instability after the universe becomes matter dominated and inhomogeneities can grow. Linear perturbation theory shows that the growing mode [5] of small density contrasts go like $\delta(a) \propto a$, in the Einstein-de Sitter limit [25, 26]. There are slight deviations for $a \gg a_{\rm eq}$, if $\Omega_{\rm M} \neq 1$ or $\Omega_{\Lambda} \neq 0$, but we will not be concerned with them here. The important observation is that, since the density contrast at last scattering is of order $\delta \sim 10^{-5}$, and the scale factor has grown since then only a factor $z_{\rm dec} \sim 10^3$, one would expect a density contrast today of order $\delta_0 \sim 10^{-2}$. Instead, we observe structures like galaxies, where $\delta \sim 10^6$. So how can this be possible? The microwave background shows anisotropies due to fluctuations in the baryonic matter component only (to which photons couple, electromagnetically). If there is an additional matter component that only couples through very weak interactions, fluctuations in that component could grow as soon as it decoupled from the plasma, well before photons decoupled from baryons. The reason why baryonic inhomogeneities cannot grow is because of photon pressure: as baryons collapse towards denser regions, radiation pressure eventually halts the contraction and sets up acoustic oscillations in the plasma that prevent the growth of perturbations, until photon decoupling. On the other hand, a weakly interacting cold dark matter component could start gravitational collapse much earlier, even before matter-radiation equality, and thus reach the density contrast amplitudes observed today. The resolution of this mismatch is one of the strongest arguments for the existence of a weakly interacting cold dark matter component of the universe.

How much dark matter there is in the universe can be deduced from the actual power spectrum (the Fourier transform of the two-point correlation function of density perturbations) of the observed large scale structure. One can decompose the density contrast in Fourier components, see Eq. (99). This is very convenient since in linear perturbation theory individual Fourier components evolve independently. A comoving wavenumber k is said to "enter the horizon" when $k = d_H^{-1}(a) = aH(a)$. If a certain perturbation, of wavelength $\lambda = k^{-1} < d_H(a_{\rm eq})$, enters the horizon before matter-radiation equality, the fast radiation-driven expansion prevents dark-matter perturbations from collapsing. Since light can only cross regions that are smaller than the horizon, the suppression of growth due to radiation is restricted to scales smaller than the horizon, while large-scale perturbations remain unaffected. This is the reason why the horizon size at equality, Eq. (95), sets an important scale for structure growth,

$$k_{\rm eq} = d_H^{-1}(a_{\rm eq}) \simeq 0.083 \, (\Omega_{\rm M} h) \, h \, {\rm Mpc}^{-1} \,. \tag{101}$$

The suppression factor can be easily computed as $f_{\rm sup} = (a_{\rm enter}/a_{\rm eq})^2 = (k_{\rm eq}/k)^2$. In other words, the

---

[5] The decaying mode goes like $\delta(t) \sim t^{-1}$, for all ω.
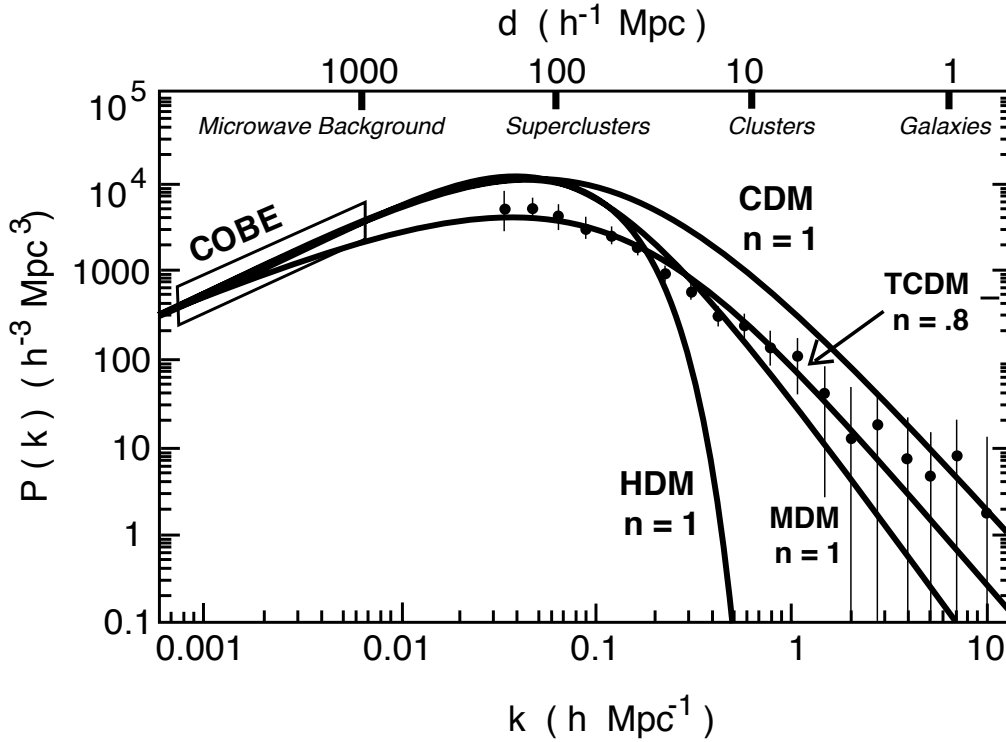
Fig. 11: The power spectrum for cold dark matter (CDM), tilted cold dark matter (TCDM), hot dark matter (HDM), and mixed hot plus cold dark matter (MDM), normalized to COBE, for large-scale structure formation. From Ref. [29].

processed power spectrum $P(k)$ will have the form:

$$P(k) \propto \begin{cases} k, & k \ll k_{eq} \\ k^{-3}, & k \gg k_{eq} \end{cases} \tag{102}$$

This is precisely the shape that large-scale galaxy catalogs are bound to test in the near future, see Fig. 11. Furthermore, since relativistic Hot Dark Matter (HDM) transfer energy between clumps of matter, they will wipe out small scale perturbations, and this should be seen as a distinctive signature in the matter power spectra of future galaxy catalogs. On the other hand, non-relativistic Cold Dark Matter (CDM) allow structure to form on *all* scales via gravitational collapse. The dark matter will then pull in the baryons, which will later shine and thus allow us to see the galaxies.

Naturally, when baryons start to collapse onto dark matter potential wells, they will convert a large fraction of their potential energy into kinetic energy of protons and electrons, ionizing the medium. As a consequence, we expect to see a large fraction of those baryons constituting a hot ionized gas surrounding large clusters of galaxies. This is indeed what is observed, and confirms the general picture of structure formation.

3 Determination of Cosmological Parameters

In this Section, I will restrict myself to those recent measurements of the cosmological parameters by means of standard cosmological techniques, together with a few instances of new results from recently applied techniques. We will see that a large host of observations are determining the cosmological parameters with some reliability of the order of 10%. However, the majority of these measurements are dominated by large systematic errors. Most of the recent work in observational cosmology has been

the search for virtually systematic-free observables, like those obtained from the microwave background anisotropies, and discussed in Section 4.4. I will devote, however, this Section to the more 'classical' measurements of the following cosmological parameters: The rate of expansion H_0; the matter content Ω_M; the cosmological constant Ω_Λ; the spatial curvature Ω_K, and the age of the universe t_0.

3.1   The rate of expansion H_0

Over most of last century the value of H_0 has been a constant source of disagreement [30]. Around 1929, Hubble measured the rate of expansion to be $H_0 = 500 \ \mathrm{km\,s^{-1}Mpc^{-1}}$, which implied an age of the universe of order $t_0 \sim 2$ Gyr, in clear conflict with geology. Hubble's data was based on Cepheid standard candles that were incorrectly calibrated with those in the Large Magellanic Cloud. Later on, in 1954 Baade recalibrated the Cepheid distance and obtained a lower value, $H_0 = 250 \ \mathrm{km\,s^{-1}Mpc^{-1}}$, still in conflict with ratios of certain unstable isotopes. Finally, in 1958 Sandage realized that the brightest stars in galaxies were ionized HII regions, and the Hubble rate dropped down to $H_0 = 60 \ \mathrm{km\,s^{-1}}$ $\mathrm{Mpc^{-1}}$, still with large (factor of two) systematic errors. Fortunately, in the past 15 years there has been significant progress towards the determination of H_0, with systematic errors approaching the 10% level. These improvements come from two directions. First, technological, through the replacement of photographic plates (almost exclusively the source of data from the 1920s to 1980s) with charged couple devices (CCDs), i.e. solid state detectors with excellent flux sensitivity per pixel, which were previously used successfully in particle physics detectors. Second, by the refinement of existing methods for measuring extragalactic distances (e.g. parallax, Cepheids, supernovae, etc.). Finally, with the development of completely new methods to determine H_0, which fall into totally independent and very broad categories: a) Gravitational lensing; b) Sunyaev-Zel'dovich effect; c) Extragalactic distance scale, mainly Cepheid variability and type Ia Supernovae; d) Microwave background anisotropies. I will review here the first three, and leave the last method for Section 4.4, since it involves knowledge about the primordial spectrum of inhomogeneities.

3.1.1   Gravitational lensing

Imagine a quasi-stellar object (QSO) at large redshift ($z \gg 1$) whose light is lensed by an intervening galaxy at redshift $z \sim 1$ and arrives to an observer at $z = 0$. There will be at least two different images of the same background *variable* point source. The arrival times of photons from two different gravitationally lensed images of the quasar depend on the different path lengths and the gravitational potential traversed. Therefore, a measurement of the time delay and the angular separation of the different images of a variable quasar can be used to determine H_0 with great accuracy. This method, proposed in 1964 by Refsdael [31], offers tremendous potential because it can be applied at great distances and it is based on very solid physical principles [32].

Unfortunately, there are very few systems with both a favourable geometry (i.e. a known mass distribution of the intervening galaxy) and a variable background source with a measurable time delay. That is the reason why it has taken so much time since the original proposal for the first results to come out. Fortunately, there are now very powerful telescopes that can be used for these purposes. The best candidate to-date is the QSO $0957 + 561$, observed with the 10m Keck telescope, for which there is a model of the lensing mass distribution that is consistent with the measured velocity dispersion. Assuming a flat space with $\Omega_M = 0.25$, one can determine [33]

$$H_0 = 72 \pm 7 \ (1\sigma \ \text{statistical}) \ \pm 15\% \ (\text{systematic}) \ \mathrm{km\,s^{-1}Mpc^{-1}} \,. \tag{103}$$

The main source of systematic error is the degeneracy between the mass distribution of the lens and the value of H_0. Knowledge of the velocity dispersion within the lens as a function of position helps constrain the mass distribution, but those measurements are very difficult and, in the case of lensing by a cluster of galaxies, the dark matter distribution in those systems is usually unknown, associated with a

complicated cluster potential. Nevertheless, the method is just starting to give promising results and, in the near future, with the recent discovery of several systems with optimum properties, the prospects for measuring H_0 and lowering its uncertainty with this technique are excellent.

3.1.2 Sunyaev-Zel'dovich effect

As discussed in the previous Section, the gravitational collapse of baryons onto the potential wells generated by dark matter gave rise to the reionization of the plasma, generating an X-ray halo around rich clusters of galaxies, see Fig. 12. The inverse-Compton scattering of microwave background photons off the hot electrons in the X-ray gas results in a measurable distortion of the blackbody spectrum of the microwave background, known as the Sunyaev-Zel'dovich (SZ) effect. Since photons acquire extra energy from the X-ray electrons, we expect a shift towards higher frequencies of the spectrum, $(\Delta\nu/\nu) \simeq (k_B T_{\text{gas}}/m_e c^2) \sim 10^{-2}$. This corresponds to a *decrement* of the microwave background temperature at low frequencies (Rayleigh-Jeans region) and an increment at high frequencies, see Ref. [34].
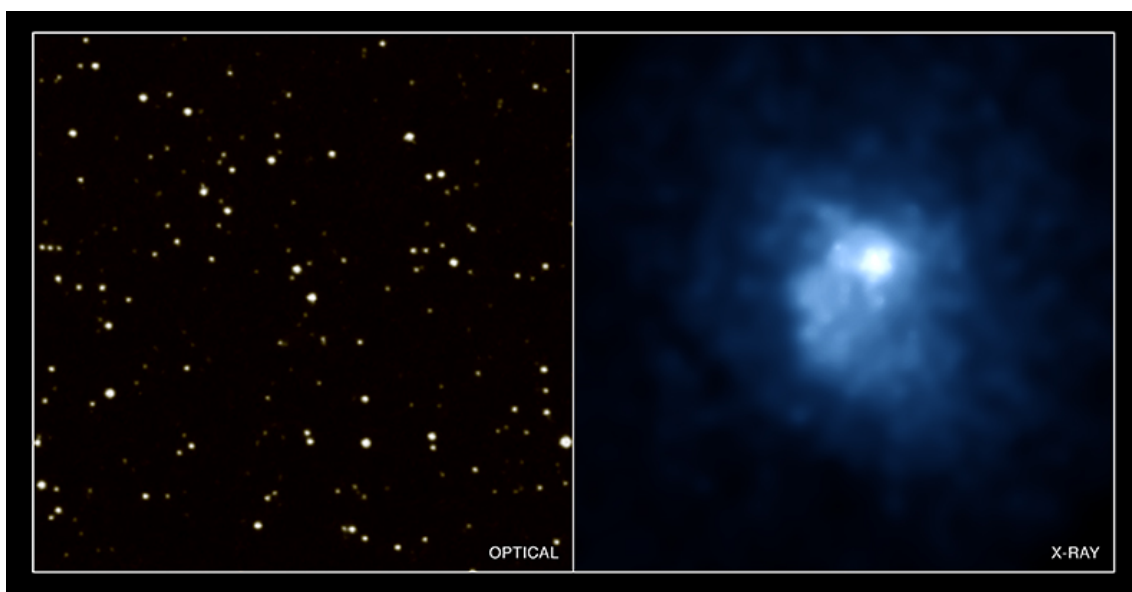
Fig. 12: The 3C438 cluster of galaxies, seen here in an optical image (left) and an X-ray image (right), taken by Chandra X-ray Observatory. It is clear that the gas in the center of the cluster is very hot and has no optical counterpart with any particular galaxy. From Ref. [35].

Measuring the *spatial* distribution of the SZ effect (3 K spectrum), together with a high resolution X-ray map (10^8 K spectrum) of the cluster, one can determine the density and temperature distribution of the hot gas. Since the X-ray flux is distance-dependent ($\mathcal{F} = \mathcal{L}/4\pi d_L^2$), while the SZ decrement is not (because the energy of the CMB photons increases as we go back in redshift, $\nu = \nu_0(1 + z)$, and exactly compensates the redshift in energy of the photons that reach us), one can determine from there the distance to the cluster, and thus the Hubble rate H_0.

The advantages of this method are that it can be applied to large distances and it is based on clear physical principles. The main systematics come from possible clumpiness of the gas (which would reduce H_0), projection effects (if the clusters are prolate, H_0 could be larger), the assumption of hydrostatic equilibrium of the X-ray gas, details of models for the gas and electron densities, and possible contaminations from point sources. Present measurements give the value [34]

$$H_0 = 60 \pm 10 \ (1\sigma \text{ statistical}) \ \pm 20\% \ (\text{systematic}) \ \text{km s}^{-1}\text{Mpc}^{-1}, \tag{104}$$

compatible with other determinations. A great advantage of this completely new and independent method is that nowadays more and more clusters are observed in the X-ray, and soon we will have high-resolution 2D maps of the SZ decrement from several balloon flights, as well as from future microwave background satellites, together with precise X-ray maps and spectra from the Chandra X-ray observatory recently launched by NASA, as well as from the European X-ray satellite XMM launched a few months ago by ESA, which will deliver orders of magnitude better resolution than the existing Einstein X-ray satellite.

3.1.3 *Cepheid variability*

Cepheids are low-mass variable stars with a period-luminosity relation based on the helium ionization cycles inside the star, as it contracts and expands. This time variability can be measured, and the star's absolute luminosity determined from the calibrated relationship. From the observed flux one can then deduce the luminosity distance, see Eq. (28), and thus the Hubble rate H_0. The Hubble Space Telescope (HST) was launched by NASA in 1990 (and repaired in 1993) with the specific project of calibrating the extragalactic distance scale and thus determining the Hubble rate with 10% accuracy. The most recent results from HST are the following [36]

$$H_0 = 71 \pm 4 \text{ (random) } \pm 7 \text{ (systematic) } \text{ km s}^{-1}\text{Mpc}^{-1}. \tag{105}$$

The main source of systematic error is the distance to the Large Magellanic Cloud, which provides the fiducial comparison for Cepheids in more distant galaxies. Other systematic uncertainties that affect the value of H_0 are the internal extinction correction method used, a possible metallicity dependence of the Cepheid period-luminosity relation and cluster population incompleteness bias, for a set of 21 galaxies within 25 Mpc, and 23 clusters within $z \lesssim 0.03$.

With better telescopes already taking data, like the Very Large Telescope (VLT) interferometer of the European Southern Observatory (ESO) in the Chilean Atacama desert, with 8 synchronized telescopes, and others coming up soon, like the Next Generation Space Telescope (NGST) proposed by NASA for 2008, and the Gran TeCan of the European Northern Observatory in the Canary Islands, for 2010, it is expected that much better resolution and therefore accuracy can be obtained for the determination of H_0.

3.2 Dark Matter

In the 1920s Hubble realized that the so called nebulae were actually distant galaxies very similar to our own. Soon afterwards, in 1933, Zwicky found dynamical evidence that there is possibly ten to a hundred times more mass in the Coma cluster than contributed by the luminous matter in galaxies [37]. However, it was not until the 1970s that the existence of dark matter began to be taken more seriously. At that time there was evidence that rotation curves of galaxies did not fall off with radius and that the dynamical mass was increasing with scale from that of individual galaxies up to clusters of galaxies. Since then, new possible extra sources to the matter content of the universe have been accumulating:

$$
\begin{aligned}
\Omega_M \ &= \ \Omega_{B, \text{ lum}} && \text{(stars in galaxies)} && (106) \\
&+ \ \Omega_{B, \text{ dark}} && \text{(MACHOs?)} && (107) \\
&+ \ \Omega_{CDM} && \text{(weakly interacting : axion, neutralino?)} && (108) \\
&+ \ \Omega_{HDM} && \text{(massive neutrinos?)} && (109)
\end{aligned}
$$

The empirical route to the determination of Ω_M is nowadays one of the most diversified of all cosmological parameters. The matter content of the universe can be deduced from the mass-to-light ratio of various objects in the universe; from the rotation curves of galaxies; from microlensing and the direct search of Massive Compact Halo Objects (MACHOs); from the cluster velocity dispersion with the use of the Virial theorem; from the baryon fraction in the X-ray gas of clusters; from weak gravitational

lensing; from the observed matter distribution of the universe via its power spectrum; from the cluster abundance and its evolution; from direct detection of massive neutrinos at SuperKamiokande; from direct detection of Weakly Interacting Massive Particles (WIMPs) at CDMS, DAMA or UKDMC, and finally from microwave background anisotropies. I will review here just a few of them.

3.2.1 Rotation curves of spiral galaxies

The flat rotation curves of spiral galaxies provide the most direct evidence for the existence of large amounts of dark matter. Spiral galaxies consist of a central bulge and a very thin disk, stabilized against gravitational collapse by angular momentum conservation, and surrounded by an approximately spherical halo of dark matter. One can measure the orbital velocities of objects orbiting around the disk as a function of radius from the Doppler shifts of their spectral lines.

The rotation curve of the Andromeda galaxy was first measured by Babcock in 1938, from the stars in the disk. Later it became possible to measure galactic rotation curves far out into the disk, and a trend was found [39, 40]. The orbital velocity rose linearly from the center outward until it reached a typical value of 200 km/s, and then remained flat out to the largest measured radii. This was completely unexpected since the observed surface luminosity of the disk falls off exponentially with radius [39], $I(r) = I_0 \exp(-r/r_D)$. Therefore, one would expect that most of the galactic mass is concentrated within a few disk lengths r_D, such that the rotation velocity is determined as in a Keplerian orbit, $v_{\rm rot} = (GM/r)^{1/2} \propto r^{-1/2}$. No such behaviour is observed. In fact, the most convincing observations come from radio emission (from the 21 cm line) of neutral hydrogen in the disk, which has been measured to much larger galactic radii than optical tracers. A typical case is that of the spiral galaxy NGC 6503, where $r_D = 1.73$ kpc, while the furthest measured hydrogen line is at $r = 22.22$ kpc, about 13 disk lengths away. Nowadays, thousands of galactic rotation curves are known, see Fig. 14, and all suggest the existence of about ten times more mass in the halos of spiral galaxies than in the stars of the disk. The connection with dark matter halos was emphasized in Ref. [41]. Recent numerical simulations of galaxy formation in a CDM cosmology [42] suggest that galaxies probably formed by the infall of material in an overdense region of the universe that had decoupled from the overall expansion.

The dark matter is supposed to undergo violent relaxation and create a virialized system, i.e. in hydrostatic equilibrium. This picture has led to a simple model of dark-matter halos as isothermal spheres, with density profile $\rho(r) = \rho_c/(r_c^2 + r^2)$, where r_c is a core radius and $\rho_c = v_\infty^2/4\pi G$, with v_∞ equal to the plateau value of the flat rotation curve. This model is consistent with the universal rotation curves seen in Fig. 6. At large radii the dark matter distribution leads to a flat rotation curve. The question is for how long. In dense galaxy clusters one expects the galactic halos to overlap and form a continuum, and therefore the rotation curves should remain flat from one galaxy to another. However, in field galaxies, far from clusters, one can study the rotation velocities of substructures (like satellite dwarf galaxies) around a given galaxy, and determine whether they fall off at sufficiently large distances according to Kepler's law, as one would expect, once the edges of the dark matter halo have been reached. These observations are rather difficult because of uncertainties in distinguishing between true satellites and interlopers. Recently, a group from the Sloan Digital Sky Survey Collaboration claim that they have seen the edges of the dark matter halos around field galaxies by confirming the fall-off at large distances of their rotation curves [43]. These results, if corroborated by further analysis, would constitute a tremendous support to the idea of dark matter as a fluid surrounding galaxies and clusters, while at the same time eliminates the need for modifications of Newtonian of even Einstenian gravity at the scales of galaxies, to account for the flat rotation curves.

That's fine, but how much dark matter is there at the galactic scale? Adding up all the matter in galactic halos up to a maximum radii, one finds

$$\Omega_{\rm halo} \simeq 10 \, \Omega_{\rm lum} \geq 0.03 - 0.05 \,. \tag{110}$$

Of course, it would be extraordinary if we could confirm, through direct detection, the existence of dark
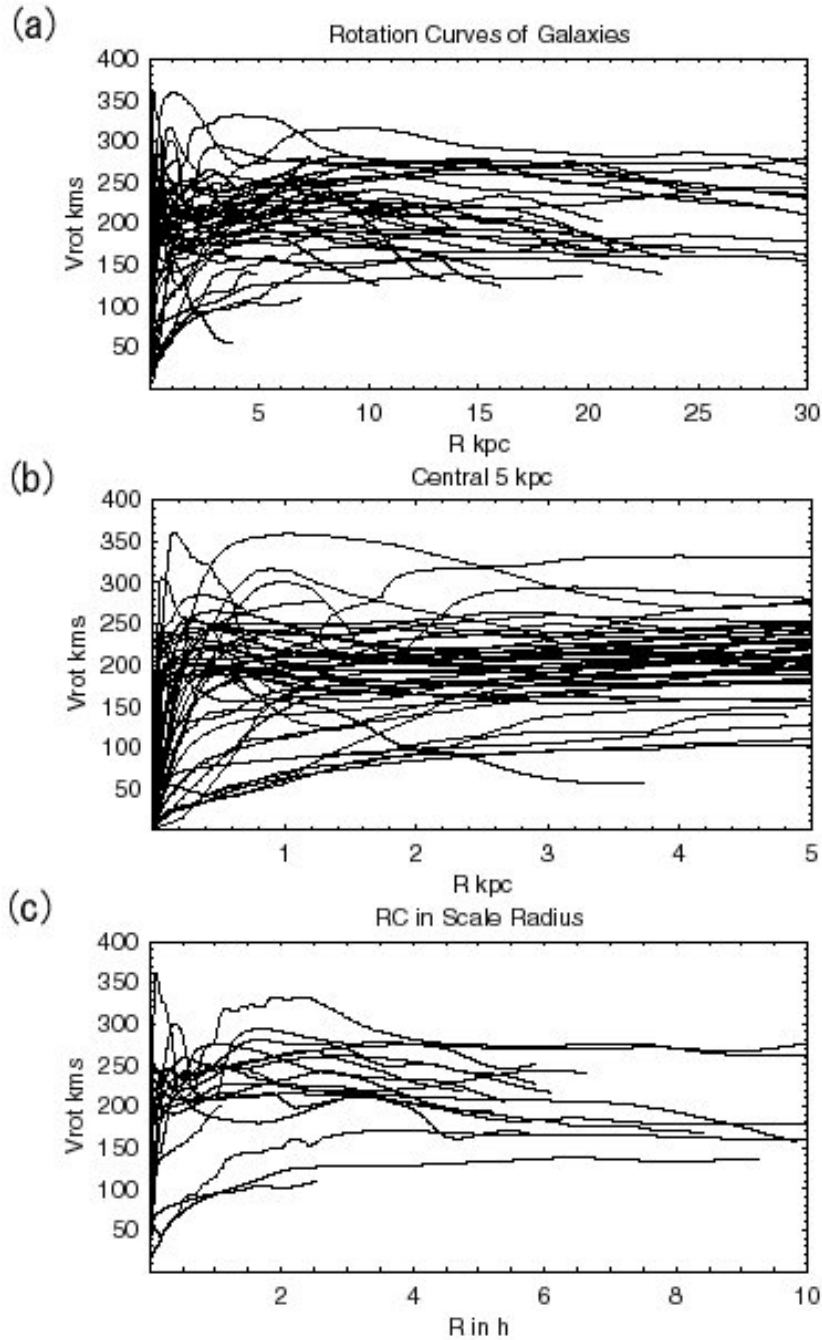
Fig. 13: The rotation curves of several hundred galaxies. Upper panel: As a function of their radii in kpc. Middle panel: The central 5 kpc. Lower panel: As a function of scale radius.

matter in our own galaxy. For that purpose, one should measure its rotation curve, which is much more difficult because of obscuration by dust in the disk, as well as problems with the determination of reliable galactocentric distances for the tracers. Nevertheless, the rotation curve of the Milky Way has been measured and conforms to the usual picture, with a plateau value of the rotation velocity of 220 km/s. For dark matter searches, the crucial quantity is the dark matter density in the solar neighbourhood, which turns out to be (within a factor of two uncertainty depending on the halo model) $\rho_{\mathrm{DM}} = 0.3$ GeV/cm^3. We will come back to direct searched of dark matter in a later subsection.

3.2.2 Baryon fraction in clusters

Since large clusters of galaxies form through gravitational collapse, they scoop up mass over a large volume of space, and therefore the ratio of baryons over the total matter in the cluster should be representative of the entire universe, at least within a 20% systematic error. Since the 1960s, when X-ray telescopes became available, it is known that galaxy clusters are the most powerful X-ray sources in the sky [44]. The emission extends over the whole cluster and reveals the existence of a hot plasma with temperature $T \sim 10^7 - 10^8$ K, where X-rays are produced by electron bremsstrahlung. Assuming the gas to be in hydrostatic equilibrium and applying the virial theorem one can estimate the total mass in the cluster, giving general agreement (within a factor of 2) with the virial mass estimates. From these estimates one can calculate the baryon fraction of clusters

$$f_\mathrm{B} h^{3/2} = 0.08 \quad \Rightarrow \quad \frac{\Omega_B}{\Omega_M} \approx 0.14\,, \quad \text{for} \quad h = 0.70\,. \tag{111}$$

Since $\Omega_\mathrm{lum} \simeq 0.002 - 0.006$, the previous expression suggests that clusters contain far more baryonic matter in the form of hot gas than in the form of stars in galaxies. Assuming this fraction to be representative of the entire universe, and using the Big Bang nucleosynthesis value of $\Omega_B = 0.04 \pm 0.01$, for $h = 0.7$, we find

$$\Omega_M = 0.3 \pm 0.1 \ (\text{statistical}) \ \pm 20\% \ (\text{systematic})\,. \tag{112}$$

This value is consistent with previous determinations of Ω_M. If some baryons are ejected from the cluster during gravitational collapse, or some are actually bound in nonluminous objects like planets, then the actual value of Ω_M is smaller than this estimate.

3.2.3 Weak gravitational lensing

Since the mid 1980s, deep surveys with powerful telescopes have observed huge arc-like features in galaxy clusters. The spectroscopic analysis showed that the cluster and the giant arcs were at very different redshifts. The usual interpretation is that the arc is the image of a distant background galaxy which is in the same line of sight as the cluster so that it appears distorted and magnified by the gravitational lens effect: the giant arcs are essentially partial Einstein rings. From a systematic study of the cluster mass distribution one can reconstruct the shear field responsible for the gravitational distortion [45]. This analysis shows that there are large amounts of dark matter in the clusters, in rough agreement with the virial mass estimates, although the lensing masses tend to be systematically larger. At present, the estimates indicate $\Omega_M = 0.2 - 0.3$ on scales $\lesssim 6\,h^{-1}$ Mpc.

3.2.4 Large scale structure formation and the matter power spectrum

Although the isotropic microwave background indicates that the universe in the *past* was extraordinarily homogeneous, we know that the universe *today* is far from homogeneous: we observe galaxies, clusters and superclusters on large scales. These structures are expected to arise from very small primordial inhomogeneities that grow in time via gravitational instability, and that may have originated from tiny ripples in the metric, as matter fell into their troughs. Those ripples must have left some trace as temperature anisotropies in the microwave background, and indeed such anisotropies were finally discovered by the COBE satellite in 1992. However, not all kinds of matter and/or evolution of the universe can give rise to the structure we observe today. If we define the density contrast as

$$\delta(\vec{x}, a) \equiv \frac{\rho(\vec{x}, a) - \bar{\rho}(a)}{\bar{\rho}(a)} = \int d^3\vec{k}\, \delta_k(a)\, e^{i\vec{k}\cdot\vec{x}}\,, \tag{113}$$

where $\bar{\rho}(a) = \rho_0\, a^{-3}$ is the average cosmic density, we need a theory that will grow a density contrast with amplitude $\delta \sim 10^{-5}$ at the last scattering surface ($z = 1100$) up to density contrasts of the order of
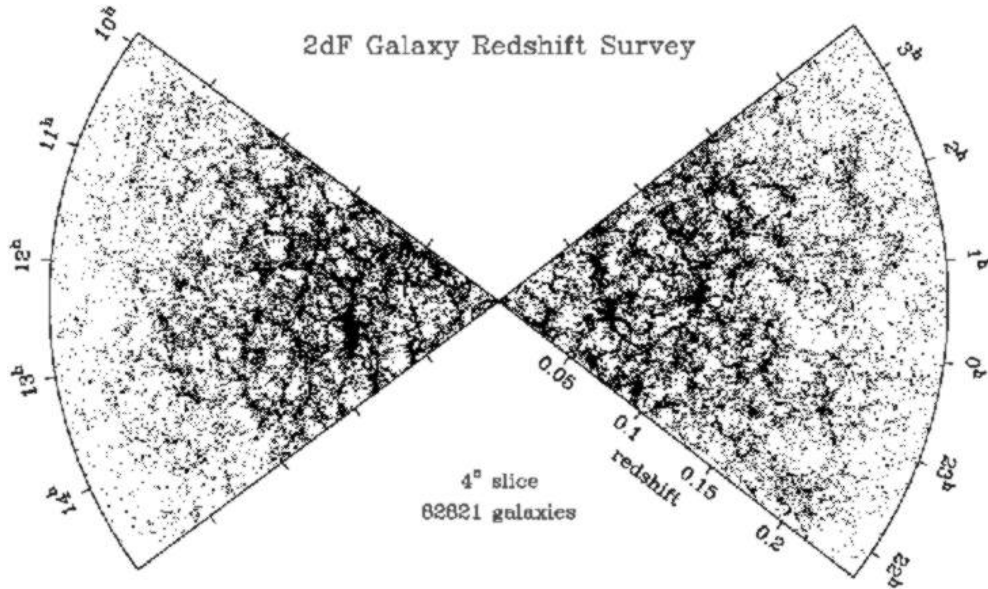
Fig. 14: The 2 degree Field Galaxy Redshift Survey contains some 250,000 galaxies, covering a large fraction of the sky up to redshifts of $z \leq 0.25$. From Ref. [46].

$\delta \sim 10^6$ for galaxies at redshifts $z \ll 1$, i.e. today. This is a *necessary* requirement for any consistent theory of structure formation.

Furthermore, the anisotropies observed by the COBE satellite correspond to a small-amplitude scale-invariant primordial power spectrum of inhomogeneities

$$P(k) = \langle |\delta_k|^2 \rangle \propto k^n \,, \quad \text{with} \quad n = 1 \,, \tag{114}$$

These inhomogeneities are like waves in the space-time metric. When matter fell in the troughs of those waves, it created density perturbations that collapsed gravitationally to form galaxies and clusters of galaxies, with a spectrum that is also scale invariant. Such a type of spectrum was proposed in the early 1970s by Edward R. Harrison, and independently by the Russian cosmologist Yakov B. Zel'dovich [27], to explain the distribution of galaxies and clusters of galaxies on very large scales in our observable universe, see Fig. 14.

Since the primordial spectrum is very approximately represented by a scale-invariant *Gaussian random field*, the best way to present the results of structure formation is by working with the 2-point correlation function in Fourier space, the so-called *power spectrum*. If the reprocessed spectrum of inhomogeneities remains Gaussian, the power spectrum is all we need to describe the galaxy distribution. Non-Gaussian effects are expected to arise from the non-linear gravitational collapse of structure, and may be important at small scales. The power spectrum measures the degree of inhomogeneity in the mass distribution on different scales, see Fig. 15. It depends upon a few basic ingredientes: a) the primordial spectrum of inhomogeneities, whether they are Gaussian or non-Gaussian, whether *adiabatic* (perturbations in the energy density) or *isocurvature* (perturbations in the entropy density), whether the primordial spectrum has *tilt* (deviations from scale-invariance), etc.; b) the recent creation of inhomogeneities, whether *cosmic strings* or some other topological defect from an early phase transition are responsible for the formation of structure today; and c) the cosmic evolution of the inhomogeneity, whether the universe has been dominated by cold or hot dark matter or by a cosmological constant since the beginning of structure formation, and also depending on the rate of expansion of the universe.

The working tools used for the comparison between the observed power spectrum and the predicted one are very precise N-body numerical simulations and theoretical models that predict the *shape*
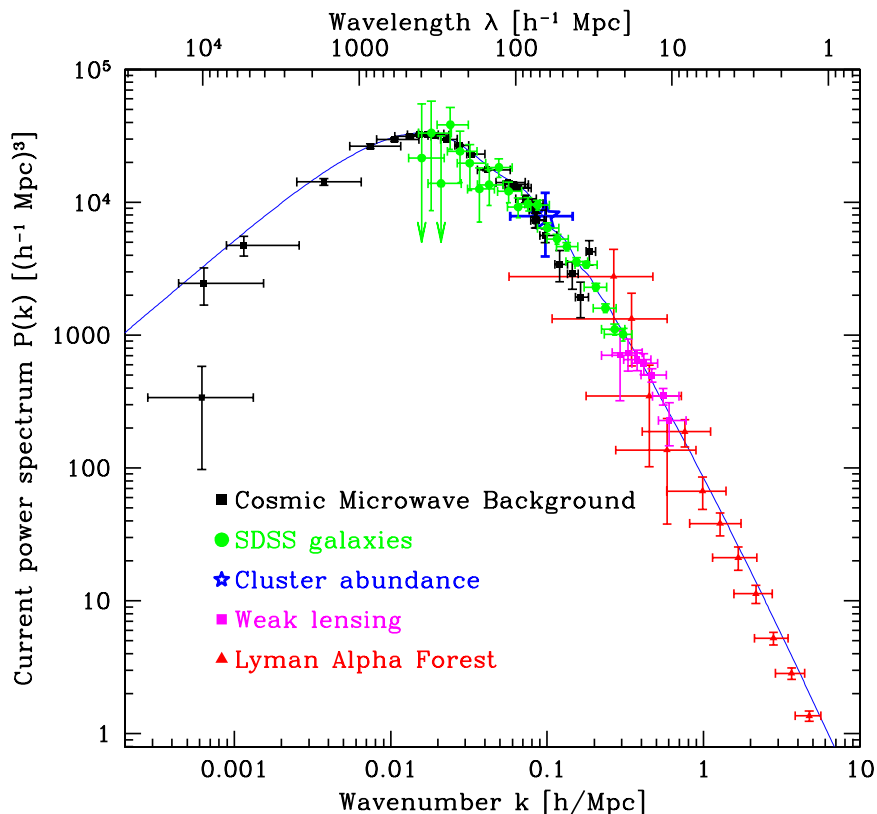
Fig. 15: The measured power spectrum $P(k)$ as a function of wavenumber k. From observations of the Sloan Digital Sky Survey, CMB anisotropies, cluster abundance, gravitational lensing and Lyman-α forest. From Ref. [47].

but not the *amplitude* of the present power spectrum. Even though a large amount of work has gone into those analyses, we still have large uncertainties about the nature and amount of matter necessary for structure formation. A model that has become a working paradigm is a flat cold dark matter model with a cosmological constant and $\Omega_M \sim 0.3$. This model is now been confronted with the recent very precise measurements from 2dFGRS [46] and SDSS [47].

3.2.5 The new redshift catalogs, 2dF and Sloan Digital Sky Survey

Our view of the large-scale distribution of luminous objects in the universe has changed dramatically during the last 25 years: from the simple pre-1975 picture of a distribution of field and cluster galaxies, to the discovery of the first single superstructures and voids, to the most recent results showing an almost regular web-like network of interconnected clusters, filaments and walls, separating huge nearly empty volumes. The increased efficiency of redshift surveys, made possible by the development of spectrographs and – specially in the last decade – by an enormous increase in multiplexing gain (i.e. the ability to collect spectra of several galaxies at once, thanks to fibre-optic spectrographs), has allowed us not only to do *cartography* of the nearby universe, but also to statistically characterize some of its properties. At the same time, advances in theoretical modeling of the development of structure, with large high-resolution gravitational simulations coupled to a deeper yet limited understanding of how to form galaxies within the dark matter halos, have provided a more realistic connection of the models to the observable quantities. Despite the large uncertainties that still exist, this has transformed the study of cosmology and large-scale structure into a truly quantitative science, where theory and observations can progress together.
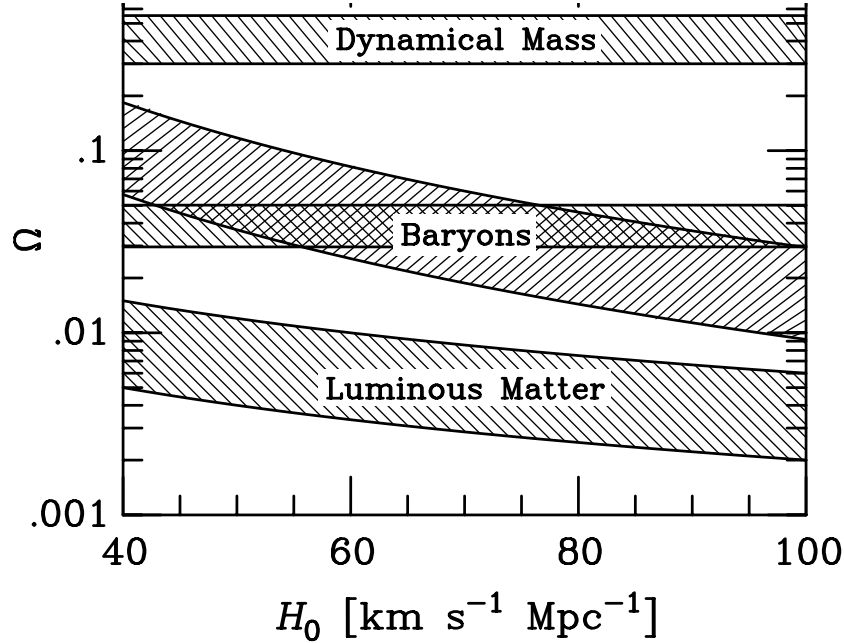
Fig. 16: The observed cosmic matter components as functions of the Hubble expansion parameter. The luminous matter component is given by $0.002 \leq \Omega_{\text{lum}} \leq 0.006$; the galactic halo component is the horizontal band, $0.03 \leq \Omega_{\text{halo}} \leq 0.05$, crossing the baryonic component from BBN, $\Omega_B h^2 = 0.0244 \pm 0.0024$; and the dynamical mass component from large scale structure analysis is given by $\Omega_M = 0.3 \pm 0.1$. Note that in the range $H_0 = 71 \pm 3$ km/s/Mpc, there are *three* dark matter problems, see the text. From Ref. [48].

3.2.6 Summary of the matter content

We can summarize the present situation with Fig. 16, for Ω_M as a function of H_0. There are four bands, the luminous matter Ω_{lum}; the baryon content Ω_B, from BBN; the galactic halo component Ω_{halo}, and the dynamical mass from clusters, Ω_M. From this figure it is clear that there are in fact *three* dark matter problems: The first one is where are 90% of the baryons? Between the fraction predicted by BBN and that seen in stars and diffuse gas there is a huge fraction which is in the form of dark baryons. They could be in small clumps of hydrogen that have not started thermonuclear reactions and perhaps constitute the dark matter of spiral galaxies' halos. Note that although Ω_B and Ω_{halo} coincide at $H_0 \simeq 70$ km/s/Mpc, this could be just a coincidence. The second problem is what constitutes 90% of matter, from BBN baryons to the mass inferred from cluster dynamics? This is the standard dark matter problem and could be solved in the future by direct detection of a weakly interacting massive particle in the laboratory. And finally, since we know from observations of the CMB that the universe is flat, the rest, up to $\Omega_0 = 1$, must be a diffuse vacuum energy, which affects the very large scales and late times, and seems to be responsible for the present acceleration of the universe, see Section 3. Nowadays, multiple observations seem to converge towards a common determination of $\Omega_M = 0.25 \pm 0.08$ (95% c.l.), see Fig. 17.

3.2.7 Massive neutrinos

One of the 'usual suspects' when addressing the problem of dark matter are neutrinos. They are the only candidates known to exist. If neutrinos have a mass, could they constitute the missing matter? We know from the Big Bang theory, see Section 2.6.5, that there is a cosmic neutrino background at a temperature of approximately 2K. This allows one to compute the present number density in the form of neutrinos, which turns out to be, for massless neutrinos, $n_\nu(T_\nu) = \frac{3}{11} n_\gamma(T_\gamma) = 112 \text{ cm}^{-3}$, per species of neutrino. Since neutrinos have mass, see Fig. 18, the cosmic energy density in massive neutrinos
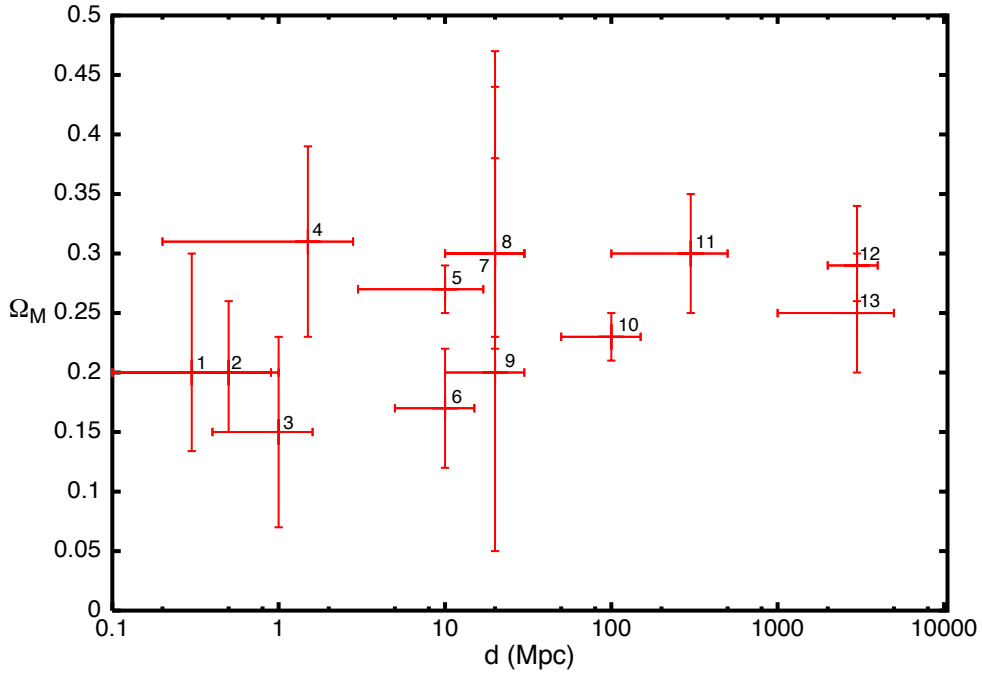
Fig. 17: Different determinations of Ω_M as a function of distance, from various sources: 1. peculiar velocities; 2. weak gravitacional lensing; 3. shear autocorrelation function; 4. local group of galaxies; 5. baryon mass fraction; 6. cluster mass function; 7. virgocentric flow; 8. mean relative velocities; 9. redshift space distortions; 10. mass power spectrum; 11. integrated Sachs-Wolfe effect; 12. angular diameter distance: SNe; 13. cluster baryon fraction. While a few years ago the dispersion among observed values was huge and strongly dependent on scale, at present the observed value of the matter density parameter falls well within a narrow range, $\Omega_M = 0.25 \pm 0.07$ (95% c.l.) and is essentially independent on scale, from 100 kpc to 5000 Mpc. Adapted from Ref. [49].

would be $\rho_\nu = \sum n_\nu m_\nu = \frac{3}{11} n_\gamma \sum m_\nu$, and therefore its contribution today,

$$\Omega_\nu h^2 = \frac{\sum m_\nu}{93.2 \text{ eV}} . \tag{115}$$

The discussion in the previous Sections suggest that $\Omega_{\rm M} \leq 0.4$, and thus, for any of the three families of neutrinos, $m_\nu \leq 40$ eV. Note that this limit improves by six orders of magnitude the present bound on the tau-neutrino mass [20]. Supposing that the missing mass in non-baryonic cold dark matter arises from a single particle dark matter (PDM) component, its contribution to the critical density is bounded by $0.05 \leq \Omega_{\rm PDM} h^2 \leq 0.4$, see Fig. 17.

I will now go through the various logical arguments that exclude neutrinos as the *dominant* component of the missing dark matter in the universe. Is it possible that neutrinos with a mass $4\,\text{eV} \leq m_\nu \leq 40$ eV be the non-baryonic PDM component? For instance, could massive neutrinos constitute the dark matter halos of galaxies? For neutrinos to be gravitationally bound to galaxies it is necessary that their velocity be less that the escape velocity $v_{\rm esc}$, and thus their maximum momentum is $p_{\max} = m_\nu v_{\rm esc}$. How many neutrinos can be packed in the halo of a galaxy? Due to the Pauli exclusion principle, the maximum number density is given by that of a completely degenerate Fermi gas with momentum $p_{\rm F} = p_{\max}$, i.e. $n_{\max} = p_{\max}^3/3\pi^2$. Therefore, the maximum local density in dark matter neutrinos is $\rho_{\max} = n_{\max} m_\nu = m_\nu^4 v_{\rm esc}^3/3\pi^2$, which must be greater than the typical halo density $\rho_{\rm halo} = 0.3\,\text{GeV cm}^{-3}$. For a typical spiral galaxy, this constraint, known as the Tremaine-Gunn limit, gives $m_\nu \geq 40$ eV, see Ref. [51]. However, this mass, even for a single species, say the tau-neutrino,
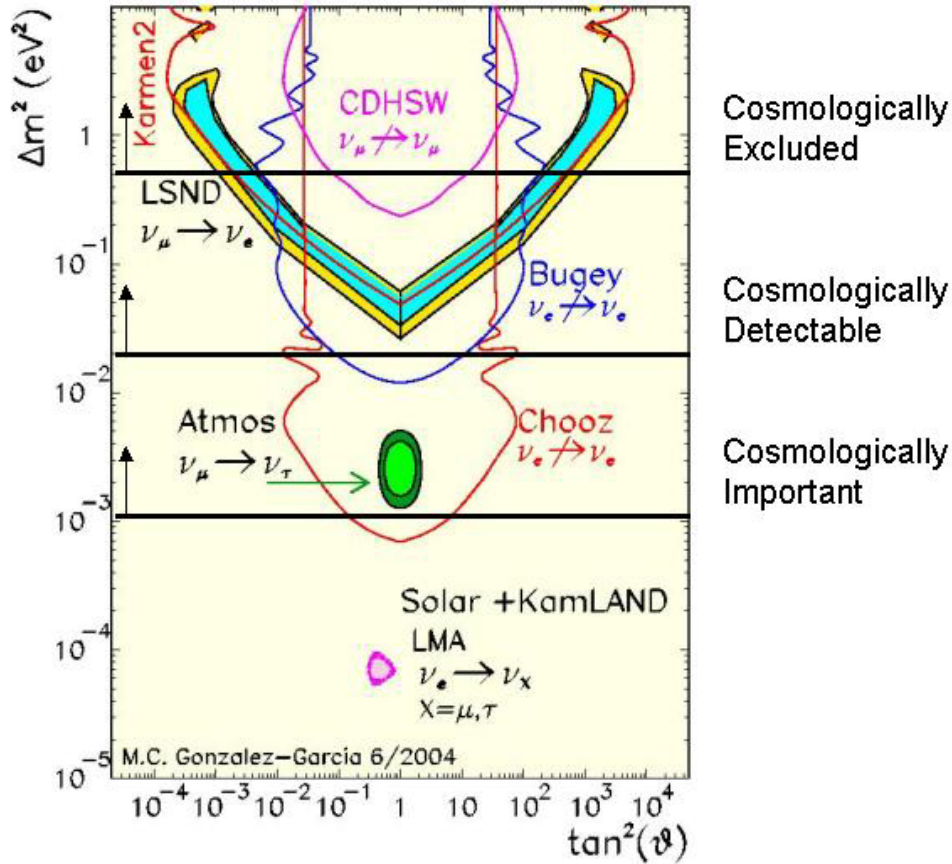
Fig. 18: The neutrino parameter space, mixing angle against Δm^2, including the results from the different solar and atmospheric neutrino oscillation experiments. Note the threshold of cosmologically important masses, cosmologically detectable neutrinos (by CMB and LSS observations), and cosmologically excluded range of masses. Adapted from Refs. [50] and [95].

gives a value for $\Omega_\nu h^2 = 0.5$, which is far too high for structure formation. Neutrinos of such a low mass would constitute a relativistic hot dark matter component, which would wash-out structure below the supercluster scale, against evidence from present observations, see Fig. 18. Furthermore, applying the same phase-space argument to the neutrinos as dark matter in the halo of dwarf galaxies gives $m_\nu \geq 100$ eV, beyond closure density (115). We must conclude that the simple idea that light neutrinos could constitute the particle dark matter on all scales is ruled out. They could, however, still play a role as a sub-dominant hot dark matter component in a flat CDM model. In that case, a neutrino mass of order 1 eV is not cosmological excluded, see Fig. 18.

Another possibility is that neutrinos have a large mass, of order a few GeV. In that case, their number density at decoupling, see Section 2.5.1, is suppressed by a Boltzmann factor, $\sim \exp(-m_\nu/T_{\text{dec}})$. For masses $m_\nu > T_{\text{dec}} \simeq 0.8$ MeV, the present energy density has to be computed as a solution of the corresponding Boltzmann equation. Apart from a logarithmic correction, one finds $\Omega_\nu h^2 \simeq 0.1(10\,\text{GeV}/m_\nu)^2$ for Majorana neutrinos and slightly smaller for Dirac neutrinos. In either case, neutrinos could be the dark matter only if their mass was a few GeV. Laboratory limits for ν_τ of around 18 MeV [20], and much more stringent ones for ν_μ and ν_e, exclude the known light neutrinos. However, there is always the possibility of a fourth unknown heavy and stable (perhaps sterile) neutrino. If it couples to the Z boson and has a mass below 45 GeV for Dirac neutrinos (39.5 GeV for Majorana neu-

trinos), then it is ruled out by measurements at LEP of the invisible width of the Z. There are two logical alternatives, either it is a sterile neutrino (it does not couple to the Z), or it does couple but has a larger mass. In the case of a Majorana neutrino (its own antiparticle), their abundance, for this mass range, is too small for being cosmologically relevant, $\Omega_\nu h^2 \leq 0.005$. If it were a Dirac neutrino there could be a lepton asymmetry, which may provide a higher abundance (similar to the case of baryogenesis). However, neutrinos scatter on nucleons via the weak axial-vector current (spin-dependent) interaction. For the small momentum transfers imparted by galactic WIMPs, such collisions are essentially coherent over an entire nucleus, leading to an enhancement of the effective cross section. The relatively large detection rate in this case allowes one to exclude fourth-generation Dirac neutrinos for the galactic dark matter [52]. Anyway, it would be very implausible to have such a massive neutrino today, since it would have to be stable, with a life-time greater than the age of the universe, and there is no theoretical reason to expect a massive sterile neutrino that does not oscillate into the other neutrinos.

Of course, the definitive test to the possible contribution of neutrinos to the overall density of the universe would be to measure *directly* their mass in laboratory experiments. There are at present two types of experiments: neutrino oscillation experiments, which measure only *differences* in squared masses, and direct mass-searches experiments, like the tritium β-spectrum and the neutrinoless double-β decay experiments, which measure directly the mass of the electron neutrino. The former experiments give a bound $m_{\nu_e} \lesssim 2.3$ eV (95% c.l.) [53], while the latter claim [54] they have a positive evidence for a Majorana neutrino of mass $m_\nu = 0.05 - 0.89$ eV (95% c.l.), although this result still awaits confirmation by other experiments. Neutrinos with such a mass could very well constitute the HDM component of the universe, $\Omega_{\rm HDM} \lesssim 0.15$. The oscillation experiments give a range of possibilities for $\Delta m_\nu^2 = 0.3 - 3$ eV^2 from LSND (not yet confirmed by Miniboone), to the atmospheric neutrino oscillations from SuperKamiokande ($\Delta m_\nu^2 \simeq 2.2 \pm 0.5 \times 10^{-3}$ eV^2, $\tan^2\theta = 1.0 \pm 0.3$) and the solar neutrino oscillations from KamLAND and the Sudbury Neutrino Observatory ($\Delta m_\nu^2 \simeq 8.2 \pm 0.3 \times 10^{-5}$ eV^2, $\tan^2\theta = 0.39 \pm 0.05$), see Ref. [50]. Only the first two possibilities would be cosmologically relevant, see Fig. 18. Thanks to recent observations by WMAP, SDSS and Planck, we can put stringent limits on the absolute scale of neutrino masses, see below (Section 3.4).

3.2.8 *Weakly Interacting Massive Particles*

Unless we drastically change the theory of gravity on large scales, baryons cannot make up the bulk of the dark matter. Massive neutrinos are the only alternative among the known particles, but they are essentially ruled out as a universal dark matter candidate, even if they may play a subdominant role as a hot dark matter component. There remains the mystery of what is the physical nature of the dominant cold dark matter component. Something like a heavy stable neutrino, a generic Weakly Interacting Massive Particle (WIMP), could be a reasonable candidate because its present abundance could fall within the expected range,

$$\Omega_{\rm PDM} h^2 \sim \frac{G^{3/2} T_0^3 h^2}{H_0^2 \langle \sigma_{\rm ann} v_{\rm rel}\rangle} = \frac{3 \times 10^{-27}\ {\rm cm^3 s^{-1}}}{\langle \sigma_{\rm ann} v_{\rm rel}\rangle}\,. \tag{116}$$

Here $v_{\rm rel}$ is the relative velocity of the two incoming dark matter particles and the brackets $\langle \cdot \rangle$ denote a thermal average at the freeze-out temperature, $T_{\rm f} \simeq m_{\rm PDM}/20$, when the dark matter particles go out of equilibrium with radiation. The value of $\langle \sigma_{\rm ann} v_{\rm rel}\rangle$ needed for $\Omega_{\rm PDM} \approx 1$ is remarkably close to what one would expect for a WIMP with a mass $m_{\rm PDM} = 100$ GeV, $\langle \sigma_{\rm ann} v_{\rm rel}\rangle \sim \alpha^2/8\pi\, m_{\rm PDM} \sim 3 \times 10^{-27}$ cm^{3}s^{-1}. We still do not know whether this is just a coincidence or an important hint on the nature of dark matter.

There are a few theoretical candidates for WIMPs, like the neutralino, coming from supersymmetric extensions of the standard model of particle physics,[6] but at present there is no empirical evidence that

---

[6] For a review of Supersymmetry (SUSY), see Kazakov's contribution to these Proceedings.
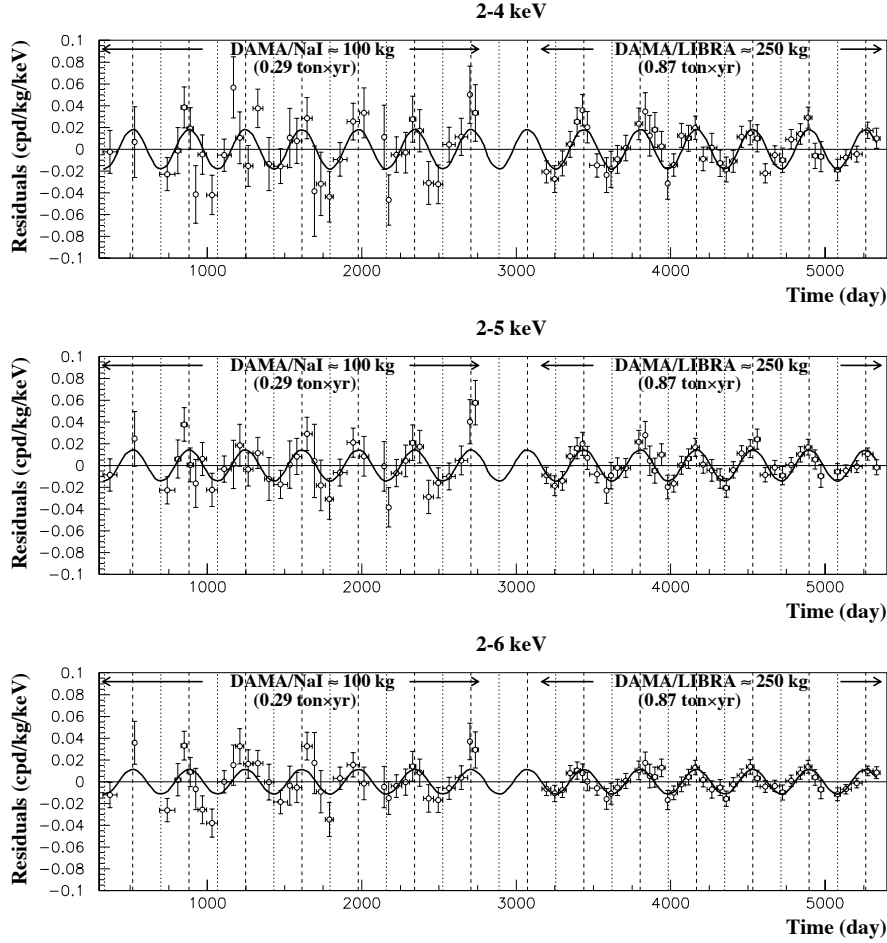
Fig. 19: The annual-modulation signal accumulated over 12 years is consistent with a neutralino of mass of $m_\chi = 59\ ^{+17}_{-14}$ GeV and a proton cross section of $\xi\sigma_p = 7.0\ ^{+0.4}_{-1.2} \times 10^{-6}$ pb, according to DAMA. From Ref. [55].

such extensions are indeed realized in nature. In fact, the non-observation of supersymmetric particles at current accelerators places stringent limits on the neutralino mass and interaction cross section [56]. If WIMPs constitute the dominant component of the halo of our galaxy, it is expected that some may cross the Earth at a reasonable rate to be detected. The direct experimental search for them rely on elastic WIMP collisions with the nuclei of a suitable target. Dark matter WIMPs move at a typical galactic "virial" velocity of around $200 - 300$ km/s, depending on the model. If their mass is in the range $10 - 100$ GeV, the recoil energy of the nuclei in the elastic collision would be of order 10 keV. Therefore, one should be able to identify such energy depositions in a macroscopic sample of the target. There are at present three different methods: First, one could search for scintillation light in NaI crystals or in liquid xenon; second, search for an ionization signal in a semiconductor, typically a very pure germanium crystal; and third, use a cryogenic detector at 10 mK and search for a measurable temperature increase of the sample. The main problem with such a type of experiment is the low expected signal rate, with a typical number below 1 event/kg/day. To reduce natural radioactive contamination one must use extremely pure substances, and to reduce the background caused by cosmic rays requires that these experiments be located deeply underground.

The best limits on WIMP scattering cross sections come from some germanium experiments, like the Criogenic Dark Matter Search (CDMS) collaboration at Stanford and the Soudan mine [57], as well as from the NaI scintillation detectors of the UK dark matter collaboration (UKDMC) in the Boulby salt

mine in England [58], and the DAMA experiment in the Gran Sasso laboratory in Italy [55]. Current experiments already touch the parameter space expected from supersymmetric particles, see Fig. 20, and therefore there is a chance that they actually discover the nature of the missing dark matter. The problem, of course, is to attribute a tentative signal unambiguously to galactic WIMPs rather than to some unidentified radioactive background.
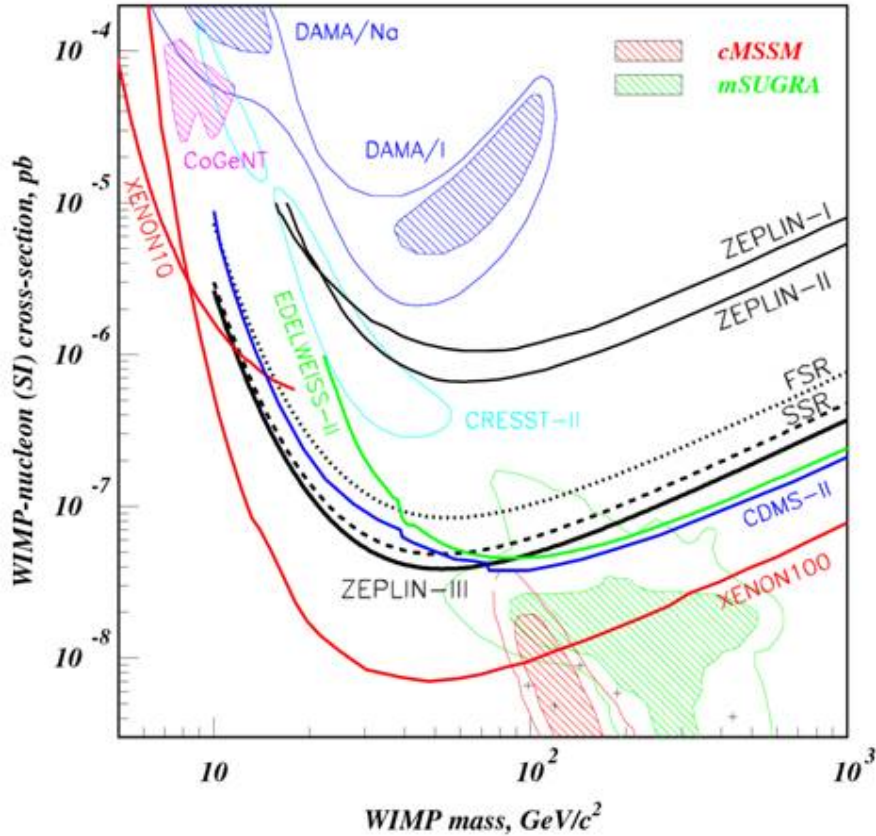
Fig. 20: Exclusion range for the spin-independent WIMP scattering cross section per nucleon from the NaI experiments and the Ge detectors. The blue lines come from the CDMS experiment, which exclude the DAMA region at more than 3 sigma. Also shown in yellow and red is the range of expected counting rates for neutralinos in the MSSM. From Ref. [57].

One specific signature is the annual modulation which arises as the Earth moves around the Sun.[7] Therefore, the net speed of the Earth relative to the galactic dark matter halo varies, causing a modulation of the expected counting rate. The DAMA/NaI experiment has actually reported such a modulation signal, from the combined analysis of their 12-year data, see Fig. 19 and Ref. [55], which provides a confidence level of 99.6% for a neutralino mass of $m_\chi = 52\,^{+10}_{-8}$ GeV and a proton cross section of $\xi\sigma_p = 7.2\,^{+0.4}_{-0.9} \times 10^{-6}$ pb, where $\xi = \rho_\chi/0.3$ GeV cm^{-3} is the local neutralino energy density in units of the galactic halo density. There has been no confirmation yet of this result from other dark matter search groups. In fact, the CDMS collaboration claims an exclusion of the DAMA region at the 3 sigma level, see Fig. 20. Hopefully in the near future we will have much better sensitivity at low masses from the Cryogenic Rare Event Search with Superconducting Thermometers (CRESST) experiment at Gran Sasso. The CRESST experiment [59] uses sapphire crystals as targets and a new

---

[7]The time scale of the Sun's orbit around the center of the galaxy is too large to be relevant in the analysis.

method to simultaneously measure the phonons and the scintillating light from particle interactions inside the crystal, which allows excellent background discrimination. Very recently there has been also the proposal of a completely new method based on a Superheated Droplet Detector (SDD), which claims to have already a similar sensitivity as the more standard methods described above, see Ref. [60].

There exist other *indirect* methods to search for galactic WIMPs [61]. Such particles could self-annihilate at a certain rate in the galactic halo, producing a potentially detectable background of high energy photons or antiprotons. The absence of such a background in both gamma ray satellites and the Alpha Matter Spectrometer [62] imposes bounds on their density in the halo. Alternatively, WIMPs traversing the solar system may interact with the matter that makes up the Earth or the Sun so that a small fraction of them will lose energy and be trapped in their cores, building up over the age of the universe. Their annihilation in the core would thus produce high energy neutrinos from the center of the Earth or from the Sun which are detectable by neutrino telescopes. In fact, SuperKamiokande already covers a large part of SUSY parameter space. In other words, neutrino telescopes are already competitive with direct search experiments. In particular, the AMANDA experiment at the South Pole [63], which has approximately 10^3 Cherenkov detectors several km deep in very clear ice, over a volume $\sim 1 \text{ km}^3$, is competitive with the best direct searches proposed. The advantages of AMANDA are also directional, since the arrays of Cherenkov detectors will allow one to reconstruct the neutrino trajectory and thus its source, whether it comes from the Earth or the Sun. AMANDA recently reported the detection of TeV neutrinos [63].

3.3   The age of the universe t_0

The universe must be older than the oldest objects it contains. Those are believed to be the stars in the oldest clusters in the Milky Way, globular clusters. The most reliable ages come from the application of theoretical models of stellar evolution to observations of old stars in globular clusters. For about 30 years, the ages of globular clusters have remained reasonable stable, at about 15 Gyr [64]. However, recently these ages have been revised downward [65].

During the 1980s and 1990s, the globular cluster age estimates have improved as both new observations have been made with CCDs, and since refinements to stellar evolution models, including opacities, consideration of mixing, and different chemical abundances have been incorporated [66]. From the theory side, uncertainties in globular cluster ages come from uncertainties in convection models, opacities, and nuclear reaction rates. From the observational side, uncertainties arise due to corrections for dust and chemical composition. However, the dominant source of systematic errors in the globular cluster age is the uncertainty in the cluster distances. Fortunately, the Hipparcos satellite recently provided geometric parallax measurements for many nearby old stars with low metallicity, typical of glubular clusters, thus allowing for a new calibration of the ages of stars in globular clusters, leading to a downward revision to $10 - 13$ Gyr [66]. Moreover, there were very few stars in the Hipparcos catalog with both small parallax erros and low metal abundance. Hence, an increase in the sample size could be critical in reducing the statatistical uncertaintites for the calibration of the globular cluster ages. There are already proposed two new parallax satellites, NASA's Space Interferometry Mission (SIM) and ESA's mission, called GAIA, that will give 2 or 3 orders of magnitude more accurate parallaxes than Hipparcos, down to fainter magnitude limits, for several orders of magnitude more stars. Until larger samples are available, however, distance errors are likely to be the largest source of systematic uncertainty to the globular cluster age [30].

The supernovae groups can also determine the age of the universe from their high redshift observations. The high confidence regions in the $(\Omega_\text{M}, \Omega_\Lambda)$ plane are almost parallel to the contours of constant age. For any value of the Hubble constant less than $H_0 = 70$ km/s/Mpc, the implied age of the universe is greater than 13 Gyr, allowing enough time for the oldest stars in globular clusters to evolve [66]. Integrating over Ω_M and Ω_Λ, the best fit value of the age in Hubble-time units is $H_0 t_0 = 0.93 \pm 0.06$ or equivalently $t_0 = 14.1 \pm 1.0 \, (0.65 \, h^{-1})$ Gyr, see Ref. [7]. Furthermore, a combination of 8 independent

recent measurements: CMB anisotropies, type Ia SNe, cluster mass-to-light ratios, cluster abundance evolution, cluster baryon fraction, deuterium-to-hidrogen ratios in quasar spectra, double-lobed radio sources and the Hubble constant, can be used to determine the present age of the universe [67]. The best fit value for the age of the universe is, according to this analysis, $t_0 = 13.4 \pm 1.6$ Gyr, about a billion years younger than other recent estimates [67].
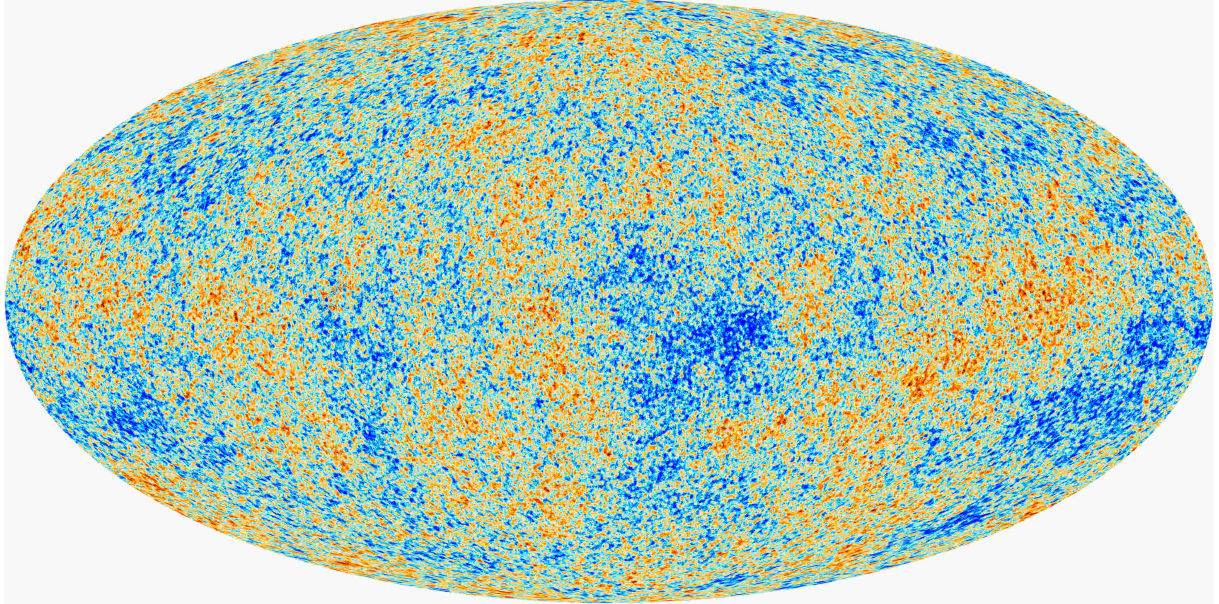

Fig. 21: The anisotropies of the microwave background measured by the Planck satellite with 4 arcminute resolution. It shows the intrinsic CMB anisotropies at the level of a few parts in 10^5. The galactic foreground has been properly subtracted. The amount of information contained in this map is enough to determine most of the cosmological parameters to few percent accuracy. From Ref. [93].

3.4 Cosmic Microwave Background Anisotropies

The cosmic microwave background has become in the last five years the Holy Grail of Cosmology, since precise observations of the temperature and polarization anisotropies allow in principle to determine the parameters of the Standard Model of Cosmology with very high accuracy. Recently, the WMAP satellite has provided with a very detailed map of the microwave anisotropies in the sky, see Fig. 21, and indeed has fulfilled our expectations, see Table 2.

The physics of the CMB anisotropies is relatively simple [68]. The universe just before recombination is a very tightly coupled fluid, due to the large electromagnetic Thomson cross section $\sigma_T = 8\pi\alpha^2/3m_e^2 \simeq 0.7$ barn. Photons scatter off charged particles (protons and electrons), and carry energy, so they feel the gravitational potential associated with the perturbations imprinted in the metric during inflation. An overdensity of baryons (protons and neutrons) does not collapse under the effect of gravity until it enters the causal Hubble radius. The perturbation continues to grow until radiation pressure opposes gravity and sets up acoustic oscillations in the plasma, very similar to sound waves. Since overdensities of the same size will enter the Hubble radius at the same time, they will oscillate in phase. Moreover, since photons scatter off these baryons, the acoustic oscillations occur also in the photon field and induces a pattern of peaks in the temperature anisotropies in the sky, at different angular scales, see Fig. 22. There are three different effects that determine the temperature anisotropies we observe in the CMB. First, *gravity*: photons fall in and escape off gravitational potential wells, characterized by Φ in the comoving gauge, and as a consequence their frequency is gravitationally blue- or red-shifted, $\delta\nu/\nu = \Phi$. If the gravitational potential is not constant, the photons will escape from a larger or smaller potential
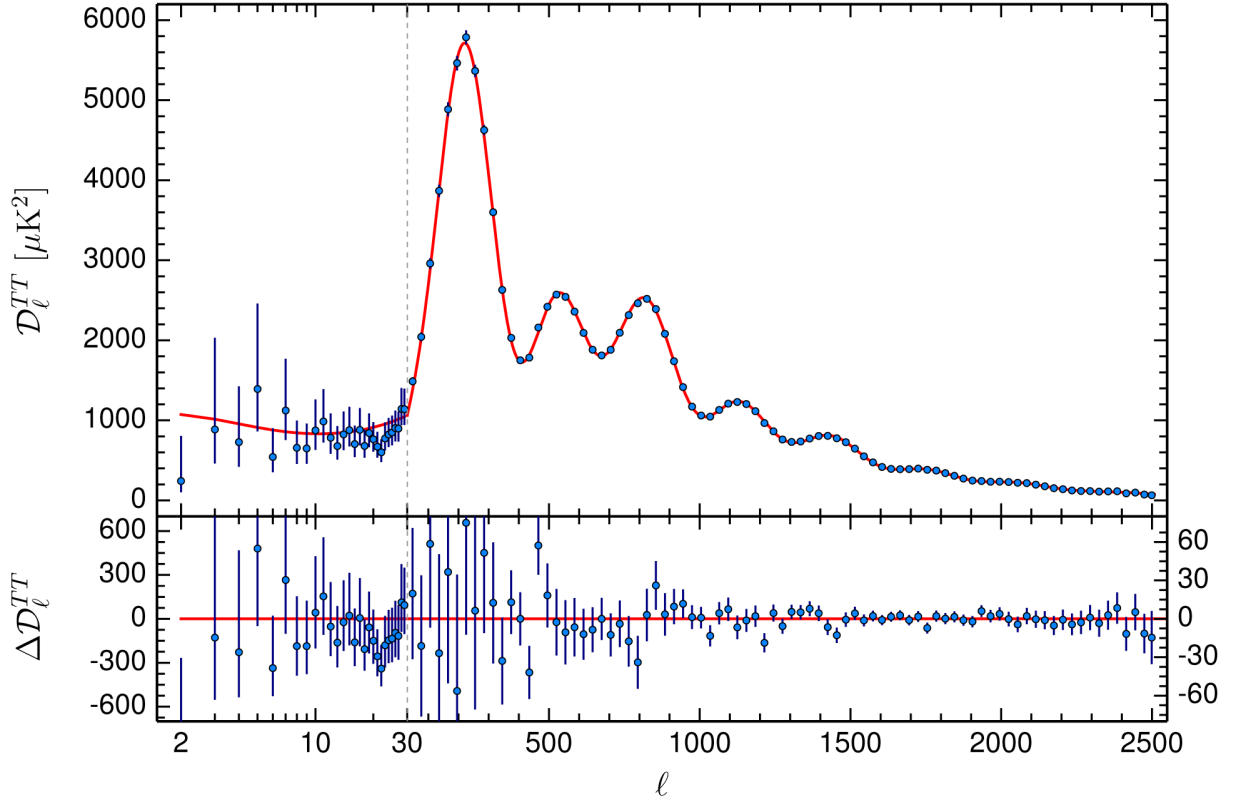
Fig. 22: The Angular Power Spectrum of CMB temperature anisotropies, $\mathcal{D}_l = l(l+1)\mathcal{C}_l/2\pi$, and their residuals w.r.t. a Standard Cosmological Model with just six parameters. The low ($l < 30$) multipoles are not used in the fit. The agreement is impressive, except for a small anomalous deviation around $l \simeq 20$. From Ref. [93].

well than they fell in, so their frequency is also blue- or red-shifted, a phenomenon known as the Rees-Sciama effect. Second, *pressure*: photons scatter off baryons which fall into gravitational potential wells and the two competing forces create acoustic waves of compression and rarefaction. Finally, *velocity*: baryons accelerate as they fall into potential wells. They have minimum velocity at maximum compression and rarefaction. That is, their velocity wave is exactly 90° off-phase with the acoustic waves. These waves induce a Doppler effect on the frequency of the photons. The temperature anisotropy induced by these three effects is therefore given by [68]

$$\frac{\delta T}{T}(\mathbf{r}) = \Phi(\mathbf{r}, t_{\text{dec}}) + 2 \int_{t_{\text{dec}}}^{t_0} \dot{\Phi}(\mathbf{r}, t)dt + \frac{1}{3}\frac{\delta\rho}{\rho} - \frac{\mathbf{r} \cdot \mathbf{v}}{c}\,. \tag{117}$$

Metric perturbations of different wavelengths enter the horizon at different times. The largest wavelengths, of size comparable to our present horizon, are entering now. There are perturbations with wavelengths comparable to the size of the horizon at the time of last scattering, of projected size about 1° in the sky today, which entered precisely at decoupling. And there are perturbations with wavelengths much smaller than the size of the horizon at last scattering, that entered much earlier than decoupling, all the way to the time of radiation-matter equality, which have gone through several acoustic oscillations before last scattering. All these perturbations of different wavelengths leave their imprint in the CMB anisotropies.

The baryons at the time of decoupling do not feel the gravitational attraction of perturbations with wavelength greater than the size of the horizon at last scattering, because of causality. Perturbations with exactly that wavelength are undergoing their first contraction, or acoustic compression, at decoupling. Those perturbations induce a large peak in the temperature anisotropies power spectrum, see Fig. 22.

Table 2: The parameters of the Standard Cosmological Model. The standard model of cosmology has about 20 different parameters, needed to describe the background space-time, the matter content and the spectrum of metric perturbations. We include here the present range of the most relevant parameters (with 1σ errors), as determined by both WMAP and Planck satellites. The rate of expansion is written in units of $H = 100\,h$ km/s/Mpc.

physical quantity	symbol	WMAP	Planck
total density	Ω_0	1.02 ± 0.02	0.9992 ± 0.0040
baryonic matter	$\Omega_{\rm B} h^2$	0.0228 ± 0.0020	0.02230 ± 0.00014
cosmological constant	Ω_Λ	0.73 ± 0.04	0.6911 ± 0.0062
cold dark matter	$\Omega_{\rm M}$	0.23 ± 0.04	0.3089 ± 0.0062
hot dark matter	$\Omega_\nu h^2$	< 0.00276 (95% c.l.)	< 0.0021 (95% c.l.)
number relativistic species	$N_{\rm eff}$	3.84 ± 0.40	3.04 ± 0.33
sum of neutrino masses	$\sum m_\nu$ (eV)	< 0.23 (95% c.l.)	< 0.194 (95% c.l.)
CMB temperature	$T_0\,(K)$	2.725 ± 0.002	2.7235 ± 0.0012
baryon to photon ratio	$10^{10}\eta$	6.1 ± 0.3	6.09 ± 0.04
baryon to matter ratio	$\Omega_{\rm B}/\Omega_{\rm M}$	0.17 ± 0.01	0.157 ± 0.002
spatial curvature	Ω_K	< 0.02 (95% c.l.)	0.0008 ± 0.0040
rate of expansion	h	0.71 ± 0.03	0.6774 ± 0.0046
age of the universe	t_0 (Gyr)	13.7 ± 0.6	13.799 ± 0.021
age at decoupling	$t_{\rm dec}$ (kyr)	379 ± 8	380.000 ± 0.120
age at reionization	$t_{\rm r}$ (Myr)	180 ± 100	212 ± 43
spectral amplitude	$10^{10}A_s^2$	17.35 ± 3.72	21.30 ± 0.53
spectral tilt	$n_{\rm s}$	0.98 ± 0.03	0.9667 ± 0.0040
spectral tilt variation	$dn_{\rm s}/d\ln k$	-0.031 ± 0.017	-0.002 ± 0.013
tensor-scalar ratio	r	< 0.71 (95% c.l.)	< 0.113 (95% c.l.)
reionization optical depth	τ	0.17 ± 0.04	0.06 ± 0.012
redshift of equality	$z_{\rm eq}$	3233 ± 200	3371 ± 23
redshift of decoupling	$z_{\rm dec}$	1089 ± 1	1089.90 ± 0.23
width of decoupling	$\Delta z_{\rm dec}$	195 ± 2	144.81 ± 0.24
redshift of reionization	$z_{\rm r}$	20 ± 10	8.8 ± 1.2

Perturbations with wavelengths smaller than these will have gone, after they entered the Hubble scale, through a series of acoustic compressions and rarefactions, which can be seen as secondary peaks in the power spectrum. Since the surface of last scattering is not a sharp discontinuity, but a region of $\Delta z \sim 100$, there will be scales for which photons, traveling from one energy concentration to another, will erase the perturbation on that scale, similarly to what neutrinos or HDM do for structure on small scales. That is the reason why we don't see all the acoustic oscillations with the same amplitude, but in fact they decay exponentially towards smaller angular scales, an effect known as Silk damping, due to photon diffusion [68, 69].

From the observations of the CMB anisotropies it is possible to determine most of the parameters of the Standard Cosmological Model with few percent accuracy, see Table 2. However, there are many degeneracies between parameters and it is difficult to disentangle one from another. For instance, as mentioned above, the first peak in the photon distribution corresponds to overdensities that have undergone half an oscillation, that is, a compression, and appear at a scale associated with the size of the horizon at last scattering, about $1°$ projected in the sky today. Since photons scatter off baryons, they will also feel the acoustic wave and create a peak in the correlation function. The height of the peak is proportional to the amount of baryons: the larger the baryon content of the universe, the higher the

peak. The position of the peak in the power spectrum depends on the geometrical size of the particle horizon at last scattering. Since photons travel along geodesics, the projected size of the causal horizon at decoupling depends on whether the universe is flat, open or closed. In a flat universe the geodesics are straight lines and, by looking at the angular scale of the first acoustic peak, we would be measuring the actual size of the horizon at last scattering. In an open universe, the geodesics are inward-curved trajectories, and therefore the projected size on the sky appears smaller. In this case, the first acoustic peak should occur at higher multipoles or smaller angular scales. On the other hand, for a closed universe, the first peak occurs at smaller multipoles or larger angular scales. The dependence of the position of the first acoustic peak on the spatial curvature can be approximately given by $l_{\text{peak}} \simeq 220 \, \Omega_0^{-1/2}$, where $\Omega_0 = \Omega_{\text{M}} + \Omega_{\Lambda} = 1 - \Omega_K$. Present observations by WMAP and other experiments give $\Omega_0 = 1.00 \pm 0.02$ at one standard deviation [21].
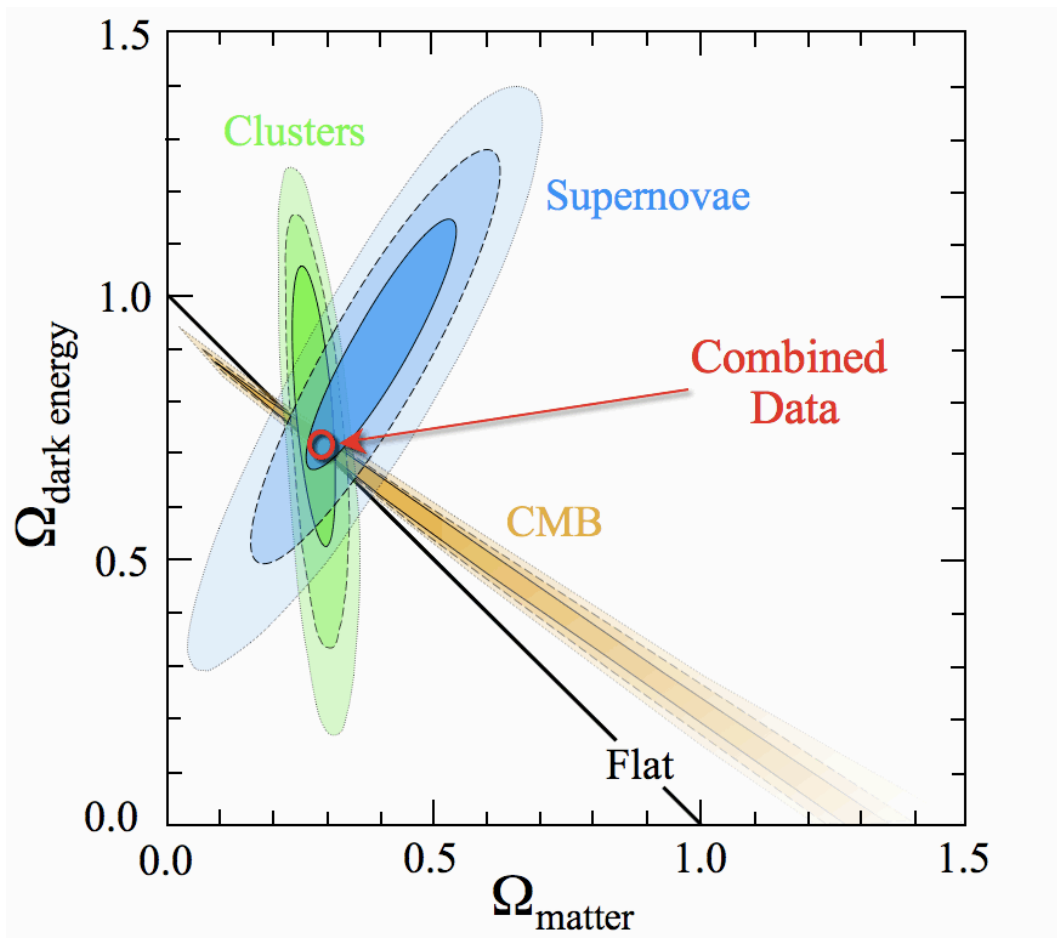
Fig. 23: The $(\Omega_M, \Omega_\Lambda)$ plane with the present data set of cosmological observations $-$ the acceleration of the universe, the large scale structure and the CMB anisotropies $-$ on the fundamental parameters which define our Standard Model of Cosmology.

The other acoustic peaks occur at harmonics of this, corresponding to smaller angular scales. Since the amplitude and position of the primary and secondary peaks are directly determined by the sound speed (and, hence, the equation of state) and by the geometry and expansion of the universe, they can be used as a powerful test of the density of baryons and dark matter, and other cosmological parameters. With the joined data from WMAP, VSA, CBI and ACBAR, we have rather good evidence of the existence of the second and third acoustic peaks, which confirms one of the most important predictions of inflation $-$ the non-causal origin of the primordial spectrum of perturbations $-$, and rules out cosmological defects

as the dominant source of structure in the universe [70]. Moreover, since the observations of CMB anisotropies now cover almost three orders of magnitude in the size of perturbations, we can determine the much better accuracy the value of the spectral tilt, $n = 0.98 \pm 0.03$, which is compatible with the approximate scale invariant spectrum needed for structure formation, and is a prediction of the simplest models of inflation. Soon after the release of data from WMAP, there was some expectation at the claim of a scale-dependent tilt. Nowadays, with better resolution in the linear matter power spectrum from SDSS [71], we can not conclude that the spectral tilt has any observable dependence on scale.

The microwave background has become also a testing ground for theories of particle physics. In particular, it already gives stringent constraints on the mass of the neutrino, when analysed together with large scale structure observations. Assuming a flat ΛCDM model, the 2-sigma upper bounds on the sum of the masses of light neutrinos is $\sum m_\nu < 1.0$ eV for degenerate neutrinos (i.e. without a large hierachy between them) if we don't impose any priors, and it comes down to $\sum m_\nu < 0.6$ eV if one imposes the bounds coming from the HST measurements of the rate of expansion and the supernova data on the present acceleration of the universe [72]. The final bound on the neutrino density can be expressed as $\Omega_\nu h^2 = \sum m_\nu / 93.2$ eV ≤ 0.01. In the future, both with Planck and with the Atacama Cosmology Telescope (ACT) we will be able to put constraints on the neutrino masses down to the 0.1 eV level.

Moreover, the present data is good enough that we can start to put constraints on the models of inflation that give rise to structure. In particular, multifield models of inflation predict a mixture of adiabatic and isocurvature perturbations,[8] and their signatures in the cosmic microwave background anisotropies and the matter power spectrum of large scale structure are specific and perfectly distinguishable. Nowadays, thanks to precise CMB, LSS and SNIa data, one can put rather stringent limits on the relative fraction and correlation of the isocurvature modes to the dominant adiabatic perturbations [73].

We can summarize this Section by showing the region in parameter space where we stand nowadays, thanks to the recent cosmological observations. We have plotted that region in Fig. 23. One could also superimpose the contour lines corresponding to equal $t_0 H_0$ lines, as a cross check. It is extraordinary that only in the last few months we have been able to reduce the concordance region to where it stands today, where all the different observations seem to converge. There are still many uncertainties, mainly systematic; however, those are quickly decreasing and becoming predominantly statistical. In the near future, with precise observations of the anisotropies in the microwave background temperature and polarization anisotropies, thanks to Planck satellite, we will be able to reduce those uncertainties to the level of one percent. This is the reason why cosmologists are so excited and why it is claimed that we live in the Golden Age of Cosmology.

4 The Inflationary Paradigm

The hot Big Bang theory is nowadays a very robust edifice, with many independent observational checks: the expansion of the universe; the abundance of light elements; the cosmic microwave background; a predicted age of the universe compatible with the age of the oldest objects in it, and the formation of structure via gravitational collapse of initially small inhomogeneities. Today, these observations are confirmed to within a few percent accuracy, and have helped establish the hot Big Bang as the preferred model of the universe. All the physics involved in the above observations is routinely tested in the laboratory (atomic and nuclear physics experiments) or in the solar system (general relativity).

However, this theory leaves a range of crucial questions unanswered, most of which are initial conditions' problems. There is the reasonable assumption that these cosmological problems will be solved or explained by *new physical principles* at high energies, in the early universe. This assumption leads to the natural conclusion that accurate observations of the present state of the universe may shed light onto processes and physical laws at energies above those reachable by particle accelerators, present

---

[8]This mixture is generic, unless all the fields thermalize simultaneously at reheating, just after inflation, in which case the entropy perturbations that would give rise to the isocurvature modes disappear.

or future. We will see that this is a very optimistic approach indeed, and that there are many unresolved issues related to those problems. However, there might be in the near future reasons to be optimistic.

4.1 Shortcomings of Big Bang Cosmology

The Big Bang theory could not explain the origin of matter and structure in the universe; that is, the origin of the matter–antimatter asymmetry, without which the universe today would be filled by a uniform radiation continuosly expanding and cooling, with no traces of matter, and thus without the possibility to form gravitationally bound systems like galaxies, stars and planets that could sustain life. Moreover, the standard Big Bang theory assumes, but cannot explain, the origin of the extraordinary smoothness and flatness of the universe on the very large scales seen by the microwave background probes and the largest galaxy catalogs. It cannot explain the origin of the primordial density perturbations that gave rise to cosmic structures like galaxies, clusters and superclusters, via gravitational collapse; the quantity and nature of the dark matter that we believe holds the universe together; nor the origin of the Big Bang itself.

A summary [10] of the problems that the Big Bang theory cannot explain is:

- The global structure of the universe.
  - Why is the universe so close to spatial flatness?
  - Why is matter so homogeneously distributed on large scales?
- The origin of structure in the universe.
  - How did the primordial spectrum of density perturbations originate?
- The origin of matter and radiation.
  - Where does all the energy in the universe come from?
  - What is the nature of the dark matter in the universe?
  - How did the matter-antimatter asymmetry arise?
- The initial singularity.
  - Did the universe have a beginning?
  - What is the global structure of the universe beyond our observable patch?

Let me discuss one by one the different issues:

4.1.1 The Flatness Problem

The Big Bang theory assumes but cannot explain the extraordinary spatial flatness of our local patch of the universe. In the general FRW metric (2) the parameter K that characterizes spatial curvature is a free parameter. There is nothing in the theory that determines this parameter a priori. However, it is directly related, via the Friedmann equation (8), to the dynamics, and thus the matter content, of the universe,

$$K = \frac{8\pi G}{3}\rho a^2 - H^2 a^2 = \frac{8\pi G}{3}\rho a^2 \left(\frac{\Omega - 1}{\Omega}\right). \qquad (118)$$

We can therefore define a new variable,

$$x \equiv \frac{\Omega - 1}{\Omega} = \frac{\text{const.}}{\rho a^2}, \qquad (119)$$

whose time evolution is given by

$$x' = \frac{dx}{dN} = (1 + 3\omega)\, x, \qquad (120)$$

where $N = \ln(a/a_i)$ characterizes the *number of e-folds* of universe expansion ($dN = H\,dt$) and where we have used Eq. (7) for the time evolution of the total energy, ρa^3, which only depends on the barotropic ratio ω. It is clear from Eq. (120) that the phase-space diagram (x, x') presents an unstable

critical (saddle) point at $x = 0$ for $\omega > -1/3$, i.e. for the radiation ($\omega = 1/3$) and matter ($\omega = 0$) eras. A small perturbation from $x = 0$ will drive the system towards $x = \pm\infty$. Since we know the universe went through both the radiation era (because of primordial nucleosynthesis) and the matter era (because of structure formation), tiny deviations from $\Omega = 1$ would have grown since then, such that today

$$x_0 = \frac{\Omega_0 - 1}{\Omega_0} = x_{\mathrm{in}} \left(\frac{T_{\mathrm{in}}}{T_{\mathrm{eq}}}\right)^2 (1 + z_{\mathrm{eq}}) . \tag{121}$$

In order that today's value be in the range $0.1 < \Omega_0 < 1.2$, or $x_0 \approx \mathcal{O}(1)$, it is required that at, say, primordial nucleosynthesis ($T_{\mathrm{NS}} \simeq 10^6 \, T_{\mathrm{eq}}$) its value be

$$\Omega(t_{\mathrm{NS}}) = 1 \pm 10^{-15} , \tag{122}$$

which represents a tremendous finetuning. Perhaps the universe indeed started with such a peculiar initial condition, but it is epistemologically more satisfying if we give a fundamental dynamical reason for the universe to have started so close to spatial flatness. These arguments were first used by Robert Dicke in the 1960s, much before inflation. He argued that the most natural initial condition for the spatial curvature should have been the Planck scale curvature, $^{(3)}R = 6K/l_{\mathrm{P}}^2$, where the Planck length is $l_{\mathrm{P}} = (\hbar G/c^3)^{1/2} = 1.62 \times 10^{-33}$ cm, that is, 60 orders of magnitude smaller than the present size of the universe, $a_0 = 1.38 \times 10^{28}$ cm. A universe with this immense curvature would have collapsed within a Planck time, $t_{\mathrm{P}} = (\hbar G/c^5)^{1/2} = 5.39 \times 10^{-44}$ s, again 60 orders of magnitude smaller than the present age of the universe, $t_0 = 4.1 \times 10^{17}$ s. Therefore, the flatness problem is also related to the Age Problem, why is it that the universe is so old and flat when, under ordinary circumstances (based on the fundamental scale of gravity) it should have lasted only a Planck time and reached a size of order the Planck length? As we will see, inflation gives a dynamical reason to such a peculiar initial condition.

4.1.2   The Homogeneity Problem

An expanding universe has *particle horizons*, that is, spatial regions beyond which causal communication cannot occur. The horizon distance can be defined as the maximum distance that light could have travelled since the origin of the universe [16],

$$d_{\mathrm{H}}(t) \equiv a(t) \int_0^t \frac{dt'}{a(t')} \sim H^{-1}(t) , \tag{123}$$

which is proportional to the Hubble scale.[9] For instance, at the beginning of nucleosynthesis the horizon distance is a few light-seconds, but grows *linearly* with time and by the end of nucleosynthesis it is a few light-minutes, i.e. a factor 100 larger, while the scale factor has increased *only* a factor of 10. The fact that the causal horizon increases faster, $d_{\mathrm{H}} \sim t$, than the scale factor, $a \sim t^{1/2}$, implies that at any given time the universe contains regions within itself that, according to the Big Bang theory, were *never* in causal contact before. For instance, the number of causally disconnected regions at a given redshift z present in our causal volume today, $d_{\mathrm{H}}(t_0) \equiv a_0$, is

$$N_{\mathrm{CD}}(z) \sim \left(\frac{a(t)}{d_{\mathrm{H}}(t)}\right)^3 \simeq (1 + z)^{3/2} , \tag{124}$$

which, for the time of decoupling, is of order $N_{\mathrm{CD}}(z_{\mathrm{dec}}) \sim 10^5 \gg 1$.

This phenomenon is particularly acute in the case of the observed microwave background. Information cannot travel faster than the speed of light, so the causal region at the time of photon decoupling could not be larger than $d_{\mathrm{H}}(t_{\mathrm{dec}}) \sim 3 \times 10^5$ light years across, or about $1°$ projected in the sky today. So why should regions that are separated by more than $1°$ in the sky today have exactly the same temperature, to within 10 ppm, when the photons that come from those two distant regions could not have been in causal contact when they were emitted? This constitutes the so-called horizon problem, see Fig. 24, and was first discussed by Robert Dicke in the 1970s as a profound inconsistency of the Big Bang theory.

---

[9]For the radiation era, the horizon distance is equal to the Hubble scale. For the matter era it is twice the Hubble scale.
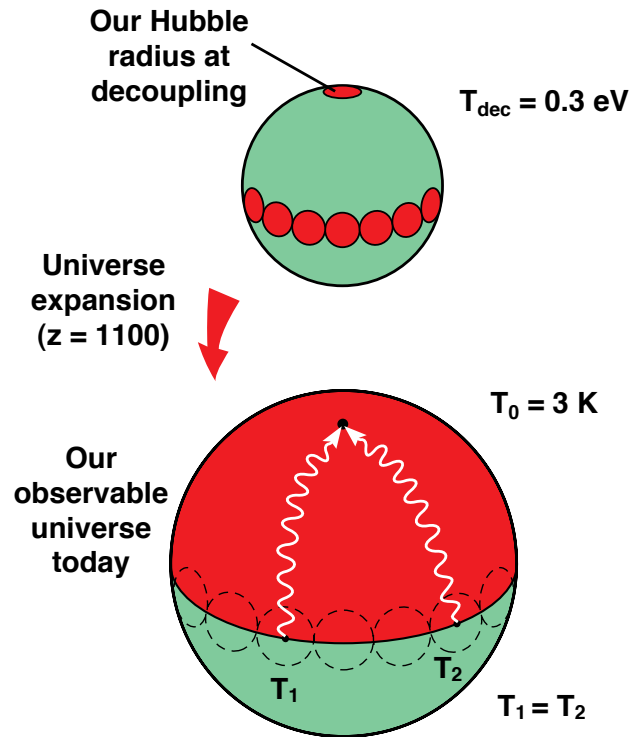
Fig. 24: Perhaps the most acute problem of the Big Bang theory is explaining the extraordinary homogeneity and isotropy of the microwave background, see Fig. 9. At the time of decoupling, the volume that gave rise to our present universe contained many causally disconnected regions (top figure). Today we observe a blackbody spectrum of photons coming from those regions and they appear to have the same temperature, $T_1 = T_2$, to one part in 10^5. Why is the universe so homogeneous? This constitutes the so-called horizon problem, which is spectacularly solved by inflation. From Ref. [74].

4.2 Cosmological Inflation

In the 1980s, a new paradigm, deeply rooted in fundamental physics, was put forward by Alan H. Guth [75], Andrei D. Linde [76] and others [77–79], to address these fundamental questions. According to the inflationary paradigm, the early universe went through a period of exponential expansion, driven by the approximately constant energy density of a scalar field called the inflaton. In modern physics, elementary particles are represented by quantum fields, which resemble the familiar electric, magnetic and gravitational fields. A field is simply a function of space and time whose quantum oscillations are interpreted as particles. In our case, the inflaton field has, associated with it, a large potential energy density, which drives the exponential expansion during inflation, see Fig. 25. We know from general relativity that the density of matter determines the expansion of the universe, but a constant energy density acts in a very peculiar way: as a repulsive force that makes any two points in space separate at exponentially large speeds. (This does not violate the laws of causality because there is no information carried along in the expansion, it is simply the stretching of space-time.)

This superluminal expansion is capable of explaining the large scale homogeneity of our observable universe and, in particular, why the microwave background looks so isotropic: regions separated today by more than $1°$ in the sky were, in fact, in causal contact before inflation, but were stretched to
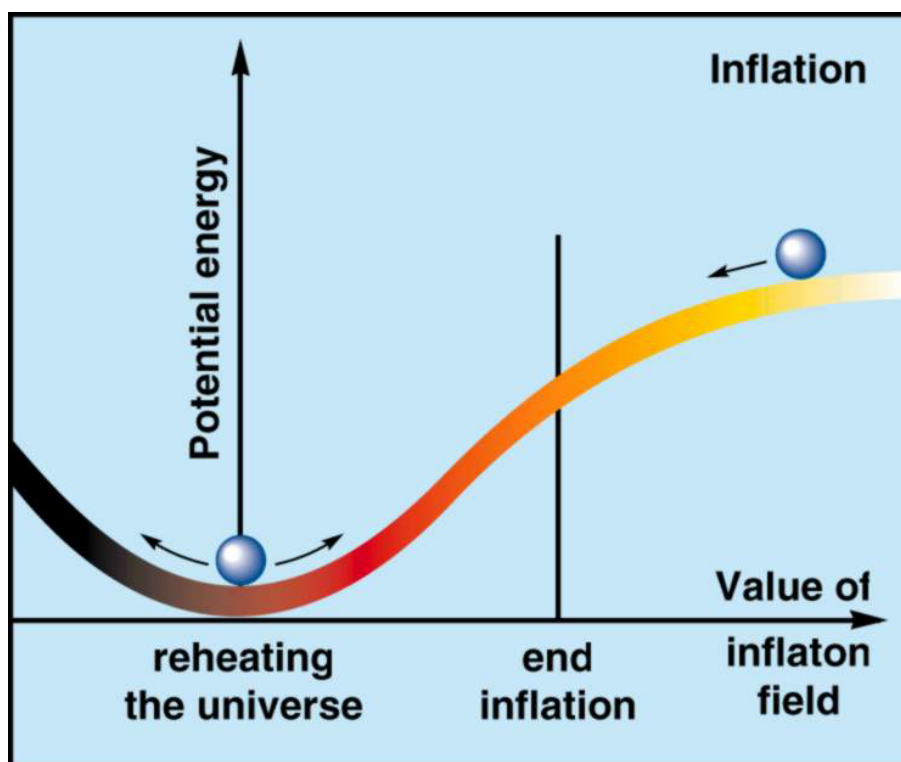
Fig. 25: The inflaton field can be represented as a ball rolling down a hill. During inflation, the energy density is approximately constant, driving the tremendous expansion of the universe. When the ball starts to oscillate around the bottom of the hill, inflation ends and the inflaton energy decays into particles. In certain cases, the coherent oscillations of the inflaton could generate a resonant production of particles which soon thermalize, reheating the universe. From Ref. [74].

cosmological distances by the expansion. Any inhomogeneities present before the tremendous expansion would be washed out. This explains why photons from supposedly causally disconneted regions have actually the same spectral distribution with the same temperature, see Fig. 24.

Moreover, in the usual Big Bang scenario a flat universe, one in which the gravitational attraction of matter is exactly balanced by the cosmic expansion, is unstable under perturbations: a small deviation from flatness is amplified and soon produces either an empty universe or a collapsed one. As we discussed above, for the universe to be nearly flat today, it must have been extremely flat at nucleosynthesis, deviations not exceeding more than one part in 10^{15}. This extreme fine tuning of initial conditions was also solved by the inflationary paradigm, see Fig. 26. Thus inflation is an extremely elegant hypothesis that explains how a region much, much greater that our own observable universe could have become smooth and flat without recourse to *ad hoc* initial conditions. Furthermore, inflation dilutes away any "unwanted" relic species that could have remained from early universe phase transitions, like monopoles, cosmic strings, etc., which are predicted in grand unified theories and whose energy density could be so large that the universe would have become unstable, and collapsed, long ago. These relics are diluted by the superluminal expansion, which leaves at most one of these particles per causal horizon, making them harmless to the subsequent evolution of the universe.

The only thing we know about this peculiar scalar field, the *inflaton*, is that it has a mass and a self-interaction potential $V(\phi)$ but we ignore everything else, even the scale at which its dynamics
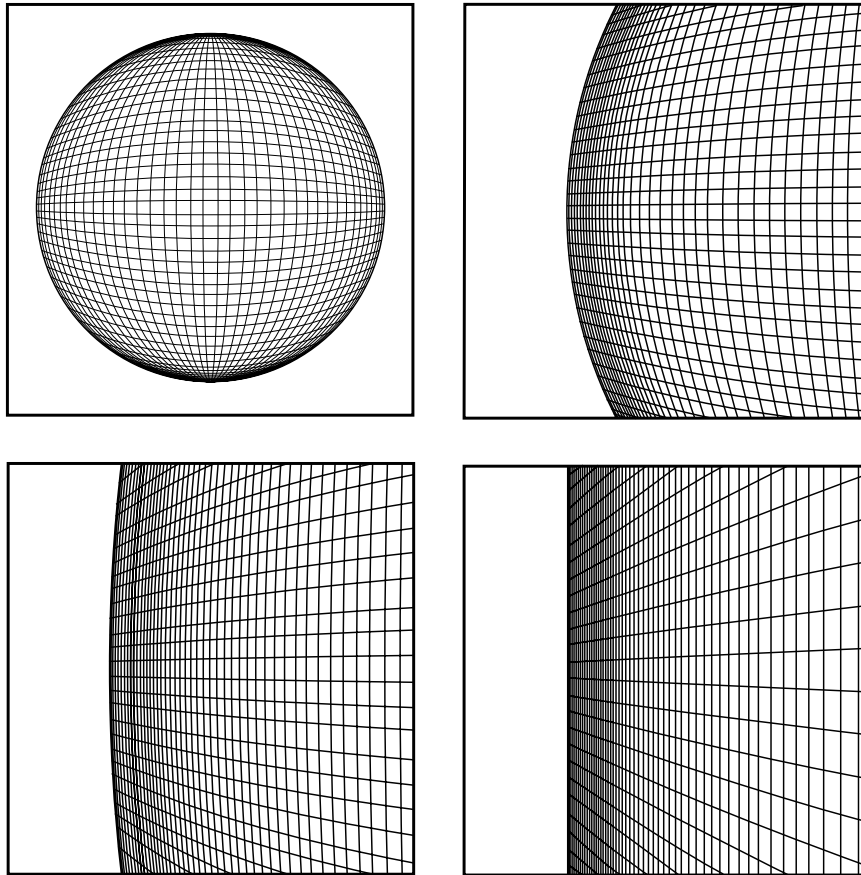
Fig. 26: The exponential expansion during inflation made the radius of curvature of the universe so large that our observable patch of the universe today appears essentialy flat, analogous (in three dimensions) to how the surface of a balloon appears flatter and flatter as we inflate it to enormous sizes. This is a crucial prediction of cosmological inflation that will be tested to extraordinary accuracy in the next few years. From Ref. [74, 78].

determines the superluminal expansion. In particular, we still do not know the nature of the inflaton field itself, is it some new *fundamental* scalar field in the electroweak symmetry breaking sector, or is it just some *effective* description of a more fundamental high energy interaction? Hopefully, in the near future, experiments in particle physics might give us a clue to its nature. Inflation had its original inspiration in the Higgs field, the scalar field supposed to be responsible for the masses of elementary particles (quarks and leptons) and the breaking of the electroweak symmetry. Such a field has not been found yet, and its discovery at the future particle colliders would help understand one of the truly fundamental problems in physics, the origin of masses. If the experiments discover something completely new and unexpected, it would automatically affect the idea of inflation at a fundamental level.

4.2.1 *Homogeneous scalar field dynamics*

In this subsection I will describe the theoretical basis for the phenomenon of inflation. Consider a scalar field ϕ, a singlet under any given interaction, with an effective potential $V(\phi)$. The Lagrangian for such a field in a curved background is

$$\mathcal{L}_{\text{inf}} = \frac{1}{2}\, g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\,, \tag{125}$$

whose evolution equation in a Friedmann-Robertson-Walker metric (2) and for a *homogeneous* field $\phi(t)$ is given by

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0\,, \tag{126}$$

where H is the rate of expansion, together with the Einstein equations,

$$H^2 = \frac{\kappa^2}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right), \tag{127}$$

$$\dot{H} = -\frac{\kappa^2}{2}\dot{\phi}^2\,, \tag{128}$$

where $\kappa^2 \equiv 8\pi G$. The dynamics of inflation can be described as a perfect fluid (5) with a time dependent pressure and energy density given by

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi)\,, \tag{129}$$

$$p = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,. \tag{130}$$

The field evolution equation (126) can then be written as the energy conservation equation,

$$\dot{\rho} + 3H(\rho + p) = 0\,. \tag{131}$$

If the potential energy density of the scalar field dominates the kinetic energy, $V(\phi) \gg \dot{\phi}^2$, then we see that

$$p \simeq -\rho \quad \Rightarrow \quad \rho \simeq \text{const.} \quad \Rightarrow \quad H(\phi) \simeq \text{const.}\,, \tag{132}$$

which leads to the solution

$$a(t) \sim \exp(Ht) \quad \Rightarrow \quad \frac{\ddot{a}}{a} > 0 \quad \text{accelerated expansion}\,. \tag{133}$$

Using the definition of the number of e-folds, $N = \ln(a/a_i)$, we see that the scale factor grows exponentially, $a(N) = a_i \exp(N)$. This solution of the Einstein equations solves immediately the flatness problem. Recall that the problem with the radiation and matter eras is that $\Omega = 1$ ($x = 0$) is an unstable critical point in phase-space. However, during inflation, with $p \simeq -\rho \Rightarrow \omega \simeq -1$, we have that $1 + 3\omega \geq 0$ and therefore $x = 0$ is a stable *attractor* of the equations of motion, see Eq. (120). As a consequence, what seemed an *ad hoc* initial condition, becomes a natural *prediction* of inflation. Suppose that during inflation the scale factor increased N e-folds, then

$$x_0 = x_{\text{in}}\, e^{-2N} \left(\frac{T_{\text{rh}}}{T_{\text{eq}}}\right)^2 (1 + z_{\text{eq}}) \simeq e^{-2N}\, 10^{56} \leq 1 \quad \Rightarrow \quad N \geq 65\,, \tag{134}$$

where we have assumed that inflation ended at the scale V_{end}, and the transfer of the inflaton energy density to thermal radiation at reheating occurred almost instantaneously[10] at the temperature $T_{\text{rh}} \sim V_{\text{end}}^{1/4} \sim 10^{15}$ GeV. Note that we can now have initial conditions with a large uncertainty, $x_{\text{in}} \simeq 1$, and still have today $x_0 \simeq 1$, thanks to the inflationary attractor towards $\Omega = 1$. This can be understood very easily by realizing that the three curvature evolves during inflation as

$$^{(3)}R = \frac{6K}{a^2} = \,^{(3)}R_{\text{in}}\, e^{-2N} \quad \longrightarrow \quad 0\,, \quad \text{for } N \gg 1\,. \tag{135}$$

Therefore, if cosmological inflation lasted over 65 e-folds, as most models predict, then today the universe (or at least our local patch) should be exactly flat, see Fig. 26, a prediction that can be tested with

---

[10]There could be a small delay in thermalization, due to the intrinsic inefficiency of reheating, but this does not change significantly the required number of e-folds.

great accuracy in the near future and for which already seems to be some evidence from observations of the microwave background [91].

Furthermore, inflation also solves the homogeneity problem in a spectacular way. First of all, due to the superluminal expansion, any inhomogeneity existing prior to inflation will be washed out,

$$\delta_k \sim \left(\frac{k}{aH}\right)^2 \Phi_k \propto e^{-2N} \longrightarrow 0, \quad \text{for } N \gg 1. \tag{136}$$

Moreover, since the scale factor grows exponentially, while the horizon distance remains essentially constant, $d_H(t) \simeq H^{-1} = \text{const.}$, any scale within the horizon during inflation will be stretched by the superluminal expansion to enormous distances, in such a way that at photon decoupling all the causally disconnected regions that encompass our present horizon actually come from a single region during inflation, about 65 e-folds before the end. This is the reason why two points separated more than $1°$ in the sky have the same backbody temperature, as observed by the COBE satellite: they were actually in causal contact during inflation. There is at present no other proposal known that could solve the homogeneity problem without invoquing an acausal mechanism like inflation.

Finally, any relic particle species (relativistic or not) existing prior to inflation will be diluted by the expansion,

$$\rho_{\mathrm{M}} \propto a^{-3} \sim e^{-3N} \longrightarrow 0, \quad \text{for } N \gg 1, \tag{137}$$

$$\rho_{\mathrm{R}} \propto a^{-4} \sim e^{-4N} \longrightarrow 0, \quad \text{for } N \gg 1. \tag{138}$$

Note that the vacuum energy density ρ_v remains constant under the expansion, and therefore, very soon it is the only energy density remaining to drive the expansion of the universe.

4.2.2 *The slow-roll approximation*

In order to simplify the evolution equations during inflation, we will consider the slow-roll approximation (SRA). Suppose that, during inflation, the scalar field evolves very slowly down its effective potential, then we can define the slow-roll parameters [80],

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{\kappa^2}{2}\frac{\dot{\phi}^2}{H^2} \ll 1, \tag{139}$$

$$\delta \equiv -\frac{\ddot{\phi}}{H\dot{\phi}} \ll 1, \tag{140}$$

$$\xi \equiv \frac{\dddot{\phi}}{H^2\dot{\phi}} - \delta^2 \ll 1. \tag{141}$$

It is easy to see that the condition

$$\epsilon < 1 \quad \Longleftrightarrow \quad \frac{\ddot{a}}{a} > 0 \tag{142}$$

characterizes inflation: it is all you need for superluminal expansion, i.e. for the horizon distance to grow more slowly than the scale factor, in order to solve the homogeneity problem, as well as for the spatial curvature to decay faster than usual, in order to solve the flatness problem.

The number of e-folds during inflation can be written with the help of Eq. (139) as

$$N = \ln\frac{a_{\mathrm{end}}}{a_i} = \int_{t_i}^{t_e} H dt = \int_{\phi_i}^{\phi_e} \frac{\kappa d\phi}{\sqrt{2\epsilon(\phi)}}, \tag{143}$$

which is an exact expression in terms of $\epsilon(\phi)$.

In the limit given by Eqs. (139), the evolution equations (126) and (127) become

$$H^2\left(1-\frac{\epsilon}{3}\right) \;\simeq\; H^2 \;=\; \frac{\kappa^2}{3}\,V(\phi)\,, \tag{144}$$

$$3H\dot\phi\left(1-\frac{\delta}{3}\right) \;\simeq\; 3H\dot\phi \;=\; -\,V'(\phi)\,. \tag{145}$$

Note that this corresponds to a reduction of the dimensionality of phase-space from two to one dimensions, $H(\phi,\dot\phi) \to H(\phi)$. In fact, it is possible to prove a theorem, for single-field inflation, which states that the slow-roll approximation is an attractor of the equations of motion, and thus we can always evaluate the inflationary trajectory in phase-space within the SRA, therefore reducing the number of initial conditions to just one, the initial value of the scalar field. If $H(\phi)$ only depends on ϕ, then $H'(\phi) = -\kappa^2\dot\phi/2$ and we can rewrite the slow-roll parameters (139) as

$$\epsilon \;=\; \frac{2}{\kappa^2}\left(\frac{H'(\phi)}{H(\phi)}\right)^2 \;\simeq\; \frac{1}{2\kappa^2}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \;\equiv\; \epsilon_V \;\ll\; 1\,, \tag{146}$$

$$\delta \;=\; \frac{2}{\kappa^2}\frac{H''(\phi)}{H(\phi)} \;\simeq\; \frac{1}{\kappa^2}\frac{V''(\phi)}{V(\phi)} - \frac{1}{2\kappa^2}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \;\equiv\; \eta_V - \epsilon_V \;\ll\; 1\,, \tag{147}$$

$$\xi \;=\; \frac{4}{\kappa^4}\frac{H'(\phi)H'''(\phi)}{H^2(\phi)} \;\simeq\; \frac{1}{\kappa^4}\frac{V'(\phi)V'''(\phi)}{V^2(\phi)} - \frac{3}{2\kappa^4}\frac{V''(\phi)}{V(\phi)}\left(\frac{V'(\phi)}{V(\phi)}\right)^2$$

$$+\; \frac{3}{4\kappa^4}\left(\frac{V'(\phi)}{V(\phi)}\right)^4 \;\equiv\; \xi_V - 3\eta_V\epsilon_V + 3\epsilon_V^2 \;\ll\; 1\,. \tag{148}$$

These expressions define the new slow-roll parameters ϵ_V, η_V and ξ_V. The number of e-folds can also be rewritten in this approximation as

$$N \;\simeq\; \int_{\phi_i}^{\phi_e}\frac{\kappa\,d\phi}{\sqrt{2\epsilon_V(\phi)}} \;=\; \kappa^2\int_{\phi_i}^{\phi_e}\frac{V(\phi)\,d\phi}{V'(\phi)}\,, \tag{149}$$

a very useful expression for evaluating N for a given effective scalar potential $V(\phi)$.

4.3 The origin of density perturbations

If cosmological inflation made the universe so extremely flat and homogeneous, where did the galaxies and clusters of galaxies come from? One of the most astonishing predictions of inflation, one that was not even expected, is that quantum fluctuations of the inflaton field are stretched by the exponential expansion and generate large-scale perturbations in the metric. Inflaton fluctuations are small wave packets of energy that, according to general relativity, modify the space-time fabric, creating a whole spectrum of curvature perturbations. The use of the word spectrum here is closely related to the case of light waves propagating in a medium: a spectrum characterizes the amplitude of each given wavelength. In the case of inflation, the inflaton fluctuations induce waves in the space-time metric that can be decomposed into different wavelengths, all with approximately the same amplitude, that is, corresponding to a scale-invariant spectrum. These patterns of perturbations in the metric are like fingerprints that unequivocally characterize a period of inflation. When matter fell in the troughs of these waves, it created density perturbations that collapsed gravitationally to form galaxies, clusters and superclusters of galaxies, with a spectrum that is also scale invariant. Such a type of spectrum was proposed in the early 1970s (before inflation) by Harrison and Zel'dovich [27], to explain the distribution of galaxies and clusters of galaxies on very large scales in our observable universe. Perhaps the most interesting aspect of structure formation is the possibility that the detailed knowledge of what seeded galaxies and clusters of galaxies will allow us to test the idea of inflation.

4.3.1 *Reparametrization invariant perturbation theory*

Until now we have considered only the unperturbed FRW metric described by a scale factor $a(t)$ and a homogeneous scalar field $\phi(t)$,

$$ds^2 = a^2(\eta)[-d\eta^2 + \gamma_{ij}\,dx^i dx^j]\,, \tag{150}$$

$$\phi = \phi(\eta)\,, \tag{151}$$

where $\eta = \int dt/a(t)$ is the conformal time, under which the background equations of motion can be written as

$$\mathcal{H}^2 = \frac{\kappa^2}{3}\left(\frac{1}{2}\phi'^2 + a^2 V(\phi)\right)\,, \tag{152}$$

$$\mathcal{H}' - \mathcal{H}^2 = -\frac{\kappa^2}{2}\phi'^2\,, \tag{153}$$

$$\phi'' + 2\mathcal{H}\phi' + a^2 V'(\phi) = 0\,, \tag{154}$$

where $\mathcal{H} = aH$ and $\phi' = a\dot{\phi}$.

During inflation, the quantum fluctuations of the scalar field will induce metric perturbations which will backreact on the scalar field. Let us consider, in linear perturbation theory, the most general line element with both scalar and tensor metric perturbations [81],[11] together with the scalar field perturbations

$$ds^2 = a^2(\eta)\left[-(1+2A)d\eta^2 + 2B_{|i}dx^i d\eta + \left\{(1+2\mathcal{R})\gamma_{ij} + 2E_{|ij} + 2h_{ij}\right\}dx^i dx^j\right]\,, \tag{155}$$

$$\phi = \phi(\eta) + \delta\phi(\eta, x^i)\,. \tag{156}$$

The indices $\{i,j\}$ label the three-dimensional spatial coordinates with metric γ_{ij}, and the $|i$ denotes covariant derivative with respect to that metric. The gauge invariant tensor perturbation h_{ij} corresponds to a transverse traceless gravitational wave, $\nabla^i h_{ij} = h^i_i = 0$. The four scalar perturbations $(A, B, \mathcal{R}, E)$ are *gauge dependent* functions of (η, x^i). Under a general coordinate (gauge) transformation [81, 82]

$$\tilde{\eta} = \eta + \xi^0(\eta, x^i)\,, \tag{157}$$

$$\tilde{x}^i = x^i + \gamma^{ij}\xi_{|j}(\eta, x^i)\,, \tag{158}$$

with arbitrary functions (ξ^0, ξ), the scalar and tensor perturbations transform, to linear order, as

$$\tilde{A} = A - \xi^{0\prime} - \mathcal{H}\xi^0\,, \qquad \tilde{B} = B + \xi^0 - \xi'\,, \tag{159}$$

$$\tilde{\mathcal{R}} = \mathcal{R} - \mathcal{H}\xi^0\,, \qquad\qquad \tilde{E} = E - \xi\,, \tag{160}$$

$$\tilde{h}_{ij} = h_{ij}\,, \tag{161}$$

where a prime denotes derivative with respect to conformal time. It is possible to construct, however, two gauge-invariant gravitational potentials [81, 82],

$$\Phi = A + (B - E')' + \mathcal{H}(B - E')\,, \tag{162}$$

$$\Psi = \mathcal{R} + \mathcal{H}(B - E')\,, \tag{163}$$

which are related through the perturbed Einstein equations,

$$\Phi = \Psi\,, \tag{164}$$

$$\frac{k^2 - 3K}{a^2}\Psi = \frac{\kappa^2}{2}\delta\rho\,, \tag{165}$$

---

[11]Note that inflation cannot generate, to linear order, a vector perturbation.

where $\delta\rho$ is the gauge-invariant density perturbation, and the latter expression is nothing but the Poisson equation for the gravitational potential, written in relativistic form.

During inflation, the energy density is given in terms of a scalar field, and thus the gauge-invariant equations for the perturbations on comoving hypersurfaces (constant energy density hypersurfaces) are

$$\Phi'' + 3\mathcal{H}\Phi' + (\mathcal{H}' + 2\mathcal{H}^2)\Phi = \frac{\kappa^2}{2}[\phi'\delta\phi' - a^2 V'(\phi)\delta\phi],\tag{166}$$

$$-\nabla^2\Phi + 3\mathcal{H}\Phi' + (\mathcal{H}' + 2\mathcal{H}^2)\Phi = -\frac{\kappa^2}{2}[\phi'\delta\phi' + a^2 V'(\phi)\delta\phi],\tag{167}$$

$$\Phi' + \mathcal{H}\Phi = \frac{\kappa^2}{2}\phi'\delta\phi,\tag{168}$$

$$\delta\phi'' + 2\mathcal{H}\delta\phi' - \nabla^2\delta\phi = 4\phi'\Phi' - 2a^2 V'(\phi)\Phi - a^2 V''(\phi)\delta\phi.\tag{169}$$

This system of equations seem too difficult to solve at first sight. However, there is a gauge invariant combination of variables that allows one to find exact solutions. Let us define [82]

$$u \equiv a\delta\phi + z\Phi,\tag{170}$$

$$z \equiv a\frac{\phi'}{\mathcal{H}}.\tag{171}$$

Under this redefinition, the above equations simplify enormously to just three independent equations,

$$u'' - \nabla^2 u - \frac{z''}{z}u = 0,\tag{172}$$

$$\nabla^2\Phi = \frac{\kappa^2}{2}\frac{\mathcal{H}}{a^2}(zu' - z'u),\tag{173}$$

$$\left(\frac{a^2\Phi}{\mathcal{H}}\right)' = \frac{\kappa^2}{2}zu.\tag{174}$$

From Equation (172) we can find a solution $u(z)$, which substituted into (174) can be integrated to give $\Phi(z)$, and together with $u(z)$ allow us to obtain $\delta\phi(z)$.

4.3.2 Quantum Field Theory in curved space-time

Until now we have treated the perturbations as classical, but we should in fact consider the perturbations Φ and $\delta\phi$ as quantum fields. Note that the perturbed action for the scalar mode u can be written as

$$\delta S = \frac{1}{2}\int d^3x\, d\eta\left[(u')^2 - (\nabla u)^2 + \frac{z''}{z}u^2\right].\tag{175}$$

In order to quantize the field u in the curved background defined by the metric (150), we can write the operator

$$\hat{u}(\eta,\mathbf{x}) = \int\frac{d^3\mathbf{k}}{(2\pi)^{3/2}}\left[u_k(\eta)\,\hat{a}_\mathbf{k}\,e^{i\mathbf{k}\cdot\mathbf{x}} + u_k^*(\eta)\,\hat{a}_\mathbf{k}^\dagger\,e^{-i\mathbf{k}\cdot\mathbf{x}}\right],\tag{176}$$

where the creation and annihilation operators satisfy the commutation relation of bosonic fields, and the scalar field's Fock space is defined through the vacuum condition,

$$[\hat{a}_\mathbf{k},\hat{a}_{\mathbf{k}'}^\dagger] = \delta^3(\mathbf{k} - \mathbf{k}'),\tag{177}$$

$$\hat{a}_\mathbf{k}|0\rangle = 0.\tag{178}$$

Note that we are not assuming that the inflaton is a fundamental scalar field, but that is can be written as a quantum field with its commutation relations (as much as a pion can be described as a quantum field).

The equations of motion for each mode $u_k(\eta)$ are decoupled in linear perturbation theory,

$$u_k'' + \left(k^2 - \frac{z''}{z} \right) u_k = 0 \,. \tag{179}$$

The ratio z''/z acts like a time-dependent potential for this Schrödinger like equation. In order to find exact solutions to the mode equation, we will use the slow-roll parameters (139), see Ref. [80]

$$\epsilon = 1 - \frac{\mathcal{H}'}{\mathcal{H}^2} = \frac{\kappa^2}{2} \frac{z^2}{a^2} \,, \tag{180}$$

$$\delta = 1 - \frac{\phi''}{\mathcal{H}\phi'} = 1 + \epsilon - \frac{z'}{\mathcal{H}z} \,, \tag{181}$$

$$\xi = - \left(2 - \epsilon - 3\delta + \delta^2 - \frac{\phi'''}{\mathcal{H}^2\phi'} \right) \,. \tag{182}$$

In terms of these parameters, the conformal time and the effective potential for the u_k mode can be written as

$$\eta = \frac{-1}{\mathcal{H}} + \int \frac{\epsilon da}{a\mathcal{H}} \,, \tag{183}$$

$$\frac{z''}{z} = \mathcal{H}^2 \left[(1 + \epsilon - \delta)(2 - \delta) + \mathcal{H}^{-1}(\epsilon' - \delta') \right] \,. \tag{184}$$

Note that the slow-roll parameters, (180) and (181), can be taken as *constant*,[12] to order ϵ^2,

$$\begin{aligned} \epsilon' &= 2\mathcal{H} \left(\epsilon^2 - \epsilon\delta \right) = \mathcal{O}(\epsilon^2) \,, \\ \delta' &= \mathcal{H} \left(\epsilon\delta - \xi \right) = \mathcal{O}(\epsilon^2) \,. \end{aligned} \tag{185}$$

In that case, for constant slow-roll parameters, we can write

$$\eta = \frac{-1}{\mathcal{H}} \frac{1}{1 - \epsilon} \,, \tag{186}$$

$$\frac{z''}{z} = \frac{1}{\eta^2} \left(\nu^2 - \frac{1}{4} \right) \,, \qquad \text{where} \qquad \nu = \frac{1 + \epsilon - \delta}{1 - \epsilon} + \frac{1}{2} \,. \tag{187}$$

We are now going to search for approximate solutions of the mode equation (179), where the effective potential (184) is of order $z''/z \simeq 2\mathcal{H}^2$ in the slow-roll approximation. In quasi-de Sitter there is a characteristic scale given by the (event) horizon size or Hubble scale during inflation, H^{-1}. There will be modes u_k with physical wavelengths much smaller than this scale, $k/a \gg H$, that are well within the de Sitter horizon and therefore do not feel the curvature of space-time. On the other hand, there will be modes with physical wavelengths much greater than the Hubble scale, $k/a \ll H$. In these two asymptotic regimes, the solutions can be written as

$$u_k = \frac{1}{\sqrt{2k}} e^{-ik\eta} \qquad k \gg aH \,, \tag{188}$$

$$u_k = C_1 z \qquad k \ll aH \,. \tag{189}$$

In the limit $k \gg aH$ the modes behave like ordinary quantum modes in Minkowsky space-time, appropriately normalized, while in the opposite limit, u/z becomes constant on superhorizon scales. For

---

[12]For instance, there are models of inflation, like power-law inflation, $a(t) \sim t^p$, where $\epsilon = \delta = 1/p < 1$, that give constant slow-roll parameters.

approximately constant slow-roll parameters one can find exact solutions to (179), with the effective potential given by (187), that interpolate between the two asymptotic solutions,

$$u_k(\eta) = \frac{\sqrt{\pi}}{2} e^{i(\nu+\frac{1}{2})\frac{\pi}{2}} (-\eta)^{1/2} H_\nu^{(1)}(-k\eta), \tag{190}$$

where $H_\nu^{(1)}(z)$ is the Hankel function of the first kind [83], and ν is given by (187) in terms of the slow-roll parameters. In the limit $k\eta \to 0$, the solution becomes

$$|u_k| = \frac{2^{\nu-\frac{3}{2}}}{\sqrt{2k}} \frac{\Gamma(\nu)}{\Gamma(\frac{3}{2})} (-k\eta)^{\frac{1}{2}-\nu} \equiv \frac{C(\nu)}{\sqrt{2k}} \left(\frac{k}{aH}\right)^{\frac{1}{2}-\nu}, \tag{191}$$

$$C(\nu) = 2^{\nu-\frac{3}{2}} \frac{\Gamma(\nu)}{\Gamma(\frac{3}{2})} (1-\epsilon)^{\nu-\frac{1}{2}} \simeq 1 \qquad \text{for} \quad \epsilon, \delta \ll 1. \tag{192}$$

We can now compute Φ and $\delta\phi$ from the super-Hubble-scale mode solution (189), for $k \ll aH$. Substituting into Eq. (174), we find

$$\Phi = C_1 \left(1 - \frac{\mathcal{H}}{a^2} \int a^2 d\eta\right) + C_2 \frac{\mathcal{H}}{a^2}, \tag{193}$$

$$\delta\phi = \frac{C_1}{a^2} \int a^2 d\eta - \frac{C_2}{a^2}. \tag{194}$$

The term proportional to C_1 corresponds to the growing solution, while that proportional to C_2 corresponds to the decaying solution, which can soon be ignored. These quantities are gauge invariant but evolve with time outside the horizon, during inflation, and before entering again the horizon during the radiation or matter eras. We would like to write an expression for a gauge invariant quantity that is also *constant* for superhorizon modes. Fortunately, in the case of adiabatic perturbations, there is such a quantity:

$$\zeta \equiv \Phi + \frac{1}{\epsilon\mathcal{H}} (\Phi' + \mathcal{H}\Phi) = \frac{u}{z}, \tag{195}$$

which is constant, see Eq. (189), for $k \ll aH$. In fact, this quantity ζ is identical, for superhorizon modes, to the gauge invariant curvature metric perturbation $\mathcal{R}_c$ on comoving (constant energy density) hypersurfaces, see Ref. [81, 84],

$$\zeta = \mathcal{R}_c + \frac{1}{\epsilon\mathcal{H}^2} \nabla^2 \Phi. \tag{196}$$

Using Eq. (173) we can write the evolution equation for $\zeta = \frac{u}{z}$ as $\zeta' = \frac{1}{\epsilon\mathcal{H}} \nabla^2\Phi$, which confirms that ζ is constant for (adiabatic[13]) superhorizon modes, $k \ll aH$. Therefore, we can evaluate the Newtonian potential Φ_k when the perturbation reenters the horizon during radiation/matter eras in terms of the curvature perturbation $\mathcal{R}_k$ when it left the Hubble scale during inflation,

$$\Phi_k = \left(1 - \frac{\mathcal{H}}{a^2} \int a^2 d\eta\right) \mathcal{R}_k = \frac{3+3\omega}{5+3\omega} \mathcal{R}_k = \begin{cases} \frac{2}{3} \mathcal{R}_k & \text{radiation era}, \\ \frac{3}{5} \mathcal{R}_k & \text{matter era}. \end{cases} \tag{197}$$

Let us now compute the tensor or gravitational wave metric perturbations generated during inflation. The perturbed action for the tensor mode can be written as

$$\delta S = \frac{1}{2} \int d^3x \, d\eta \, \frac{a^2}{2\kappa^2} \left[(h'_{ij})^2 - (\nabla h_{ij})^2\right], \tag{198}$$

---

[13]This conservation fails for entropy or isocurvature perturbations, see Ref. [84].

with the tensor field h_{ij} considered as a quantum field,

$$\hat{h}_{ij}(\eta, \mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \sum_{\lambda=1,2} \left[h_k(\eta)\, e_{ij}(\mathbf{k}, \lambda)\, \hat{a}_{\mathbf{k},\lambda}\, e^{i\mathbf{k}\cdot\mathbf{x}} + h.c. \right], \tag{199}$$

where $e_{ij}(\mathbf{k}, \lambda)$ are the two polarization tensors, satisfying symmetric, transverse and traceless conditions

$$e_{ij} = e_{ji}, \quad k^i e_{ij} = 0, \quad e_{ii} = 0, \tag{200}$$

$$e_{ij}(-\mathbf{k}, \lambda) = e_{ij}^*(\mathbf{k}, \lambda), \quad \sum_\lambda e_{ij}^*(\mathbf{k}, \lambda) e^{ij}(\mathbf{k}, \lambda) = 4, \tag{201}$$

while the creation and annihilation operators satisfy the usual commutation relation of bosonic fields, Eq. (177). We can now redefine our gauge invariant tensor amplitude as

$$v_k(\eta) = \frac{a}{\sqrt{2}\kappa}\, h_k(\eta), \tag{202}$$

which satisfies the following evolution equation, decoupled for each mode $v_k(\eta)$ in linear perturbation theory,

$$v_k'' + \left(k^2 - \frac{a''}{a}\right) v_k = 0. \tag{203}$$

The ratio a''/a acts like a time-dependent potential for this Schrödinger like equation, analogous to the term z''/z for the scalar metric perturbation. For constant slow-roll parameters, the potential becomes

$$\frac{a''}{a} = 2\mathcal{H}^2 \left(1 - \frac{\epsilon}{2}\right) = \frac{1}{\eta^2}\left(\mu^2 - \frac{1}{4}\right), \tag{204}$$

$$\mu = \frac{1}{1-\epsilon} + \frac{1}{2}. \tag{205}$$

We can solve equation (203) in the two asymptotic regimes,

$$v_k = \frac{1}{\sqrt{2k}}\, e^{-ik\eta} \qquad k \gg aH, \tag{206}$$

$$v_k = C\, a \qquad\qquad k \ll aH. \tag{207}$$

In the limit $k \gg aH$ the modes behave like ordinary quantum modes in Minkowsky space-time, appropriately normalized, while in the opposite limit, the metric perturbation h_k becomes *constant* on superhorizon scales. For constant slow-roll parameters one can find exact solutions to (203), with effective potential given by (204), that interpolate between the two asymptotic solutions. These are identical to Eq. (190) except for the substitution $\nu \to \mu$. In the limit $k\eta \to 0$, the solution becomes

$$|v_k| = \frac{C(\mu)}{\sqrt{2k}} \left(\frac{k}{aH}\right)^{\frac{1}{2} - \mu}. \tag{208}$$

Since the mode h_k becomes constant on superhorizon scales, we can evaluate the tensor metric perturbation when it reentered during the radiation or matter era directly in terms of its value during inflation.

4.3.3 *Power spectrum of scalar and tensor metric perturbations*

Not only do we expect to measure the amplitude of the metric perturbations generated during inflation and responsible for the anisotropies in the CMB and density fluctuations in LSS, but we should also be able to measure its power spectrum, or two-point correlation function in Fourier space. Let us consider

first the scalar metric perturbations $\mathcal{R}_k$, which enter the horizon at $a = k/H$. Its correlator is given by [80]

$$\langle 0|\mathcal{R}_k^* \mathcal{R}_{k'}|0\rangle = \frac{|u_k|^2}{z^2}\,\delta^3(\mathbf{k}-\mathbf{k}') \equiv \frac{\mathcal{P}_\mathcal{R}(k)}{4\pi k^3}\,(2\pi)^3\,\delta^3(\mathbf{k}-\mathbf{k}')\,, \tag{209}$$

$$\mathcal{P}_\mathcal{R}(k) = \frac{k^3}{2\pi^2}\,\frac{|u_k|^2}{z^2} = \frac{\kappa^2}{2\epsilon}\left(\frac{H}{2\pi}\right)^2\left(\frac{k}{aH}\right)^{3-2\nu} \equiv A_S^2\left(\frac{k}{aH}\right)^{n_s-1}\,, \tag{210}$$

where we have used $\mathcal{R}_k = \zeta_k = \frac{u_k}{z}$ and Eq. (191). This last equation determines the power spectrum in terms of its amplitude at horizon-crossing, A_S, and a tilt,

$$n_s - 1 \equiv \frac{d\ln\mathcal{P}_\mathcal{R}(k)}{d\ln k} = 3 - 2\nu = 2\left(\frac{\delta-2\epsilon}{1-\epsilon}\right) \simeq 2\eta_V - 6\epsilon_V\,, \tag{211}$$

see Eqs. (146), (147). Note from this equation that it is possible, in principle, to obtain from inflation a scalar tilt which is either positive ($n > 1$) or negative ($n < 1$). Furthermore, depending on the particular inflationary model [85], we can have significant departures from scale invariance.

Note that at horizon entry $k\eta = -1$, and thus we can alternatively evaluate the tilt as

$$n_s - 1 \equiv -\frac{d\ln\mathcal{P}_\mathcal{R}}{d\ln\eta} = -2\eta\mathcal{H}\left[(1-\epsilon)-(\epsilon-\delta)-1\right] = 2\left(\frac{\delta-2\epsilon}{1-\epsilon}\right) \simeq 2\eta_V - 6\epsilon_V\,, \tag{212}$$

and the running of the tilt

$$\frac{dn_s}{d\ln k} = -\frac{dn_s}{d\ln\eta} = -\eta\mathcal{H}\left(2\xi + 8\epsilon^2 - 10\epsilon\delta\right) \simeq 2\xi_V + 24\epsilon_V^2 - 16\eta_V\epsilon_V\,, \tag{213}$$

where we have used Eqs. (185).

Let us consider now the tensor (gravitational wave) metric perturbation, which enter the horizon at $a = k/H$,

$$\sum_\lambda \langle 0|h_{k,\lambda}^* h_{k',\lambda}|0\rangle = 4\,\frac{2\kappa^2}{a^2}\,|v_k|^2\delta^3(\mathbf{k}-\mathbf{k}') \equiv \frac{\mathcal{P}_g(k)}{4\pi k^3}\,(2\pi)^3\,\delta^3(\mathbf{k}-\mathbf{k}')\,, \tag{214}$$

$$\mathcal{P}_g(k) = 8\kappa^2\left(\frac{H}{2\pi}\right)^2\left(\frac{k}{aH}\right)^{3-2\mu} \equiv A_T^2\left(\frac{k}{aH}\right)^{n_T}\,, \tag{215}$$

where we have used Eqs. (202) and (208). Therefore, the power spectrum can be approximated by a power-law expression, with amplitude A_T and tilt

$$n_T \equiv \frac{d\ln\mathcal{P}_g(k)}{d\ln k} = 3 - 2\mu = \frac{-2\epsilon}{1-\epsilon} \simeq -2\epsilon_V < 0\,, \tag{216}$$

which is always negative. In the slow-roll approximation, $\epsilon \ll 1$, the tensor power spectrum is scale invariant.

Alternatively, we can evaluate the tensor tilt by

$$n_T \equiv -\frac{d\ln\mathcal{P}_g}{d\ln\eta} = -2\eta\mathcal{H}\left[(1-\epsilon)-1\right] = \frac{-2\epsilon}{1-\epsilon} \simeq -2\epsilon_V\,, \tag{217}$$

and its running by

$$\frac{dn_T}{d\ln k} = -\frac{dn_T}{d\ln\eta} = -\eta\mathcal{H}\left(4\epsilon^2 - 4\epsilon\delta\right) \simeq 8\epsilon_V^2 - 4\eta_V\epsilon_V\,, \tag{218}$$

where we have used Eqs. (185).

4.4 The anisotropies of the microwave background

The metric fluctuations generated during inflation are not only responsible for the density perturbations that gave rise to galaxies via gravitational collapse, but one should also expect to see such ripples in the metric as temperature anisotropies in the cosmic microwave background, that is, minute deviations in the temperature of the blackbody spectrum when we look at different directions in the sky. Such anisotropies had been looked for ever since Penzias and Wilson's discovery of the CMB, but had eluded all detection, until COBE satellite discovered them in 1992, see Fig. 9. The reason why they took so long to be discovered was that they appear as perturbations in temperature of only one part in 10^5. Soon after COBE, other groups quickly confirmed the detection of temperature anisotropies at around $30 \, \mu$K, at higher multipole numbers or smaller angular scales.

4.4.1 The Sachs-Wolfe effect

The anisotropies corresponding to large angular scales are only generated via gravitational red-shift and density perturbations through the Einstein equations, $\delta\rho/\rho = -2\Phi$ for adiabatic perturbations; we can ignore the Doppler contribution, since the perturbation is non-causal. In that case, the temperature anisotropy in the sky today is given by [86]

$$\frac{\delta T}{T}(\theta, \phi) = \frac{1}{3}\Phi(\eta_{\mathrm{LS}}) \, Q(\eta_0, \theta, \phi) + 2 \int_{\eta_{\mathrm{LS}}}^{\eta_0} dr \, \Phi'(\eta_0 - r) \, Q(r, \theta, \phi) \,, \tag{219}$$

where η_0 is the *coordinate distance* to the last scattering surface, i.e. the present conformal time, while $\eta_{\mathrm{LS}} \simeq 0$ determines that comoving hypersurface. The above expression is known as the Sachs-Wolfe effect [86], and contains two parts, the intrinsic and the Integrated Sachs-Wolfe (ISW) effect, due to integration along the line of sight of time variations in the gravitational potential.

In linear perturbation theory, the scalar metric perturbations can be separated into $\Phi(\eta, \mathbf{x}) \equiv \Phi(\eta) \, Q(\mathbf{x})$, where $Q(\mathbf{x})$ are the scalar harmonics, eigenfunctions of the Laplacian in three dimensions, $\nabla^2 Q_{klm}(r, \theta, \phi) = -k^2 \, Q_{klm}(r, \theta, \phi)$. These functions have the general form [87]

$$Q_{klm}(r, \theta, \phi) = \Pi_{kl}(r) \, Y_{lm}(\theta, \phi) \,, \tag{220}$$

where $Y_{lm}(\theta, \phi)$ are the usual spherical harmonics [83].

In order to compute the temperature anisotropy associated with the Sachs-Wolfe effect, we have to know the evolution of the metric perturbation during the matter era,

$$\Phi'' + 3\mathcal{H} \, \Phi' + a^2 \Lambda \, \Phi - 2K \, \Phi = 0 \,. \tag{221}$$

In the case of a flat universe without cosmological constant, the Newtonian potential remains constant during the matter era and only the intrinsic SW effect contributes to $\delta T/T$. In case of a non-vanishing Λ, since its contribution is negligible in the past, most of the photon's trajectory towards us is unperturbed, and the only difference with respect to the $\Lambda = 0$ case is an overall factor [90]. We will consider here the approximation $\Phi = $ constant during the matter era and ignore that factor, see Ref. [88].

In a flat universe, the radial part of the eigenfunctions (220) can be written as [87]

$$\Pi_{kl}(r) = \sqrt{\frac{2}{\pi}} \, k \, j_l(kr) \,, \tag{222}$$

where $j_l(z)$ are the spherical Bessel functions [83]. The growing mode solution of the metric perturbation that left the Hubble scale during inflation contributes to the temperature anisotropies on large scales (219) as

$$\frac{\delta T}{T}(\theta, \phi) = \frac{1}{3}\Phi(\eta_{\mathrm{LS}}) \, Q = \frac{1}{5}\mathcal{R} \, Q(\eta_0, \theta, \phi) \equiv \sum_{l=2}^{\infty} \sum_{m=-l}^{l} a_{lm} \, Y_{lm}(\theta, \phi) \,, \tag{223}$$

where we have used the fact that at reentry (at the surface of last scattering) the gauge invariant Newtonian potential Φ is related to the curvature perturbation $\mathcal{R}$ at Hubble-crossing during inflation, see Eq. (197); and we have expanded $\delta T/T$ in spherical harmonics.

We can now compute the two-point correlation function or angular power spectrum, $C(\theta)$, of the CMB anisotropies on large scales, defined as an expansion in multipole number,

$$C(\theta) = \left\langle \frac{\delta T^*}{T}(\mathbf{n}) \frac{\delta T}{T}(\mathbf{n}') \right\rangle_{\mathbf{n} \cdot \mathbf{n}' = \cos\theta} = \frac{1}{4\pi} \sum_{l=2}^{\infty} (2l+1)\, C_l\, P_l(\cos\theta) \,, \tag{224}$$

where $P_l(z)$ are the Legendre polynomials [83], and we have averaged over different universe realizations. Since the coefficients a_{lm} are isotropic (to first order), we can compute the $C_l = \langle |a_{lm}|^2 \rangle$ as

$$C_l^{(S)} = \frac{4\pi}{25} \int_0^\infty \frac{dk}{k}\, \mathcal{P}_\mathcal{R}(k)\, j_l^2(k\eta_0) \,, \tag{225}$$

where we have used Eqs. (223) and (209). In the case of scalar metric perturbation produced during inflation, the scalar power spectrum at reentry is given by $\mathcal{P}_\mathcal{R}(k) = A_S^2 (k\eta_0)^{n-1}$, in the power-law approximation, see Eq. (210). In that case, one can integrate (225) to give

$$C_l^{(S)} = \frac{2\pi}{25} A_S^2 \frac{\Gamma[\frac{3}{2}]\, \Gamma[1 - \frac{n-1}{2}]\, \Gamma[l + \frac{n-1}{2}]}{\Gamma[\frac{3}{2} - \frac{n-1}{2}]\, \Gamma[l + 2 - \frac{n-1}{2}]} \,, \tag{226}$$

$$\frac{l(l+1)\, C_l^{(S)}}{2\pi} = \frac{A_S^2}{25} = \text{constant} \,, \qquad \text{for} \quad n = 1 \,. \tag{227}$$

This last expression corresponds to what is known as the Sachs-Wolfe plateau, and is the reason why the coefficients C_l are always plotted multiplied by $l(l+1)$, see Fig. 3.4.

Tensor metric perturbations also contribute with an approximately constant angular power spectrum, $l(l+1)C_l$. The Sachs-Wolfe effect for a gauge invariant tensor perturbation is given by [86]

$$\frac{\delta T}{T}(\theta, \phi) = \int_{\eta_{\mathrm{LS}}}^{\eta_0} dr\, h'(\eta_0 - r)\, Q_{rr}(r, \theta, \phi) \,, \tag{228}$$

where Q_{rr} is the rr-component of the tensor harmonic along the line of sight [87]. The tensor perturbation h during the matter era satisfies the following evolution equation

$$h_k'' + 3\mathcal{H}\, h_k' + (k^2 + 2K)\, h_k = 0 \,, \tag{229}$$

which depends on the wavenumber k, contrary to what happens with the scalar modes, see Eq. (221). For a flat ($K = 0$) universe, the solution to this equation is $h_k(\eta) = h\, G_k(\eta)$, where h is the constant tensor metric perturbation at horizon crossing and $G_k(\eta) = 3\, j_1(k\eta)/k\eta$, normalized so that $G_k(0) = 1$ at the surface of last scattering. The radial part of the tensor harmonic Q_{rr} in a flat universe can be written as [87]

$$Q_{kl}^{rr}(r) = \left[\frac{(l-1)l(l+1)(l+2)}{\pi k^2} \right]^{1/2} \frac{j_l(kr)}{r^2} \,. \tag{230}$$

The tensor angular power spectrum can finally be expressed as

$$C_l^{(T)} = \frac{9\pi}{4} (l-1)l(l+1)(l+2) \int_0^\infty \frac{dk}{k}\, \mathcal{P}_g(k)\, I_{kl}^2 \,, \tag{231}$$

$$I_{kl} = \int_0^{x_0} dx\, \frac{j_2(x_0 - x) j_l(x)}{(x_0 - x) x^2} \,, \tag{232}$$

where $x \equiv k\eta$, and $\mathcal{P}_g(k)$ is the primordial tensor spectrum (215). For a scale invariant spectrum, $n_T = 0$, we can integrate (231) to give [89]

$$l(l+1)\, C_l^{(T)} = \frac{\pi}{36} \left(1 + \frac{48\pi^2}{385} \right) A_T^2\, B_l \,, \tag{233}$$

with $B_l = (1.1184, 0.8789, \ldots, 1.00)$ for $l = 2, 3, \ldots, 30$. Therefore, $l(l+1)\, C_l^{(T)}$ also becomes constant for large l. Beyond $l \sim 30$, the Sachs-Wolfe expression is not a good approximation and the tensor angular power spectrum decays very quickly at large l, see Fig. 28.

4.4.2 The consistency relation

In spite of the success of inflation in predicting a homogeneous and isotropic background on which to imprint a scale-invariant spectrum of inhomogeneities, it is difficult to test the idea of inflation. A CMB cosmologist before the 1980s would have argued that *ad hoc* initial conditions could have been at the origin of the homogeneity and flatness of the universe on large scales, while a LSS cosmologist would have agreed with Harrison and Zel'dovich that the most natural spectrum needed to explain the formation of structure was a scale-invariant spectrum. The surprise was that inflation incorporated an understanding of *both* the globally homogeneous and spatially flat background, and the approximately scale-invariant spectrum of perturbations in the same formalism. But that could have been just a coincidence.

What is *unique* to inflation is the fact that inflation determines not just one but *two* primordial spectra, corresponding to the scalar (density) and tensor (gravitational waves) metric perturbations, from a single continuous function, the inflaton potential $V(\phi)$. In the slow-roll approximation, one determines, from $V(\phi)$, two continuous functions, $\mathcal{P}_{\mathcal{R}}(k)$ and $\mathcal{P}_g(k)$, that in the power-law approximation reduces to two amplitudes, A_S and A_T, and two tilts, n and n_T. It is clear that there must be a relation between the four parameters. Indeed, one can see from Eqs. (233) and (227) that the ratio of the tensor to scalar contribution to the angular power spectrum is proportional to the tensor tilt [80],

$$r \equiv \frac{A_T^2}{A_S^2} = 16\epsilon \simeq -8\, n_T \,. \tag{234}$$

This is a unique prediction of inflation, which could not have been postulated a priori by any cosmologist. If we finally observe a tensor spectrum of anisotropies in the CMB, or a stochastic gravitational wave background in laser interferometers like LIGO or LISA, with sufficient accuracy to determine their spectral tilt, one might have some chance to test the idea of inflation, via the consistency relation (234). For the moment, observations of the microwave background anisotropies suggest that the Sachs-Wolfe plateau exists, see Fig. 3.4, but it is still premature to determine the tensor contribution. Perhaps in the near future, from the analysis of polarization as well as temperature anisotropies, with the CMB satellites MAP and Planck, we might have a chance of determining the validity of the consistency relation.

Assuming that the scalar contribution dominates over the tensor on large scales, i.e. $r \ll 1$, one can actually give a measure of the amplitude of the scalar metric perturbation from the observations of the Sachs-Wolfe plateau in the angular power spectrum [21],

$$\left[\frac{l(l+1)\, C_l^{(S)}}{2\pi} \right]^{1/2} = \frac{A_S}{5} = (0.926 \pm 0.0106) \times 10^{-5} \,, \tag{235}$$

$$n = 0.9667 \pm 0.0040 \,, \tag{236}$$

$$\frac{dn}{d\ln k} = -0.002 \pm 0.013 \,. \tag{237}$$

These measurements can be used to normalize the primordial spectrum and determine the parameters of the model of inflation [85]. In the near future these parameters will be determined with much better accuracy, as described in Section 4.4.5.

4.4.3 The acoustic peaks

The Sachs-Wolfe plateau is a distinctive feature of Fig. 24. These observations confirm the existence of a primordial spectrum of scalar (density) perturbations on all scales, otherwise the power spectrum would have started from zero at $l = 2$. However, we see that the spectrum starts to rise around $l = 20$ towards the first acoustic peak, where the SW approximation breaks down and the above formulae are no longer valid.

As mentioned above, the first peak in the photon distribution corresponds to overdensities that have undergone half an oscillation, that is, a compression, and appear at a scale associated with the size of the horizon at last scattering, about $1°$ projected in the sky today. Since photons scatter off baryons, they will also feel the acoustic wave and create a peak in the correlation function. The height of the peak is proportional to the amount of baryons: the larger the baryon content of the universe, the higher the peak. The position of the peak in the power spectrum depends on the geometrical size of the particle horizon at last scattering. Since photons travel along geodesics, the projected size of the causal horizon at decoupling depends on whether the universe is flat, open or closed. In a flat universe the geodesics are straight lines and, by looking at the angular scale of the first acoustic peak, we would be measuring the actual size of the horizon at last scattering. In an open universe, the geodesics are inward-curved trajectories, and therefore the projected size on the sky appears smaller. In this case, the first acoustic peak should occur at higher multipoles or smaller angular scales. On the other hand, for a closed universe, the first peak occurs at smaller multipoles or larger angular scales. The dependence of the position of the first acoustic peak on the spatial curvature can be approximately given by $l_{\rm peak} \simeq 220 \, \Omega_0^{-1/2}$, where $\Omega_0 = \Omega_{\rm M} + \Omega_\Lambda = 1 - \Omega_K$. Past observations from the balloon experiment BOOMERANG [91], suggested clearly a few years ago that the first peak was between $l = 180$ and 250 at 95% c.l., with an amplitude $\delta T = 80 \pm 10 \, \mu$K, and therefore the universe was most probably flat. However, with the high precision Planck data we can now pinpoint the spatial curvature to less than a tenth of a percent,

$$\Omega_0 = 0.9992 \pm 0.0040 \quad (95\% \text{ c.l.}) \tag{238}$$

Therefore, the universe is spatially flat (i.e. Euclidean), within 0.1% uncertainty, which is much better than we could ever do before, and is one the most robust predictions of inflation.

With WMAP and specially with Planck, we have now evidence of at least nine distinct acoustic peaks. These peaks should occur at harmonics of the first one, but are typically much lower because of Silk damping. Since the amplitude and position of the primary and secondary peaks are directly determined by the sound speed (and, hence, the equation of state) and by the geometry and expansion of the universe, they can be used as a powerful test of the density of baryons and dark matter, and other cosmological parameters. By looking at these patterns in the anisotropies of the microwave background, cosmologists can determine not only the cosmological parameters, but also the primordial spectrum of density perturbations produced during inflation. It turns out that the observed temperature anisotropies are compatible with a scale-invariant spectrum, see Eq. (236), as predicted by inflation. This is remarkable, and gives very strong support to the idea that inflation may indeed be responsible for both the CMB anisotropies and the large-scale structure of the universe. Different models of inflation have different specific predictions for the fine details associated with the spectrum generated during inflation. It is these minute differences that will allow cosmologists to differentiate between alternative models of inflation and discard those that do not agree with observations. However, most importantly, perhaps, the pattern of anisotropies predicted by inflation is completely different from those predicted by alternative models of structure formation, like cosmic defects: strings, vortices, textures, etc. These are complicated networks of energy density concentrations left over from an early universe phase transition, analogous to the defects formed in the laboratory in certain kinds of liquid crystals when they go through a phase transition. The cosmological defects have spectral properties very different from those generated by inflation. That is why it is so important to launch more sensitive instruments, and with better angular resolution, to determine the properties of the CMB anisotropies.

4.4.4 The new microwave anisotropy satellites, WMAP and Planck

The large amount of information encoded in the anisotropies of the microwave background is the reason why both NASA and the European Space Agency have decided to launch two independent satellites to measure the CMB temperature and polarization anisotropies to unprecedented accuracy. The Wilkinson Microwave Anisotropy Probe [92] was launched by NASA at the end of 2000, while Planck [93] was launched by ESA in 2009 and both have fulfilled all our expectations for temperature and E mode polarization. There are at the moment other large missiuon proposals like PRISM [99], and CORE+ [100], which should provide precision measurements of CMB polarization anisotropies and detect for the first time the primordial B modes of inflation.

As we have emphasized before, the fact that these anisotropies have such a small amplitude allow for an accurate calculation of the predicted anisotropies in linear perturbation theory. A particular cosmological model is characterized by a dozen or so parameters: the rate of expansion, the spatial curvature, the baryon content, the cold dark matter and neutrino contribution, the cosmological constant (vacuum energy), the reionization parameter (optical depth to the last scattering surface), and various primordial spectrum parameters like the amplitude and tilt of the adiabatic and isocurvature spectra, the amount of gravitational waves, non-Gaussian effects, etc. All these parameters can now be fed into very fast CMB codes called CMBFAST [97] and CAMB [98], that compute the predicted temperature and polarization anisotropies to better than 0.1% accuracy, and thus can be used to compare with observations.

These two satellites have improved both the sensitivity, down to μK, and the resolution, down to arc minutes, with respect to the previous COBE satellite, thanks to large numbers of microwave horns of various sizes, positioned at specific angles, and also thanks to recent advances in detector technology, with high electron mobility transistor amplifiers (HEMTs) for frequencies below 100 GHz and bolometers for higher frequencies. The primary advantage of HEMTs is their ease of use and speed, with a typical sensitivity of 0.5 mKs$^{1/2}$, while the advantage of bolometers is their tremendous sensitivity, better than 0.1 mKs$^{1/2}$, see Ref. [101]. This has allowed cosmologists to extract information from around 3000 multipoles! Since most of the cosmological parameters have specific signatures in the height and position of the first few acoustic peaks, the higher the resolution, the more peaks one is expected to see, and thus the better the accuracy with which one will be able to measure those parameters, see Table 2.

Although the satellite probes were designed for the accurate measurement of the CMB temperature anisotropies, there are other experiments, like balloon-borne and ground interferometers [94]. Probably the most important objective of the future satellites (beyond WMAP and Planck) will be the measurement of the CMB polarization anisotropies, discovered by DASI in November 2002 [102], and confirmed a few months later by WMAP [21], and by Planck [93] with much greater accuracy, see Fig. 27. These anisotropies were predicted by models of structure formation and indeed found at the level of microKelvin sensitivities, where the new satellites were aiming at. The complementary information contained in the polarization anisotropies already provides much more stringent constraints on the cosmological parameters than from the temperature anisotropies alone. However, in the future, PRISM and CORE+ will have much better sensitivities. In particular, the curl-curl component of the polarization power spectra is nowadays the only means we have to determine the tensor (gravitational wave) contribution to the metric perturbations responsible for temperature anisotropies. If such a component is found, one could finally confirm the inflationary paradigm [95].

4.5 From metric perturbations to large scale structure

If inflation is responsible for the metric perturbations that gave rise to the temperature anisotropies observed in the microwave background, then the primordial spectrum of density inhomogeneities induced by the same metric perturbations should also be responsible for the present large scale structure [104]. This simple connection allows for more stringent tests on the inflationary paradigm for the generation of metric perturbations, since it relates the large scales (of order the present horizon) with the smallest scales (on galaxy scales). This provides a very large lever arm for the determination of primordial spectra
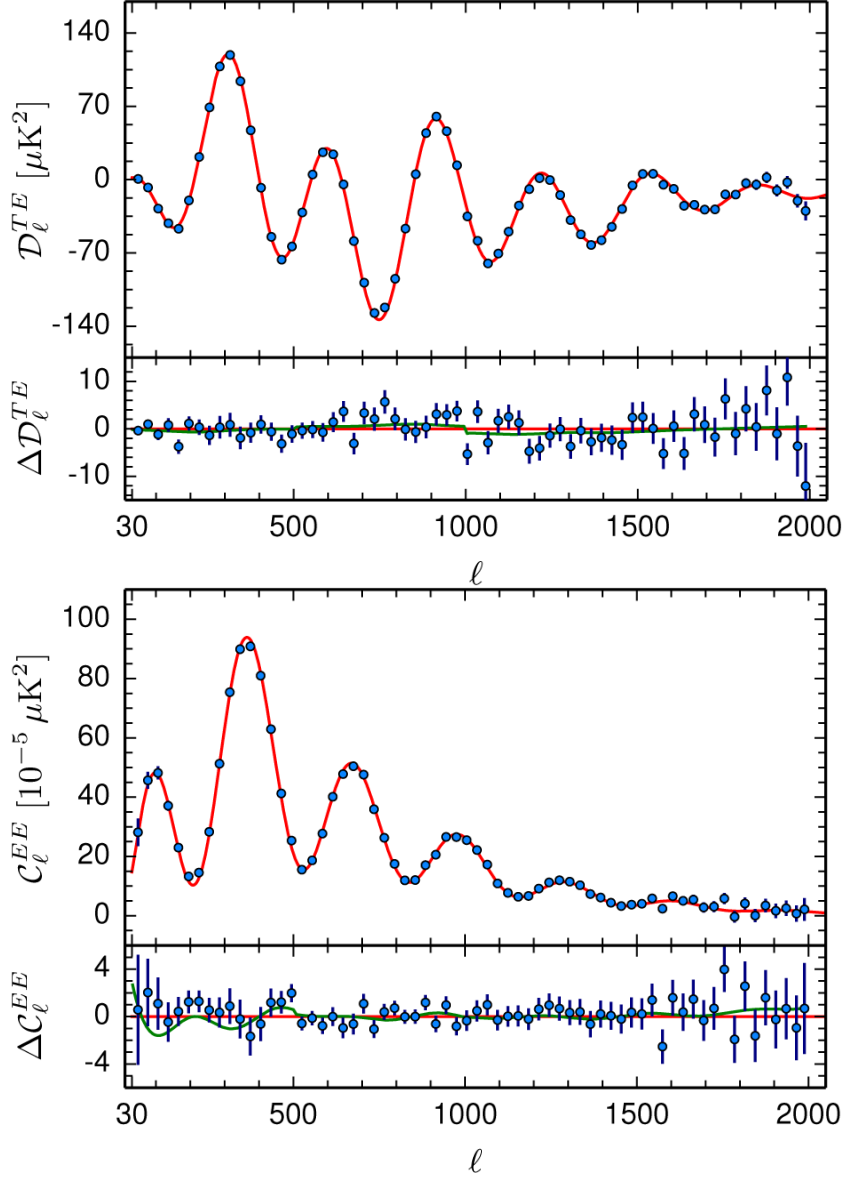
Fig. 27: Planck measurements of the TE cross-correlation and EE power spectrum of CMB polarization fluctuations, and residuals w.r.t. a Standard Cosmological Model. We only have upper bounds on the BB power spectrum. From Ref. [93].

parameters like the tilt, the nature of the perturbations, whether adiabatic or isocurvature, the geometry of the universe, as well as its matter and energy content, whether CDM, HDM or mixed CHDM.

4.5.1 The galaxy power spectrum

As metric perturbations enter the causal horizon during the radiation or matter era, they create density fluctuations via gravitational attraction of the potential wells. The density contrast δ can be deduced from the Einstein equations in linear perturbation theory, see Eq. (165),

$$\delta_k \equiv \frac{\delta\rho_k}{\rho} = \left(\frac{k}{aH}\right)^2 \frac{2}{3}\Phi_k = \left(\frac{k}{aH}\right)^2 \frac{2+2\omega}{5+3\omega}\mathcal{R}_k\,, \tag{239}$$

where we have assumed $K = 0$, and used Eq. (197). From this expression one can compute the power spectrum, at horizon crossing, of matter density perturbations induced by inflation, see Eq. (209),

$$P(k) = \langle |\delta_k|^2 \rangle = A \left(\frac{k}{aH} \right)^n , \tag{240}$$

with n given by the scalar tilt (211), $n = 1 + 2\eta - 6\epsilon$. This spectrum reduces to a Harrison-Zel'dovich spectrum (100) in the slow-roll approximation: η, $\epsilon \ll 1$.

Since perturbations evolve after entering the horizon, the power spectrum will not remain constant. For scales entering the horizon well after matter domination ($k^{-1} \gg k_{\rm eq}^{-1} \simeq 81$ Mpc), the metric perturbation has not changed significantly, so that $\mathcal{R}_k({\rm final}) = \mathcal{R}_k({\rm initial})$. Then Eq. (239) determines the final density contrast in terms of the initial one. On smaller scales, there is a linear transfer function $T(k)$, which may be defined as [80]

$$\mathcal{R}_k({\rm final}) = T(k)\, \mathcal{R}_k({\rm initial}) . \tag{241}$$

To calculate the transfer function one has to specify the initial condition with the relative abundance of photons, neutrinos, baryons and cold dark matter long before horizon crossing. The most natural condition is that the abundances of all particle species are uniform on comoving hypersurfaces (with constant total energy density). This is called the *adiabatic* condition, because entropy is conserved independently for each particle species X, i.e. $\delta\rho_X = \dot{\rho}_X \delta t$, given a perturbation in time from a comoving hypersurface, so

$$\frac{\delta\rho_X}{\rho_X + p_X} = \frac{\delta\rho_Y}{\rho_Y + p_Y} , \tag{242}$$

where we have used the energy conservation equation for each species, $\dot{\rho}_X = -3H(\rho_X + p_X)$, valid to first order in perturbations. It follows that each species of radiation has a common density contrast δ_r, and each species of matter has also a common density contrast δ_m, with the relation $\delta_m = \frac{3}{4}\delta_r$.

Given the adiabatic condition, the transfer function is determined by the physical processes occuring between horizon entry and matter domination. If the radiation behaves like a perfect fluid, its density perturbation oscillates during this era, with decreasing amplitude. The matter density contrast living in this background does not grow appreciably before matter domination because it has negligible self-gravity. The transfer function is therefore given roughly by, see Eq. (102),

$$T(k) = \begin{cases} 1 , & k \ll k_{\rm eq} \\ (k/k_{\rm eq})^2 , & k \gg k_{\rm eq} \end{cases} \tag{243}$$

The perfect fluid description of the radiation is far from being correct after horizon entry, because roughly half of the radiation consists of neutrinos whose perturbation rapidly disappears through free streeming. The photons are also not a perfect fluid because they diffuse significantly, for scales below the Silk scale, $k_S^{-1} \sim 1$ Mpc. One might then consider the opposite assumption, that the radiation has zero perturbation after horizon entry. Then the matter density perturbation evolves according to

$$\ddot{\delta}_k + 2H\dot{\delta}_k + (c_s^2\, k_{\rm ph}^2 - 4\pi G\rho)\, \delta_k = 0 , \tag{244}$$

which corresponds to the equation of a damped harmonic oscillator. The zero-frequency oscillator defines the Jeans wavenumber, $k_J = \sqrt{4\pi G\rho/c_s^2}$. For $k \ll k_J$, δ_k grows exponentially on the dynamical timescale, $\tau_{\rm dyn} = {\rm Im}\,\omega^{-1} = (4\pi G\rho)^{-1/2} = \tau_{\rm grav}$, which is the time scale for gravitational collapse. One can also define the Jeans length,

$$\lambda_J = \frac{2\pi}{k_J} = c_s \sqrt{\frac{\pi}{G\rho}} , \tag{245}$$

which separates gravitationally stable from unstable modes. If we define the pressure response timescale as the size of the perturbation over the sound speed, $\tau_{\text{pres}} \sim \lambda/c_s$, then, if $\tau_{\text{pres}} > \tau_{\text{grav}}$, gravitational collapse of a perturbation can occur before pressure forces can response to restore hydrostatic equilibrium (this occurs for $\lambda > \lambda_J$). On the other hand, if $\tau_{\text{pres}} < \tau_{\text{grav}}$, radiation pressure prevents gravitational collapse and there are damped acoustic oscillations (for $\lambda < \lambda_J$).

We will consider now the behaviour of modes within the horizon during the transition from the radiation ($c_s^2 = 1/3$) to the matter era ($c_s^2 = 0$). The growing mode solution increases only by a factor of 2 between horizon entry and the epoch when matter starts to dominate, i.e. $y = 1$. The transfer function is therefore again roughly given by Eq. (243). Since the radiation consists roughly half of neutrinos, which free streem, and half of photons, which either form a perfect fluid or just diffuse, neither the perfect fluid nor the free-streeming approximation looks very sensible. A more precise calculation is needed, including: neutrino free streeming around the epoch of horizon entry; the diffusion of photons around the same time, for scales below Silk scale; the diffusion of baryons along with the photons, and the establishment after matter domination of a common matter density contrast, as the baryons fall into the potential wells of cold dark matter. All these effects apply separately, to first order in the perturbations, to each Fourier component, so that a linear transfer function is produced. There are several parametrizations in the literature, but the one which is more widely used is that of Ref. [105],

$$T(k) = \left[1 + \left(ak + (bk)^{3/2} + (ck)^2 \right)^{\nu} \right]^{-1/\nu}, \qquad \nu = 1.13, \qquad (246)$$

$$a = 6.4\,(\Omega_M h)^{-1}\,h^{-1}\,\text{Mpc}, \qquad (247)$$

$$b = 3.0\,(\Omega_M h)^{-1}\,h^{-1}\,\text{Mpc}, \qquad (248)$$

$$c = 1.7\,(\Omega_M h)^{-1}\,h^{-1}\,\text{Mpc}. \qquad (249)$$

We see that the behaviour estimated in Eq. (243) is roughly correct, although the break at $k = k_{\text{eq}}$ is not at all sharp, see Fig. 28. The transfer function, which encodes the soltion to linear equations, ceases to be valid when the density contrast becomes of order 1. After that, the highly nonlinear phenomenon of gravitational collapse takes place, see Fig. 28.
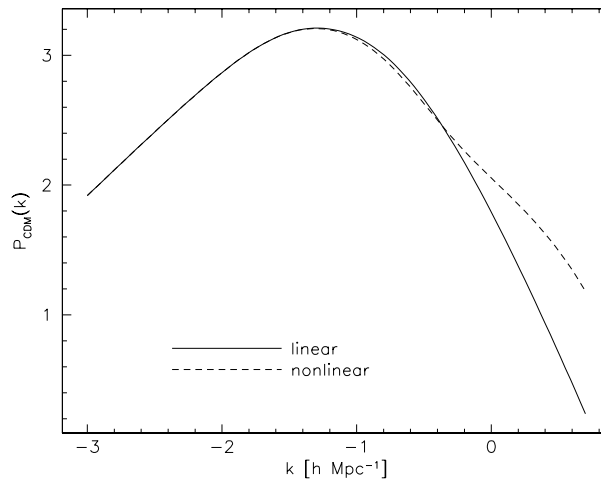
Fig. 28: The CDM power spectrum $P(k)$ as a function of wavenumber k, in logarithmic scale, normalized to the local abundance of galaxy clusters, for an Einstein-de Sitter universe with $h = 0.5$. The solid (dashed) curve shows the linear (non-linear) power spectrum. While the linear power spectrum falls off like k^{-3}, the non-linear power-spectrum illustrates the increased power on small scales due to non-linear effects, at the expense of the large-scale structures. From Ref. [45].

4.5.2   The new redshift catalogs, 2dF and Sloan Digital Sky Survey

Our view of the large-scale distribution of luminous objects in the universe has changed dramatically during the last 25 years: from the simple pre-1975 picture of a distribution of field and cluster galaxies, to the discovery of the first single superstructures and voids, to the most recent results showing an almost regular web-like network of interconnected clusters, filaments and walls, separating huge nearly empty volumes. The increased efficiency of redshift surveys, made possible by the development of spectrographs and – specially in the last decade – by an enormous increase in multiplexing gain (i.e. the ability to collect spectra of several galaxies at once, thanks to fibre-optic spectrographs), has allowed us not only to do *cartography* of the nearby universe, but also to statistically characterize some of its properties, see Ref. [106]. At the same time, advances in theoretical modeling of the development of structure, with large high-resolution gravitational simulations coupled to a deeper yet limited understanding of how to form galaxies within the dark matter halos, have provided a more realistic connection of the models to the observable quantities [107]. Despite the large uncertainties that still exist, this has transformed the study of cosmology and large-scale structure into a truly quantitative science, where theory and observations can progress side by side.

I will concentrate on two of the new catalogs, which are taking data at the moment and which have changed the field, the 2-degree-Field (2dF) Catalog and the Sloan Digital Sky Survey (SDSS). The advantages of multi-object fibre spectroscopy have been pushed to the extreme with the construction of the 2dF spectrograph for the prime focus of the Anglo-Australian Telescope [46]. This instrument is able to accommodate 400 automatically positioned fibres over a 2 degree in diameter field. This implies a density of fibres on the sky of approximately 130 deg^{-2}, and an optimal match to the galaxy counts for a magnitude $b_J \simeq 19.5$, similar to that of previous surveys like the ESP, with the difference that with such an area yield, the same number of redshifts as in the ESP survey can be collected in about 10 exposures, or slightly more than one night of telescope time with typical 1 hour exposures. This is the basis of the 2dF galaxy redshift survey. Its goal is to measure redshifts for more than 250,000 galaxies with $b_J < 19.5$. In addition, a faint redshift survey of 10,000 galaxies brighter than $R = 21$ will be done over selected fields within the two main strips of the South and North Galactic Caps. The survey has now finished, with a quarter of a million redshifts. The final result can be seen in Ref. [46].

The most ambitious and comprehensive galaxy survey currently in progress is without any doubt the Sloan Digital Sky Survey [47]. The aim of the project is, first of all, to observe photometrically the whole Northern Galactic Cap, $30°$ away from the galactic plane (about 10^4 deg^2) in five bands, at limiting magnitudes from 20.8 to 23.3. The expectation is to detect around 50 million galaxies and around 10^8 star-like sources. This has already led to the discovery of several high-redshift ($z > 4$) quasars, including the highest-redshift quasar known, at $z = 5.0$, see Ref. [47]. Using two fibre spectrographs carrying 320 fibres each, the spectroscopic part of the survey will then collect spectra from about 10^6 galaxies with $r' < 18$ and 10^5 AGNs with $r' < 19$. It will also select a sample of about 10^5 red luminous galaxies with $r' < 19.5$, which will be observed spectroscopically, providing a nearly volume-limited sample of early-type galaxies with a median redshift of $z \simeq 0.5$, that will be extremely valuable to study the evolution of clustering. The data that is coming from these catalogs is so outstanding that already cosmologists are using them for the determination of the cosmological parameters of the standard model of cosmology. The main outcome of these catalogs is the linear power spectrum of matter fluctuations that give rise to galaxies, and clusters of galaxies. It covers from the large scales of order Gigaparsecs, the realm of the unvirialised superclusters, to the small scales of hundreds of kiloparsecs, where the Lyman-α systems can help reconstruct the linear power spectrum, since they are less sensitive to the nonlinear growth of perturbations.

As often happens in particle physics, not always are observations from a single experiment sufficient to isolate and determine the precise value of the parameters of the standard model. We mentioned in the previous Section that some of the cosmological parameters created similar effects in the temperature anisotropies of the microwave background. We say that these parameters are *degenerate* with

respect to the observations. However, often one finds combinations of various experiments/observations which break the degeneracy, for example by depending on a different combination of parameters. This is precisely the case with the cosmological parameters, as measured by a combination of large-scale structure observations, microwave background anisotropies, Supernovae Ia observations and Hubble Space Telescope measurements. It is expected that in the near future we will be able to determine the parameters of the standard cosmological model with great precision from a combination of several different experiments.

5   Conclusions

In the last ten years we have seen a true revolution in the quality and quantity of cosmological data that has allowed cosmologists to determine most of the cosmological parameters with a few percent accuracy and thus fix a Standard Model of Cosmology. The art of measuring the cosmos has developed so rapidly and efficiently that one may be temped of renaming this science as Cosmonomy, leaving the word Cosmology for the theories of the Early Universe. In summary, we now know that the stuff we are made of − baryons − constitutes just about 4% of all the matter/energy in the Universe, while 25% is dark matter − perhaps a new particle species related to theories beyond the Standard Model of Particle Physics −, and the largest fraction, 70%, some form of diffuse tension also known as dark energy − perhaps a cosmological constant. The rest, about 1%, could be in the form of massive neutrinos.

Nowadays, a host of observations − from CMB anisotropies and large scale structure to the age and the acceleration of the universe − all converge towards these values, see Fig. 25. Fortunately, we will have, within this decade, new satellite experiments like Planck, CMBpol, SNAP as well as deep galaxy catalogs from Earth, to complement and precisely pin down the values of the Standard Model cosmological parameters below the percent level, see Table 2.

All these observations would not make much sense without the encompassing picture of the inflationary paradigm that determines the homogeneous and isotropic background on top of which it imprints an approximately scale invariant gaussian spectrum of adiabatic fluctuations. At present all observations are consistent with the predictions of inflation and hopefully in the near future we may have information, from the polarization anisotropies of the microwave background, about the scale of inflation, and thus about the physics responsible for the early universe dynamics.

Acknowledgements

References

[1] A. Einstein, Sitz. Preuss. Akad. Wiss. Phys. **142** (1917) (§4); Ann. Phys. **69** (1922) 436.

[2] A. Friedmann, Z. Phys. **10** (1922) 377.

[3] E.P. Hubble, Publ. Nat. Acad. Sci. **15** (1929) 168.

[4] G. Gamow, Phys. Rev. **70** (1946) 572; Phys. Rev. **74** (1948) 505.

[5] A.A. Penzias and R.W. Wilson, Astrophys. J. **142** (1965) 419.

[6] S. Weinberg, *Gravitation and Cosmology* (John Wiley & Sons, San Francisco, 1972).

[7] S. Perlmutter *et al.* [Supernova Cosmology Project], Astrophys. J. **517** (1999) 565. Home Page http://scp.berkeley.edu/

[8] W. L. Freedman *et al.*, Astrophys. J. **553** (2001) 47.

[9] A. G. Riess *et al.* [High-z Supernova Search], Astron. J. **116** (1998) 1009. Home Page http://cfa-www.harvard.edu/cfa/oir/Research/supernova/

[10] J. García-Bellido in *European School of High Energy Physics 2004*, ed. R. Fleischer (CERN report 2006-003); e-print Archive: astro-ph/0502139.

[11] R. A. Knop *et al.* [Supernova Cosmology Project Collaboration], Astrophys. J. **598** (2003) 102 [astro-ph/0309368].

[12] A. G. Riess *et al.* [Supernova Search Team Collaboration], Astrophys. J. **607** (2004) 665 [astro-ph/0402512].

[13] R. Amanullah, *et al.*, Astrophys. J. **716** (2010) 712 [arXiv:1004.1711 [astro-ph.CO]].

[14] S. Weinberg, Rev. Mod. Phys. **61** (1989) 1; S. M. Carroll, Living Rev. Rel. **4** (2001) 1; T. Padmanabhan, Phys. Rept. **380** (2003) 235; P. J. E. Peebles and B. Ratra, Rev. Mod. Phys. **75** (2003) 559.

[15] The SuperNova/Acceleration Probe Home page: `http://snap.lbl.gov/`

[16] E.W. Kolb and M.S. Turner, "The Early Universe", Addison Wesley (1990).

[17] R. Srianand, P. Petitjean and C. Ledoux, Nature **408** (2000) 931.

[18] S. Burles, K.M. Nollett, J.N. Truran, M.S. Turner, Phys. Rev. Lett. **82** (1999) 4176; S. Burles, K.M. Nollett, M.S. Turner, "Big-Bang Nucleosynthesis: Linking Inner Space and Outer Space", e-print Archive: astro-ph/9903300.

[19] K. A. Olive, G. Steigman and T. P. Walker, Phys. Rept. **333** (2000) 389; J. P. Kneller and G. Steigman, "BBN For Pedestrians," New J. Phys. **6** (2004) 117.

[20] Particle Data Group Home Page, `http://pdg.web.cern.ch/pdg/`

[21] D. N. Spergel *et al.*, Astrophys. J. Suppl. **148** (2003) 175.

[22] J.C. Mather et al., Astrophys. J. **512** (1999) 511.

[23] R.H. Dicke, P.J.E. Peebles, P.G. Roll and D.T. Wilkinson, Astrophys. J. **142** (1965) 414.

[24] C.L. Bennett et al., Astrophys. J. **464** (1996) L1.

[25] P.J.E. Peebles, "Principles of Physical Cosmology", Princeton U.P. (1993).

[26] T. Padmanabhan, "Structure Formation in the Universe", Cambridge U.P. (1993).

[27] E.R. Harrison, Phys. Rev. D **1** (1970) 2726; Ya. B. Zel'dovich, Astron. Astrophys. **5** (1970) 84.

[28] The IRAS Point Source Catalog Web page:
`http://www-astro.physics.ox.ac.uk/~wjs/pscz.html`

[29] P.J. Steinhardt, in Particle and Nuclear Astrophysics and Cosmology in the Next Millennium, ed. by E.W. kolb and R. Peccei (World Scientific, Singapore, 1995).

[30] W.L. Freedman, "Determination of cosmological parameters", Nobel Symposium (1998), e-print Archive: hep-ph/9905222.

[31] S. Refsdael, Mon. Not. R. Astr. Soc. **128** (1964) 295; **132** (1966) 101.

[32] R.D. Blandford and T. Kundić, "Gravitational Lensing and the Extragalactic Distance Scale", e-print Archive: astro-ph/9611229.

[33] N.A. Grogin and R. Narayan, Astrophys. J. **464** (1996) 92.

[34] M. Birkinshaw, Phys. Rep. **310** (1999) 97.

[35] The Chandra X-ray observatory Home Page: `http://chandra.harvard.edu/`

[36] W. L. Freedman *et al.*, Astrophys. J. **553** (2001) 47

[37] F. Zwicky, Helv. Phys. Acata **6** (1933) 110.

[38] H. Babcock, Pub. Astr. Soc. Pacific **50** (1938) 174.

[39] K. C. Freeman, Astrophys. J. **160** (1970) 811.

[40] V. C. Rubin, Science **209** (1980) 63; V. C. Rubin, N. Thonnard and W. K. . Ford, Astrophys. J. **238** (1980) 471.

[41] M. Persic, P. Salucci and F. Stel, Mon. Not. Roy. Astron. Soc. **281** (1996) 27.

[42] C.M. Baugh et al., "Ab initio galaxy formation", e-print Archive: astro-ph/9907056; Astrophys. J. **498** (1998) 405.

[43] F. Prada *et al.*, Astrophys. J. **598** (2003) 260.

[44] C.L. Sarazin, Rev. Mod. Phys. **58** (1986) 1.

[45] M. Bartelmann et al., Astron. & Astrophys. **330** (1998) 1; M. Bartelmann and P. Schneider, Phys. Rept. **340** (2001) 291

[46] M. Colless *et al.* [2dFGRS Coll.], "The 2dF Galaxy Redshift Survey: Final Data Release," Archive: astro-ph/0306581. The 2dFGRS Home Page: http://www.mso.anu.edu.au/2dFGRS/

[47] M. Tegmark *et al.* [SDSS Collaboration], Astrophys. J. **606** (2004) 702; Phys. Rev. D **69** (2004) 103501. The SDSS Home Page: `http://www.sdss.org/sdss.html`

[48] G.G. Raffelt, "Dark Matter: Motivation, Candidates and Searches", European Summer School of High Energy Physics 1997. CERN Report pp. 235-278, e-print Archive: hep-ph/9712538.

[49] P.J.E. Peebles, "Testing GR on the Scales of Cosmology," e-print Archive: astro-ph/0410284.

[50] M. C. Gonzalez-Garcia, "Global analysis of neutrino data," e-print Archive: hep-ph/0410030.

[51] S.D. Tremaine and J.E. Gunn, Phys. Rev. Lett. **42** (1979) 407; J. Madsen, Phys. Rev. D **44** (1991) 999.

[52] J. Primack, D. Seckel and B. Sadoulet, Ann. Rev. Nucl. Part. Sci. **38** (1988) 751; N.E. Booth, B. Cabrera and E. Fiorini, Ann. Rev. Nucl. Part. Sci. **46** (1996) 471.

[53] C. Kraus *et al.*, "Final results from phase II of the Mainz neutrino mass search in tritium beta decay," e-print Archive: hep-ex/0412056.

[54] H. V. Klapdor-Kleingrothaus et al., Mod. Phys. Lett. A **16** (2001) 2409; Mod. Phys. Lett. A **18** (2003) 2243.

[55] R. Bernabei, *et al.*, Int. J. Mod. Phys. A **28** (2013) 1330022
R. Bernabei *et al.*, "Dark matter search," Riv. Nuovo Cim. **26N1** (2003) 1. DAMA Home Page, `http://www.lngs.infn.it/lngs/htexts/dama/welcome.html`

[56] K. A. Olive, "Dark matter candidates in supersymmetric models," e-print Archive: hep-ph/0412054.

[57] R. Agnese *et al.* [SuperCDMS Collaboration], Phys. Rev. Lett. **112** (2014) 24, 241302.

[58] B. Ahmed *et al.*, Nucl. Phys. Proc. Suppl. **124** (2003) 193; Astropart. Phys. **19** (2003) 691. UKDMC Home Page at `http://hepwww.rl.ac.uk/ukdmc/`

[59] M. Bravin et al., Astropart. Phys. **12** (1999) 107.

[60] J. I. Collar et al., Phys. Rev. Lett. **85** (2000) 3083.

[61] G. Jungman, M. Kamionkowski and K. Griest, Phys. Rep. **267** (1996) 195.

[62] The Alpha Magnetic Spectrometer Home Page: `http://ams.cern.ch/AMS/`

[63] F. Halzen et al., Phys. Rep. **307** (1998) 243.

[64] D.A. Vandenberg, M. Bolte and P.B. Stetson, Ann. Rev. Astron. Astrophys. **34** (1996) 461; e-print Archive: astro-ph/9605064.

[65] L. M. Krauss, Phys. Rept. **333** (2000) 33.

[66] B. Chaboyer, P. Demarque, P.J. Kernan and L.M. Krauss, Science **271** (1996) 957; Astrophys. J. **494** (1998) 96.

[67] C.H. Lineweaver, Science **284** (1999) 1503.

[68] D. Scott, J. Silk and M. White, Science **268** (1995) 829; W. Hu, N. Sugiyama and J. Silk, Nature **386** (1997) 37; E. Gawiser and J. Silk, Phys. Rept. **333** (2000) 245.

[69] J. Silk, Nature **215** (1967) 1155.

[70] A. Albrecht, R. A. Battye and J. Robinson, Phys. Rev. Lett. **79** (1997) 4736; N. Turok, U. L. Pen, U. Seljak, Phys. Rev. D **58** (1998) 023506; L. Pogosian, Int. J. Mod. Phys. A **16S1C** (2001) 1043.

[71] U. Seljak et al., [SDSS Collaboration], Phys. Rev. D **71** (2005) 103515.

[72] V. Barger, D. Marfatia and A. Tregre, Phys. Lett. B **595** (2004) 55; P. Crotty, J. Lesgourgues and S. Pastor, Phys. Rev. D **69** (2004) 123007; S. Hannestad, Nucl. Phys. Proc. Suppl. **145** (2005) 313.

[73] H. V. Peiris *et al.*, Astrophys. J. Suppl. **148** (2003) 213; P. Crotty, J. García-Bellido, J. Lesgourgues and A. Riazuelo, Phys. Rev. Lett. **91** (2003) 171301; J. Valiviita and V. Muhonen, Phys. Rev. Lett. **91** (2003) 131302; M. Beltrán, J. García-Bellido, J. Lesgourgues and A. Riazuelo, Phys. Rev. D **70** (2004) 103530; K. Moodley, M. Bucher, J. Dunkley, P. G. Ferreira and C. Skordis, Phys. Rev. D **70** (2004) 103520; C. Gordon and K. A. Malik, Phys. Rev. D **69** (2004) 063508; F. Ferrer, S. Rasanen and J. Valiviita, JCAP **0410** (2004) 010; H. Kurki-Suonio, V. Muhonen and J. Valiviita, Phys. Rev. D **71** (2005) 063005; M. Beltrán, J. García-Bellido, J. Lesgourgues, A. R. Liddle and A. Slosar, Phys. Rev. D **71** (2005) 063532; M. Beltran, J. Garcia-Bellido, J. Lesgourgues and M. Viel, Phys. Rev. D **72** (2005) 103515; M. Beltran, J. Garcia-Bellido and J. Lesgourgues, Phys. Rev. D **75** (2007) 103507.

[74] J. García-Bellido, Phil. Trans. R. Soc. Lond. **A 357** (1999) 3237.

[75] A. Guth, Phys. Rev. D **23** (1981) 347.

[76] A.D. Linde, Phys. Lett. **108B** (1982) 389.

[77] A. Albrecht and P.J. Steinhardt, Phys. Rev. Lett. **48** (1982) 1220.

[78] For a personal historical account, see A. Guth, "The Inflationary Universe", Perseus Books (1997).

[79] A.D. Linde, "Particle Physics and Inflationary Cosmology", Harwood Academic Press (1990).

[80] A.R. Liddle and D.H. Lyth, Phys. Rep. **231** (1993) 1.

[81] J.M. Bardeen, Phys. Rev. D **22** (1980) 1882.

[82] V.F. Mukhanov, H.A. Feldman and R.H. Brandenberger, Phys. Rep. **215** (1992) 203.

[83] M. Abramowitz and I. Stegun, "Handbook of Mathematical Functions", Dover (1972).

[84] J. García-Bellido and D. Wands, Phys. Rev. D **53** (1996) 5437; D. Wands, K. A. Malik, D. H. Lyth and A. R. Liddle, Phys. Rev. D **62** (2000) 043527.

[85] D.H. Lyth and A. Riotto, Phys. Rep. **314** (1999) 1.

[86] R.K. Sachs and A.M. Wolfe, Astrophys. J. **147** (1967) 73.

[87] E.R. Harrison, Rev. Mod. Phys. **39** (1967) 862; L.F. Abbott and R.K. Schaefer, Astrophys. J. **308** (1986) 546.

[88] E.F. Bunn, A.R. Liddle and M. White, Phys. Rev. D **54** (1996) 5917.

[89] A.A. Starobinsky, Sov. Astron. Lett. **11** (1985) 133.

[90] S.M. Carroll, W.H. Press and E.L. Turner, Ann. Rev. Astron. Astrophys. **30** (1992) 499.

[91] P. de Bernardis *et al.*, New Astron. Rev. **43** (1999) 289; P. D. Mauskopf *et al.* [Boomerang Coll.], Astrophys. J. **536** (2000) L59. Boomerang Home Page: http://oberon.roma1.infn.it/boomerang/

[92] Microwave Anisotropy Probe Home Page: `http://map.gsfc.nasa.gov/`

[93] Planck Surveyor Home Page: `http://www.cosmos.esa.int/web/planck/publications`

[94] M. Tegmark Home Page: http://space.mit.edu/home/tegmark/

[95] M. Kamionkowski and A. Kosowsky, Ann. Rev. Nucl. Part. Sci. **49** (1999) 77.

[96] W. Hu and S. Dodelson, Ann. Rev. Astron. Astrophys. **40** (2002) 171.

[97] U. Seljak and M. Zaldarriaga, CMBFAST code Home Page: `http://www.cmbfast.org/`

[98] A. Lewis and A. Challinor, CAMB code Home Page: `http://camb.info/`

[99] P. Andre *et al.* [PRISM Collaboration], arXiv:1306.2259 [astro-ph.CO]; JCAP **1402** (2014) 006 [arXiv:1310.1554 [astro-ph.CO]].
PRISM experiment Home Page: `http://www.prism-mission.org/`

[100] F. R. Bouchet *et al.* [COrE Collaboration], "COrE (Cosmic Origins Explorer) A White Paper," arXiv:1102.2181 [astro-ph.CO]; CORE experiment Home Page:

`http://www.core-mission.org/`

[101] L.A. Page, "Measuring the anisotropy in the CMB", e-print Archive: astro-ph/9911199.

[102] J. Kovac et al., Nature **420** (2002) 772; DASI Home Page: http://astro.uchicago.edu/dasi/

[103] BICEP & Keck Array Home Page, `http://bicepkeck.org/`

[104] A.R. Liddle and D.H. Lyth, "Cosmological Inflation and Large Scale Structure", Cambridge University Press (2000).

[105] J.R. Bond and G. Efstathiou, Astrophys. J. **285** (1984) L45.

[106] G. Efstathiou *et al.* (Eds.) "Large-scale structure in the universe", Phil. Trans. R. Soc. Lond. **A 357** (1999) 1-198.

[107] B. Moore, Phil. Trans. R. Soc. Lond. **A 357** (1999) 3259.