Dark Energy

N. Straumann

Institute for Theoretical Physics, University of Zurich, Winterthurerstrasse 180, 8057 Zurich, Switzerland
norbert.straumann@freesurf.ch

1 Introduction

Cosmology is going through a fruitful and exciting period. Some of the developments are definitely also of interest to physicists outside the fields of astrophysics and cosmology.

This chapter covers some particularly fascinating and topical subjects. A central theme will be the current evidence that the recent ($z < 1$) Universe is dominated by an exotic nearly homogeneous dark energy density with *negative* pressure. The simplest candidate for this unknown so-called *dark energy* is a cosmological term in Einstein's field equations, a possibility that has been considered during all the history of relativistic cosmology. Independently of what this exotic energy density is, one thing is certain since a long time: The energy density belonging to the cosmological constant is not larger than the cosmological critical density, and thus *incredibly small by particle physics standards*. This is a profound mystery, since we expect that all sorts of *vacuum energies* contribute to the effective cosmological constant.

Since this is such an important issue it should be of interest to indicate how convincing the evidence for this finding really is, or whether one should remain skeptical. Much of this is based on the observed temperature fluctuations of the cosmic microwave background radiation (CMB), and large-scale structure formation. The first evidence for an accelerating expansion of the Universe, and still the only direct one, came from the Hubble diagram for Type Ia supernovae. When combined with other measurements a cosmological world model of the Friedmann–Lemaître variety has emerged that is spatially almost flat, with about 70% of its energy contained in the form dark energy. A detailed analysis of the existing data requires a considerable amount of theoretical machinery that is beyond the scope of this contribution. For interested readers we shall refer to some books, reviews, and articles that may be most convenient to penetrate deeper into various topics.

N. Straumann: *Dark Energy*, Lect. Notes Phys. **721**, 327–397 (2007)
DOI 10.1007/978-3-540-71117-9_13

Since this book addresses mostly readers whose main interests are outside astrophysics and cosmology, I do not presuppose a serious training in cosmology. However, I do assume some working knowledge of general relativity (GR). As a source, and for references, I usually quote my recent textbook [1]. The essentials of the Friedmann–Lemaître models will be summarized in Appendices A and B. Appendix C provides a brief introduction to *inflation*, a key idea of modern cosmology.

2 Einstein's Original Motivation of the Λ-Term

One of the contributions in the famous book *Albert Einstein: Philosopher–Scientist* [2] is a chapter by George E. Lemaître entitled "The Cosmological Constant". In the introduction he says: *"The history of science provides many instances of discoveries which have been made for reasons which are no longer considered satisfactory. It may be that the discovery of the cosmological constant is such a case."* When the book appeared in 1949 – at the occasion of Einstein's seventieth birthday – Lemaître could not be fully aware of how right he was, how profound the cosmological constant problem really is, especially since he was not a quantum physicist.

We begin this contribution in reviewing the main aspects of the history of the Λ-term, from its introduction in 1917 up to the point when it became widely clear that we are facing a deep mystery. (See also [3] and [4].) I describe first the *classical* aspect of the historical development.

Einstein introduced the cosmological term when he applied GR the first time to cosmology [5]. Presumably the main reason why Einstein turned so soon after the completion of GR to cosmology had much to do with Machian ideas on the origin of inertia, which played in those years an important role in Einstein's thinking. His intention was to eliminate all vestiges of absolute space. He was, in particular, convinced that isolated masses cannot impose a structure on space at infinity. Einstein was actually thinking about the problem regarding the choice of boundary conditions at infinity already in spring 1916. In a letter to Michele Besso on 14 May 1916 he also mentions the possibility of the world being finite. A few months later he expanded on this in letters to Willem de Sitter. It is along these lines that he postulated a Universe that is spatially finite and closed, a Universe in which no boundary conditions are needed. He then believed that this was the only way to satisfy what he later [7] named *Mach's principle*, in the sense that the metric field should be determined uniquely by the energy-momentum tensor.

In addition, Einstein assumed that the Universe was *static*. This was not unreasonable at the time, because the relative velocities of the stars as observed were small. (Recall that astronomers only learned later that spiral nebulae are independent star systems outside the Milky Way. This was definitely established when in 1924 Hubble found that there were Cepheid variables in Andromeda and also in other galaxies.)

These two assumptions were, however, not compatible with Einstein's original field equations. For this reason, Einstein added the famous Λ-term, which is compatible with the principles of GR, in particular with the energy–momentum law $\nabla_\nu T^{\mu\nu} = 0$ for matter. The modified field equations in standard notation and signature $(- + ++)$ are

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu} \ . \tag{1}$$

The cosmological term is, in four dimensions, the only possible complication of the field equations if no higher than second-order derivatives of the metric are allowed (*Lovelock theorem*). This remarkable uniqueness is one of the most attractive features of GR. (In higher dimensions additional terms satisfying this requirement are allowed.)

For the static Einstein universe the field equations (1) imply the two relations

$$4\pi G\rho = \frac{1}{a^2} = \Lambda \ , \tag{2}$$

where ρ is the mass density of the dust-filled universe (zero pressure) and a is the radius of curvature. (We remark, in passing, that the Einstein universe is the only static dust solution; one does not have to assume isotropy or homogeneity. Its instability was demonstrated by Lemaître in 1927.) Einstein was very pleased by this direct connection between the mass density and geometry, because he thought that this was in accord with Mach's philosophy.

Einstein concludes with the following sentences:

> *In order to arrive at this consistent view, we admittedly had to introduce an extension of the field equations of gravitation which is not justified by our actual knowledge of gravitation. It has to be emphasized, however, that a positive curvature of space is given by our results, even if the supplementary term is not introduced. That term is necessary only for the purpose of making possible a quasi-static distribution of matter, as required by the fact of the small velocities of the stars.*

To de Sitter, Einstein emphasized in a letter on 12 March 1917 that his cosmological model was intended primarily to settle the question "whether the basic idea of relativity can be followed through its completion, or whether it leads to contradictions". And he adds whether the model corresponds to reality was another matter.

Only later Einstein came to realize that Mach's philosophy is predicated on an antiquated ontology that seeks to reduce the metric field to an epiphenomenon of matter. It became increasingly clear to him that the metric field has an independent existence, and his enthusiasm for what he called Mach's principle later decreased. In a letter to F. Pirani he wrote in 1954: *"As a matter of fact, one should no longer speak of Mach's principle at all"* [8]. GR still preserves some remnant of Newton's absolute space and time.

3 From Static to Expanding World Models

Surprisingly to Einstein, de Sitter discovered in the same year, 1917, a completely different static cosmological model which also incorporated the cosmological constant, but was *anti-Machian*, because it contained no matter [9]. For this reason, Einstein tried to discard it on various grounds (more on this below). The original form of the metric was

$$g = -\Big[1 - (\frac{r}{R})^2\Big]dt^2 + \frac{dr^2}{1-(\frac{r}{R})^2} + r^2(d\vartheta^2 + \sin^2\vartheta d\varphi^2) \ .$$

Here, the spatial part is the standard metric of a three-sphere of radius R, with $R = (3/\Lambda)^{1/2}$. The model had one very interesting property: For light sources moving along static worldlines there is a gravitational redshift, which became known as the *de Sitter effect*. This was thought to have some bearing on the redshift results obtained by Slipher. Because the fundamental (static) worldlines in this model are not geodesic, a freely falling object released by any static observer will be seen by him to accelerate away, generating also local velocity (Doppler) redshifts corresponding to *peculiar velocities*. In the second edition of his book [10], published in 1924, Eddington writes about this

> *de Sitter's theory gives a double explanation for this motion of recession; first there is a general tendency to scatter (...); second there is a general displacement of spectral lines to the red in distant objects owing to the slowing down of atomic vibrations (...), which would erroneously be interpreted as a motion of recession.*

I do not want to enter into all the confusion over the de Sitter universe. One source of this was the apparent singularity at $r = R = (3/\Lambda)^{1/2}$. This was at first thoroughly misunderstood even by Einstein and Weyl. ('The Einstein–de Sitter–Weyl–Klein Debate' is now published in Vol. 8 of the *Collected Papers* [6].) At the end, Einstein had to acknowledge that de Sitter's solution is fully regular and matter-free and thus indeed a counter example to Mach's principle. But he still discarded the solution as physically irrelevant because it is not globally static. This is clearly expressed in a letter from Weyl to Klein, after he had discussed the issue during a visit of Einstein in Zurich [11]. An important discussion of the redshift of galaxies in de Sitter's model by H. Weyl in 1923 should be mentioned. Weyl introduced an expanding version[1] of the de Sitter model [12]. For *small* distances his result reduced to what later became known as the Hubble law. Independently of Weyl, Cornelius Lanczos introduced in 1922 also a non-stationary interpretation of de Sitter's solution in the form of a Friedmann spacetime with a positive spatial curvature

[1] I recall that the de Sitter model has many different interpretations, depending on the class of fundamental observers that is singled out.

[13]. In a second paper he also derived the redshift for the non-stationary interpretation [14].

Until about 1930 almost everybody believed that the Universe was static, in spite of the two fundamental papers by Friedmann [15] in 1922 and 1924 and Lemaître's independent work [16] in 1927. These path-breaking papers were in fact largely ignored. The history of this early period has – as is often the case – been distorted by some widely read documents. Einstein too accepted the idea of an expanding Universe only much later. After the first paper of Friedmann, he published a brief note claiming an error in Friedmann's work; when it was pointed out to him that it was his error, Einstein published a retraction of his comment, with a sentence that luckily was deleted before publication: *"[Friedmann's paper] while mathematically correct is of no physical significance"*. In comments to Lemaître during the Solvay meeting in 1927, Einstein again rejected the expanding universe solutions as physically unacceptable. According to Lemaître, Einstein was telling him, *"Vos calculs sont corrects, mais votre physique est abominable."* It appears astonishing that Einstein – after having studied carefully Friedmann's papers – did not realize that his static model is unstable, and hence that the Universe has to be expanding or contracting. On the other hand, I found in the archive of the ETH many years ago a postcard of Einstein to Weyl from 1923, related to Weyl's reinterpretation of de Sitter's solution, with the following interesting sentence: *"If there is no quasi-static world, then away with the cosmological term."*

It also is not well known that Hubble interpreted his famous results on the redshift of the radiation emitted by distant "nebulae" in the framework of the de Sitter model, as was suggested by Eddington.

The general attitude is well illustrated by the following remark of Eddington at a Royal Astronomical Society meeting in January 1930: *"One puzzling question is why there should be only two solutions. I suppose the trouble is that people look for static solutions."*

Lemaître, who had been for a short time a post-doctoral student of Eddington, read this remark in a report to the meeting published in *Observatory*, and wrote to Eddington pointing out his 1927 paper. Eddington had seen that paper, but had completely forgotten about it. But now he was greatly impressed and recommended Lemaître's work in a letter to *Nature*. He also arranged for a translation which appeared in MNRAS [17]. Eddington also "pointed out that it was immediately deducible from his [Lemaître's] formulae that Einstein's world is unstable, so that an expanding or a contracting universe is an inevitable result of Einstein's law of gravitation."

Lemaître's successful explanation of Hubble's discovery finally changed the viewpoint of the majority of workers in the field. At this point, Einstein *rejected the cosmological term as superfluous and no longer justified* [18]. At the end of the paper, in which he published his new view, Einstein adds some remarks about the age problem which was quite severe without the Λ-term, since Hubble's value of the Hubble parameter was then about seven times too large. Einstein is, however, not very worried and suggests two ways out. First he says that the matter distribution is in reality inhomogeneous and that the approximate treatment may be illusionary. Then he adds that in astronomy one should be cautious with large extrapolations in time.

Einstein repeated his new standpoint also much later [19], and this was adopted by many other influential workers, e.g. by Pauli [20]. Whether Einstein really considered the introduction of the Λ-term as "the biggest blunder of his life" appears doubtful to me. In his published work and letters I never found such a strong statement. Einstein discarded the cosmological term just for simplicity reasons. For a minority of cosmologists (O. Heckmann, for example [21]), this was not sufficient reason. Paraphrasing Rabi, one might ask, "who ordered it away"?

Einstein published his new view in the *Sitzungsberichte der Preussischen Akademie der Wissenschaften.* The correct citation is,

Einstein. A. (1931). Sitzungsber. Preuss. Akad. Wiss. 235–37.

Many authors have quoted this paper but never read it. As a result, the quotations gradually changed in an interesting, quite systematic fashion. Some steps are shown in the following sequence:

- A. Einstein. 1931. Sitzsber. Preuss. Akad. Wiss. ...
- A. Einstein. Sitzber. Preuss. Akad. Wiss. ... (1931)
- A. Einstein (1931). Sber. preuss. Akad. Wiss. ...
- Einstein. A .. 1931. Sb. Preuss. Akad. Wiss. ...
- A. Einstein. S.-B. Preuss. Akad. Wis. ...1931
- A. Einstein. S.B. Preuss. Akad. Wiss. (1931) ...
- Einstein, A., and Preuss, S.B. (1931). Akad. Wiss. **235**

Presumably, one day some historian of science will try to find out what happened with the young physicist S.B. Preuss, who apparently wrote just one important paper and then disappeared from the scene.

After the Λ-force was rejected by its inventor, other cosmologists, like Eddington, retained it. One major reason was that it solved the problem of the age of the Universe when the Hubble time scale was thought to be only 2 billion years (corresponding to the value $\mathrm{H}_0 \sim 500$ km $\mathrm{s}^{-1}\mathrm{Mpc}^{-1}$ of the Hubble constant). This was even shorter than the age of the Earth. In addition, Eddington and others overestimated the age of stars and stellar systems.

For this reason, the Λ-term was employed again and a model was revived which Lemaître had singled out from the many solutions of the Friedmann–Lemaître equations.[2] This so-called "Lemaître hesitation universe" is closed and has a repulsive Λ-force ($\Lambda > 0$), which is slightly greater than the value

[2] I recall that Friedmann included the Λ-term in his basic equations. I find it remarkable that for the negatively curved solutions he pointed out that these may be open or compact (but not simply connected).

chosen by Einstein. It begins with a big bang and has the following two stages of expansion. In the first the Λ-force is not important, the expansion is decelerated due to gravity and slowly approaches the radius of the Einstein universe. At about the same time, the repulsion becomes stronger than gravity and a second stage of expansion begins which eventually inflates. In this way a positive Λ was employed to reconcile the expansion of the Universe with the age of stars.

Repulsive Effect of a Positive Cosmological Constant

The *repulsive* effect of a positive cosmological constant can be seen from the following consequence of Einstein's field equations for the time-dependent scale factor $a(t)$ (see Appendix A):

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

where p is the pressure of all forms of matter.

Historically, the Newtonian analog of the cosmological term was regarded by Einstein, Weyl, Pauli, and others as a *Yukawa term*. This is not correct, as I now show.

For a better understanding of the action of the Λ-term it may be helpful to consider a general static spacetime with the metric (in adapted coordinates)

$$ds^2 = -\varphi^2 dt^2 + g_{ik} dx^i dx^k \ , \tag{4}$$

where φ and g_{ik} depend only on the spatial coordinates x^i. The component R_{00} of the Ricci tensor is given by $R_{00} = \bar{\Delta}\varphi/\varphi$, where $\bar{\Delta}$ is the three-dimensional Laplace operator for the spatial metric g_{ik} in (4) (see, e.g., [1]). Let us write (1) in the form

$$G_{\mu\nu} = \kappa(T_{\mu\nu} + T^{\Lambda}_{\mu\nu}) \qquad (\kappa = 8\pi G) \ , \tag{5}$$

with

$$T^{\Lambda}_{\mu\nu} = -\frac{\Lambda}{8\pi G} g_{\mu\nu} \ . \tag{6}$$

This has the form of the energy–momentum tensor of an ideal fluid, with energy density $\rho_\Lambda = \Lambda/8\pi G$ and pressure $p_\Lambda = -\rho_\Lambda$.[3] For an ideal fluid at rest Einstein's field equation implies

$$\frac{1}{\varphi}\bar{\Delta}\varphi = 4\pi G\Big[(\rho + 3p) + \underbrace{(\rho_\Lambda + 3p_\Lambda)}_{-2\rho_\Lambda}\Big] \ . \tag{7}$$

Since the energy density and the pressure appear in the combination $\rho + 3p$, we understand that a positive ρ_Λ leads to a repulsion (as in (3)). In the

[3] This way of looking at the cosmological term was soon (in 1918) emphasized by Schrödinger and also by F. Klein.

Newtonian limit we have $\varphi \simeq 1 + \phi$ (ϕ : Newtonian potential) and $p \ll \rho$, hence we obtain the modified Poisson equation

$$\Delta\phi = 4\pi G(\rho - 2\rho_\Lambda) . \tag{8}$$

This is the correct Newtonian limit.

As a result of revised values of the Hubble parameter and the development of the modern theory of stellar evolution in the 1950s, the controversy over ages was resolved and the Λ-term became again unnecessary. (Some tension remained for values of the Hubble parameter at the higher end of published values.)

However, in 1967 it was revived again in order to explain why quasars appeared to have redshifts that concentrated near the value $z = 2$. The idea was that quasars were born in the hesitation era [22]. Then quasars at greatly different distances can have almost the same redshift, because the universe was almost static during that period. Other arguments in favor of this interpretation were based on the following peculiarity. When the redshifts of emission lines in quasar spectra exceed 1.95, then redshifts of absorption lines in the same spectra were, as a rule, equal to 1.95. This was then quite understandable, because quasar light would most likely have crossed intervening galaxies during the epoch of suspended expansion, which would result in almost identical redshifts of the absorption lines. However, with more observational data evidence for the Λ-term dispersed for the third time.

4 The Mystery of the Λ-Problem

At this point I want to leave the classical discussion of the Λ-term, and turn to the quantum aspect of the Λ-problem, where it really becomes very serious.

4.1 Historical Remarks

Since quantum physicists had so many other problems, it is not astonishing that in the early years they did not worry about this subject. An exception was Pauli, who wondered in the early 1920s whether the zero-point energy of the radiation field could be gravitationally effective.

As background I recall that Planck had introduced the zero-point energy with somewhat strange arguments in 1911. The physical role of the zero-point energy was much discussed in the early years of quantum theory. There was, for instance, a paper by Einstein and Stern in 1913 [Collected Papers, Vol. 4, Doc. 11; see also the Editorial Note, p. 270] that aroused widespread interest. In this, two arguments in favor of the zero-point energy were given. The first had to do with the specific heat of rotating (diatomic) molecules. The authors developed an approximate theory of the energy of rotating molecules and came to the conclusion that the resulting specific heat agreed much better with recent experimental results by Arnold Eucken, if they included the zero-point energy. The second argument was based on a new derivation of Planck's radiation formula. In both the arguments, Einstein and Stern made a number of problematic assumptions, and in fall 1913, Einstein retracted their results. At the second Solvay Congress in late October 1913, Einstein said that he no longer believed in the zero-point energy, and in a letter to Ehrenfest [Vol. 5, Doc. 481] he wrote that the zero-point energy was "dead as a doornail".

From Charly Enz and Armin Thellung – Pauli's last two assistants – I have learned that Pauli had discussed this issue extensively with O. Stern in Hamburg. Stern had calculated, but never published, the vapor pressure difference between the isotopes 20 and 22 of Neon (using Debye theory). He came to the conclusion that without zero-point energy this difference would be large enough for easy separation of the isotopes, which is not the case in reality. These considerations penetrated into Pauli's lectures on statistical mechanics [23] (which I attended). The theme was taken up in an article by Enz and Thellung [24]. This was originally written as a birthday gift for Pauli, but because of Pauli's early death this appeared in a memorial volume of Helv.Phys.Acta.

From Pauli's discussions with Enz and Thellung we know that Pauli estimated the influence of the zero-point energy of the radiation field – cutoff at the classical electron radius – on the radius of the universe, and came to the conclusion that it *"could not even reach to the moon"*.

When, as a student, I heard about this, I checked Pauli's unpublished[4] remark by doing the following little calculation (which Pauli must have done):

In units with $\hbar = c = 1$ the vacuum energy density of the radiation field is

$$\langle \rho \rangle_{vac} = \frac{8\pi}{(2\pi)^3} \int_0^{\omega_{max}} \frac{\omega}{2} \omega^2 d\omega = \frac{1}{8\pi^2} \omega_{max}^4 \ ,$$

with

$$\omega_{max} = \frac{2\pi}{\lambda_{max}} = \frac{2\pi m_e}{\alpha} \ .$$

The corresponding radius of the Einstein universe in (2) would then be ($M_{pl} \equiv 1/\sqrt{G}$)

$$a = \frac{\alpha^2}{(2\pi)^{\frac{2}{3}}} \frac{M_{pl}}{m_e} \frac{1}{m_e} \sim 31 \text{ km} \ .$$

This is indeed less than the distance to the moon. (It would be more consistent to use the curvature radius of the static de Sitter solution; the result is the same, up to the factor $\sqrt{3/2}$.)

For decades nobody else seems to have worried about contributions of quantum fluctuations to the cosmological constant, although physicists

[4] A trace of this is in Pauli's Handbuch article [25] on wave mechanics in the section where he discusses the meaning of the zero-point energy of the quantized radiation field.

learned after Dirac's hole theory that the vacuum state in quantum field theory is not an empty medium, but has interesting physical properties. As an important example I mention the papers by Heisenberg and Euler [26] in which they calculated the modifications of Maxwell's equations due to the polarization of the vacuum. Shortly afterward, Weisskopf [27] not only simplified their calculations but also gave a thorough discussion of the physics involved in charge renormalization. Weisskopf related the modification of Maxwell's Lagrangian to the change of the energy of the Dirac sea as a function of slowly varying external electromagnetic fields. (Avoiding the old-fashioned Dirac sea, this effective Lagrangian is due to the interaction of a classical electromagnetic field with the vacuum fluctuations of the electron positron field.) After a charge renormalization this change is finite and gives rise to electric and magnetic polarization vectors of the vacuum. In particular, the refraction index for light propagating perpendicular to a static homogeneous magnetic field depends on the polarization direction. This is the vacuum analog of the well-known Cotton–Mouton effect in optics. As a result, an initially linearly polarized light beam becomes elliptic. (In spite of great efforts it has not yet been possible to observe this effect.)

Another beautiful example for the importance of vacuum energies as a function of varying external conditions is the *Casimir effect*. This is the most widely cited example of how vacuum fluctuations can have observable consequences.

The presence of conducting plates modifies the vacuum energy density in a manner which depends on the separation of the plates. This leads to an attractive force between the two plates.

Historically, this was a byproduct of some applied industrial research in the stability of colloidal suspensions used to deposit films in the manufacture of lamps and cathode tubes. This lead Casimir and Polder to reconsider the theory of van der Waals interaction with *retardation* included. They found that this causes the interaction to vary at large intermolecular separations as r^{-7}. Casimir mentioned his result to Niels Bohr during a walk, and told him that he was puzzled by the extreme simplicity of the result at large distance. According to Casimir, Bohr mumbled something about zero-point energy. That was all, but it put him on the right track.

Precision experiments have recently confirmed the theoretical prediction to about 1%. By now the literature related to the Casimir effect is enormous. For further information we refer to the recent book [28].

4.2 Has Dark Energy been Discovered in the Lab?

It has been suggested by Beck and Mackey [29] that part of the zero-point energy of the radiation field that is gravitationally active can be determined from noise measurements of Josephson junctions. This caused some widespread attention. In a reaction we [30] showed that there is no basis for this claim, by following the reasoning in [29] for a much simpler model, for which it is very obvious that the authors misinterpreted their formulae. Quite generally, the absolute value of the zero-point energy of a quantum mechanical system has no physical meaning when gravitational coupling is ignored. All that is measurable are *changes* of the zero-point energy under variations of system parameters or of external couplings, like an applied voltage. For further information on the controversy, see [31] and [32].

4.3 Vacuum Energy and Gravity

When we consider the coupling to gravity, the vacuum energy density acts like a cosmological constant. In order to see this, first consider the vacuum expectation value of the energy–momentum tensor in Minkowski spacetime. Since the vacuum state is Lorentz invariant, this expectation value is an invariant symmetric tensor, hence proportional to the metric tensor. For a curved metric this is still the case, up to higher curvature terms:

$$\langle T_{\mu\nu} \rangle_{vac} = -g_{\mu\nu}\rho_{vac} + \textit{higher curvature terms}\,. \tag{9}$$

The *effective* cosmological constant, which controls the large-scale behavior of the Universe, is given by

$$\Lambda = 8\pi G \rho_{vac} + \Lambda_0 \;, \tag{10}$$

where Λ_0 is a bare cosmological constant in Einstein's field equations.

We know from astronomical observations that $\rho_\Lambda \equiv \Lambda/8\pi G$ cannot be larger than about the critical density:

$$\begin{aligned} \rho_{crit} &= \frac{3H_0^2}{8\pi G} \\ &= 1.88 \times 10^{-29} h_0^2 g\mathrm{cm}^{-3} \\ &\simeq (3 \times 10^{-3} eV)^4 \;, \end{aligned} \tag{11}$$

where h_0 is the *reduced Hubble parameter*

$$h_0 = H_0/(100\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}) \tag{12}$$

that is close to 0.7.

It is a complete mystery as to why the two terms in (10) should almost exactly cancel. This is – more precisely stated – the famous Λ-problem.

As far as I know, the first who came back to possible contributions of the vacuum energy density to the cosmological constant was Zel'dovich. He discussed this issue in two papers [33] during the third renaissance period of the Λ-term, but before the advent of spontaneously broken gauge theories. The following remark by him is particularly interesting. Even if one assumes completely ad hoc that the zero-point contributions to the vacuum energy density are exactly cancelled by a bare term, there still remain higher-order effects.

In particular, *gravitational* interactions between the particles in the vacuum fluctuations are expected on dimensional grounds to lead to a gravitational self-energy density of order $G\mu^6$, where μ is some cutoff scale. Even for μ as low as 1 GeV (for no good reason) this is about 9 orders of magnitude larger than the observational bound.

This illustrates that there is something profound that we do not understand at all, certainly not in quantum field theory (so far also not in string theory). We are unable to calculate the vacuum energy density in quantum field theories, like the standard model of particle physics. But we can attempt to make what appear to be reasonable order-of-magnitude estimates for the various contributions. All expectations are *in gigantic conflict with the facts* (see below). Trying to arrange the cosmological constant to be zero is unnatural in a technical sense. It is like enforcing a particle to be massless, by fine-tuning the parameters of the theory when there is no symmetry principle which implies a vanishing mass. The vacuum energy density is unprotected from large quantum corrections. This problem is particularly severe in field theories with spontaneous symmetry breaking. In such models there are usually several possible vacuum states with different energy densities. Furthermore, the energy density is determined by what is called the effective potential, and this is a *dynamical* object. Nobody can see any reason why the vacuum of the standard model we ended up as the Universe cooled has – for particle physics standards – an almost vanishing energy density. Most probably, we will only have a satisfactory answer once we shall have a theory which successfully combines the concepts and laws of GR about gravity and spacetime structure with those of quantum theory.

4.4 Simple Estimates of Vacuum Energy Contributions

If we take into account the contributions to the vacuum energy from vacuum fluctuations in the fields of the standard model up to the currently explored energy, i.e., about the electroweak scale $M_F = G_F^{-1/2} \approx 300 GeV (G_F$: Fermi coupling constant), we cannot expect an almost complete cancellation, because there is *no symmetry principle* in this energy range that could require this. The only symmetry principle which would imply this is *supersymmetry*, but supersymmetry is broken (if it is realized in nature). Hence we can at best expect a very imperfect cancellation below the electroweak scale, leaving a contribution of the order of M_F^4. (The contributions at higher energies may largely cancel if supersymmetry holds in the real world.)

We would reasonably expect that the vacuum energy density is at least as large as the condensation energy density of the QCD phase transition to the broken phase of chiral symmetry. Already this is far too large: $\sim \Lambda_{QCD}^4/16\pi^2 \sim 10^{-4} GeV^4$; this is *more than 40 orders of magnitude larger* than ρ_{crit}. Beside the formation of quark condensates $< \bar{q}q >$ in the QCD vacuum which break chirality, one also expects a gluon condensate $< G_a^{\mu\nu} G_{a\mu\nu} > \sim \Lambda_{QCD}^4$. This produces a significant vacuum energy density as a result of a dilatation anomaly: If Θ^μ_μ denotes the "classical" trace of the energy–momentum tensor, we have [34]

$$T^\mu_\mu = \Theta^\mu_\mu - \frac{\beta(g_s)}{2g_s} G^{\mu\nu}_a G_{a\mu\nu} \ , \tag{13}$$

where the second term is the QCD piece of the trace anomaly. $\beta(g_s)$ is the β-function of QCD that determines the running of the strong coupling constant g_s (see the contribution of Dosch to this book). I recall that this anomaly arises because a scale transformation is no more a symmetry if quantum corrections are included. Taking the vacuum expectation value of (13), we would again naively expect that $< \Theta^\mu_\mu >$ is of the order M_F^4. Even if this should vanish for some unknown reason, the anomalous piece is cosmologically gigantic. The expectation value $< G^{\mu\nu}_a G_{a\mu\nu} >$ can be estimated with QCD sum rules [35], and gives

$$< T^\mu_\mu >^{anom} \sim -(350 MeV)^4 \ , \tag{14}$$

about 45 orders of magnitude larger than ρ_{crit}. This reasoning should show convincingly that the cosmological constant problem is indeed a profound one. (Note that there is some analogy with the (much milder) strong CP problem of QCD. However, in contrast to the Λ-problem, Peccei and Quinn [36] have shown that in this case there is a way to resolve the conundrum.)

Let us also have a look at the Higgs condensate of the electroweak theory. Recall that in the standard model we have for the Higgs doublet Φ in the broken phase for $< \Phi^* \Phi > \equiv \frac{1}{2}\phi^2$ the potential

$$V(\phi) = -\frac{1}{2} m^2 \phi^2 + \frac{\lambda}{8}\phi^4 \ . \tag{15}$$

Setting as usual $\phi = v + H$, where v is the value of ϕ where V has its minimum,

$$v = \sqrt{\frac{2m^2}{\lambda}} = 2^{-1/4} G_F^{-1/2} \sim 246 GeV \ , \tag{16}$$

we find that the Higgs mass is related to λ by $\lambda = M_H^2/v^2$. For $\phi = v$ we obtain the energy density of the Higgs condensate

$$V(\phi = v) = -\frac{m^4}{2\lambda} = -\frac{1}{8\sqrt{2}} M_F^2 M_H^2 = \mathcal{O}(M_F^4) \ . \tag{17}$$

We can, of course, add a constant V_0 to the potential (15) such that it cancels the Higgs vacuum energy in the broken phase – including higher-order corrections. This again requires an extreme fine tuning. A remainder of only $\mathcal{O}(m_e^4)$, say, would be catastrophic. This remark is also highly relevant for models of inflation and quintessence.

In attempts beyond the standard model the vacuum energy problem so far remains, and often becomes even worse. For instance, in supergravity theories with spontaneously broken supersymmetry there is the following simple relation between the gravitino mass m_g and the vacuum energy density

$$\rho_{vac} = \frac{3}{8\pi G} m_g^2 \ .$$

Comparing this with (11) we find

$$\frac{\rho_{vac}}{\rho_{crit}} \simeq 10^{122} \Big(\frac{m_g}{m_{Pl}} \Big)^2 \ .$$

Even for $m_g \sim 1\ eV$ this ratio becomes 10^{66}. (m_g is related to the parameter F characterizing the strength of the supersymmetry breaking by $m_g = (4\pi G/3)^{1/2} F$, so $m_g \sim 1\ eV$ corresponds to $F^{1/2} \sim 100\ TeV$.)

Also string theory has not yet offered convincing clues why the cosmological constant is so extremely small. The main reason is that a *low energy mechanism* is required, and since supersymmetry is broken, one again expects a magnitude of order M_F^4, which is *at least 50 orders of magnitude too large* (see also [37]). However, non-supersymmetric physics in string theory is at the very beginning and workers in the field hope that further progress might eventually lead to an understanding of the cosmological constant problem.

I hope I have convinced the reader that we are indeed facing a profound mystery. (For other recent reviews, see also [38–41]. These contain more extended lists of references.)

5 Luminosity–Redshift Relation for Type Ia Supernovae

A few years ago the Hubble diagram for Type Ia supernovae gave, as a big surprise, the first serious evidence for a currently accelerating Universe. Before presenting and discussing critically these exciting results, we develop on the basis of Appendix A some theoretical background.

5.1 Theoretical Redshift–Luminosity Relation

In cosmology several different distance measures are in use, which are all related by simple redshift factors (see Sect. A.4). The one which is relevant in this section is the *luminosity distance* D_L. We recall that this is defined by

$$D_L = (\mathcal{L}/4\pi\mathcal{F})^{1/2} \ , \tag{18}$$

where $\mathcal{L}$ is the intrinsic luminosity of the source and $\mathcal{F}$ the observed energy flux.

We want to express this in terms of the redshift z of the source and some of the cosmological parameters. If the comoving radial coordinate r is chosen such that the Friedmann–Lemaître metric takes the form

$$g = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right], \quad k = 0, \pm 1 , \tag{19}$$

then we have

$$\mathcal{F} dt_0 = \mathcal{L} dt_e \cdot \frac{1}{1+z} \cdot \frac{1}{4\pi (r_e a(t_0))^2} .$$

The second factor on the right is due to the redshift of the photon energy; the indices $0, e$ refer to the present and emission times, respectively. Using also $1+z = a(t_0)/a(t_e)$, we find in a first step:

$$D_L(z) = a_0(1+z)r(z) \quad (a_0 \equiv a(t_0)) . \tag{20}$$

We need the function $r(z)$. From

$$dz = -\frac{a_0}{a}\frac{\dot{a}}{a}dt , \quad dt = -a(t)\frac{dr}{\sqrt{1-kr^2}}$$

for light rays, we see that

$$\frac{dr}{\sqrt{1-kr^2}} = \frac{1}{a_0}\frac{dz}{H(z)} \quad (H(z) = \frac{\dot{a}}{a}) . \tag{21}$$

Now, we make use of the Friedmann equation

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

Let us decompose the total energy–mass density ρ into non-relativistic (NR), relativistic (R), Λ, quintessence (Q), and possibly other contributions

$$\rho = \rho_{NR} + \rho_R + \rho_\Lambda + \rho_Q + \cdots . \tag{23}$$

For the relevant cosmic period we can assume that the "energy equation"

$$\frac{d}{da}(\rho a^3) = -3pa^2 \tag{24}$$

also holds for the individual components $X = NR, R, \Lambda, Q, \cdots$. If $w_X \equiv p_X/\rho_X$ is constant, this implies that

$$\rho_X a^{3(1+w_X)} = \text{const} . \tag{25}$$

Therefore,

$$\rho = \sum_X \left(\rho_X a^{3(1+w_X)}\right)_0 \frac{1}{a^{3(1+w_X)}} = \sum_X (\rho_X)_0 (1+z)^{3(1+w_X)} . \tag{26}$$

Hence the Friedmann equation (22) can be written as

$$\frac{H^2(z)}{H_0^2} + \frac{k}{H_0^2 a_0^2}(1+z)^2 = \sum_X \Omega_X (1+z)^{3(1+w_X)} , \tag{27}$$

where Ω_X is the dimensionless density parameter for the species X,

$$\Omega_X = \frac{(\rho_X)_0}{\rho_{crit}} , \tag{28}$$

where ρ_{crit} is the critical density:

$$\begin{aligned} \rho_{crit} &= \frac{3H_0^2}{8\pi G} \\ &= 1.88 \times 10^{-29} \; h_0^2 \; g \; cm^{-3} \\ &= 8 \times 10^{-47} h_0^2 \; GeV^{\,4} . \end{aligned} \tag{29}$$

Here h_0 denotes the *reduced Hubble parameter*

$$h_0 = H_0/(100 \; km \; s^{-1} \; Mpc^{-1}) \simeq 0.7 . \tag{30}$$

Using also the curvature parameter $\Omega_K \equiv -k/H_0^2 a_0^2$, we obtain the useful form

$$\boxed{H^2(z) = H_0^2 E^2(z; \Omega_K, \Omega_X) ,} \tag{31}$$

with

$$E^2(z; \Omega_K, \Omega_X) = \Omega_K (1+z)^2 + \sum_X \Omega_X (1+z)^{3(1+w_X)} . \tag{32}$$

Especially for $z = 0$ this gives

$$\Omega_K + \Omega_0 = 1, \quad \Omega_0 \equiv \sum_X \Omega_X . \tag{33}$$

If we use (31) in (21), we get

$$\int_0^{r(z)} \frac{dr}{\sqrt{1-kr^2}} = \frac{1}{H_0 a_0} \int_0^z \frac{dz'}{E(z')} \tag{34}$$

and thus

$$r(z) = \mathcal{S}(\chi(z)) , \tag{35}$$

where

$$\chi(z) = \frac{1}{H_0 a_0} \int_0^z \frac{dz'}{E(z')} \tag{36}$$

and

$$\mathcal{S}(\chi) = \begin{cases} \sin\chi & : \quad k = 1 \\ \chi & : \quad k = 0 \\ \sinh\chi & : \quad k = 1 . \end{cases} \tag{37}$$

Inserting this in (20) gives finally the relation we were looking for

$$D_L(z) = \frac{1}{H_0}\mathcal{D}_L(z; \Omega_K, \Omega_X) , \tag{38}$$

with

$$\mathcal{D}_L(z; \Omega_K, \Omega_X) = (1+z)\frac{1}{|\Omega_K|^{1/2}}\mathcal{S}\left(|\Omega_K|^{1/2}\int_0^z \frac{dz'}{E(z')}\right) \tag{39}$$

for $k = \pm 1$. For a flat universe, $\Omega_K = 0$ or equivalently $\Omega_0 = 1$, the "Hubble-constant-free" luminosity distance is

$$\mathcal{D}_L(z) = (1+z)\int_0^z \frac{dz'}{E(z')} . \tag{40}$$

Astronomers use as logarithmic measures of $\mathcal{L}$ and $\mathcal{F}$ the *absolute and apparent magnitudes*,[5] denoted by M and m, respectively. The conventions are chosen such that the *distance modulus* $m - M$ is related to D_L as follows

$$m - M = 5\log\left(\frac{D_L}{1\ Mpc}\right) + 25 . \tag{41}$$

Inserting the representation (38), we obtain the following relation between the apparent magnitude m and the redshift z:

$$m = \mathcal{M} + 5\log\mathcal{D}_L(z; \Omega_K, \Omega_X) , \tag{42}$$

where, for our purpose, $\mathcal{M} = M - 5\log H_0 + 25$ is an uninteresting fit parameter. The comparison of this theoretical *magnitude redshift relation* with data will lead to interesting restrictions for the cosmological Ω-parameters. In practice often only Ω_M and Ω_Λ are kept as independent parameters, where from now on the subscript M denotes (as in most papers) non-relativistic matter.

The following remark about *degeneracy curves* in the Ω-plane is important in this context. For a fixed z in the presently explored interval, the contours defined by the equations $\mathcal{D}_L(z; \Omega_M, \Omega_\Lambda) = const$ have little curvature, and thus we can associate an approximate slope to them. For $z = 0.4$ the slope is about 1 and increases to 1.5-2 by $z = 0.8$ over the interesting range of Ω_M and Ω_Λ. Hence even quite accurate data can at best select a strip in the Ω-plane, with a slope in the range just discussed. This is the reason behind the shape of the likelihood regions shown later (Fig. 2).

In this context it is also interesting to determine the dependence of the *deceleration parameter*

$$q_0 = -\left(\frac{a\ddot{a}}{\dot{a}^2}\right)_0 \tag{43}$$

[5] Beside the (bolometric) magnitudes m, M, astronomers also use magnitudes m_B, m_V, ... referring to certain wavelength bands B (blue), V (visual), and so on.

on Ω_M and Ω_Λ. At an any cosmic time we obtain from (107) and (26)

$$-\frac{\ddot{a}a}{\dot{a}^2} = \frac{1}{2}\frac{1}{E^2(z)}\sum_X \Omega_X(1+z)^{3(1+w_X)}(1+3w_X) \ . \tag{44}$$

For $z = 0$ this gives

$$q_0 = \frac{1}{2}\sum_X \Omega_X(1+3w_X) = \frac{1}{2}(\Omega_M - 2\Omega_\Lambda + \cdots) \ . \tag{45}$$

The line $q_0 = 0$ ($\Omega_\Lambda = \Omega_M/2$) separates decelerating from accelerating universes at the present time. For given values of Ω_M, Ω_Λ, etc., (44) vanishes for z determined by

$$\Omega_M(1+z)^3 - 2\Omega_\Lambda + \cdots = 0 \ . \tag{46}$$

This equation gives the redshift at which the deceleration period ends (coasting redshift).

Generalization for Dynamical Models of Dark Energy

If the vacuum energy constitutes the missing two-thirds of the average energy density of the *present* Universe, we would be confronted with the following *cosmic coincidence* problem: Since the vacuum energy density is constant in time – at least after the QCD phase transition – while the matter energy density decreases as the Universe expands, it would be more than surprising if the two are comparable just at about the present time, while their ratio was tiny in the early Universe and would become very large in the distant future. The goal of dynamical models of dark energy is to avoid such an extreme fine-tuning. The ratio p/ρ of this component then becomes a function of redshift, which we denote by $w_Q(z)$ (because the so-called "quintessence models" are particular examples). Then the function $E(z)$ in (32) gets modified.

To see how, we start from the energy equation (24) and write this as

$$\frac{d\ln(\rho_Q a^3)}{d\ln(1+z)} = 3w_Q \ .$$

This gives

$$\rho_Q(z) = \rho_{Q0}(1+z)^3 \exp\left(\int_0^{\ln(1+z)} 3w_Q(z')d\ln(1+z')\right)$$

or

$$\rho_Q(z) = \rho_{Q0} \exp\left(3\int_0^{\ln(1+z)} (1+w_Q(z'))d\ln(1+z')\right) \ . \tag{47}$$

Hence, we have to perform on the right of (32) the following substitution:

$$\Omega_Q(1+z)^{3(1+w_Q)} \to \Omega_Q \exp\left(3\int_0^{\ln(1+z)} (1+w_Q(z'))d\ln(1+z')\right) . \quad (48)$$

As indicated above, a much discussed class of dynamical models for dark energy are *quintessence models*. In many ways people thereby repeat what has been done in inflationary cosmology. The main motivation there was (see Appendix C) to avoid excessive fine tunings of standard big bang cosmology (horizon and flatness problems). In Appendix D we give a brief discussion of this class of models. It has to be emphasized, however, that quintessence models do *not* solve the vacuum energy problem, so far also not the coincidence puzzle.

5.2 Type Ia Supernovas as Standard Candles

It has long been recognized that supernovas of type Ia are excellent standard candles and are visible to cosmic distances [42] (the record is at present at a redshift of about 1.7). At relatively closed distances they can be used to measure the Hubble constant, by calibrating the absolute magnitude of nearby supernovas with various distance determinations (e.g., Cepheids). There is still some dispute over these calibration resulting in differences of about 10% for H_0. (For recent papers and references, see [43].)

In 1979, Tammann [44] and Colgate [45] independently suggested that at higher redshifts this subclass of supernovas can be used to determine also the deceleration parameter. In recent years this program became feasible, thanks to the development of new technologies which made it possible to obtain digital images of faint objects over sizable angular scales, and by making use of big telescopes such as Hubble and Keck.

There are two major teams investigating high-redshift SNe Ia, namely the "Supernova Cosmology Project" (SCP) and the "High-Z Supernova search Team" (HZT). Each team has found a large number of SNe, and both groups have published almost identical results. (For up-to-date information, see the home pages [46] and [47].)

Before discussing the most recent results, a few remarks about the nature and properties of type Ia SNe should be made. Observationally, they are characterized by the absence of hydrogen in their spectra, and the presence of some strong silicon lines near maximum. The immediate progenitors are most probably carbon–oxygen white dwarfs in close binary systems, but it must be said that these have not yet been clearly identified.[6]

In the standard scenario a white dwarf accretes matter from a non-degenerate companion until it approaches the critical Chandrasekhar mass

[6] This is perhaps not so astonishing, because the progenitors are presumably faint compact dwarf stars.

and ignites carbon burning deep in its interior of highly degenerate matter. This is followed by an outward-propagating nuclear flame leading to a total disruption of the white dwarf. Within a few seconds the star is converted largely into nickel and iron. The dispersed nickel radioactively decays to cobalt and then to iron in a few hundred days. A lot of effort has been invested to simulate these complicated processes. Clearly, the physics of thermonuclear runaway burning in degenerate matter is complex. In particular, since the thermonuclear combustion is highly turbulent, multidimensional simulations are required. This is an important subject of current research. (One gets a good impression of the present status from several articles in [48]. See also the review [49].) The theoretical uncertainties are such that, for instance, predictions for possible evolutionary changes are not reliable.

It is conceivable that in some cases a type Ia supernova is the result of a merging of two carbon–oxygen-rich white dwarfs with a combined mass surpassing the Chandrasekhar limit. Theoretical modeling indicates, however, that such a merging would lead to a collapse, rather than an SN Ia explosion. But this issue is still debated.

In view of the complex physics involved, it is not astonishing that type Ia supernovas are not perfect standard candles. Their peak absolute magnitudes have a dispersion of 0.3–0.5 mag, depending on the sample. Astronomers have, however, learned in recent years to reduce this dispersion by making use of empirical correlations between the absolute peak luminosity and light curve shapes. Examination of nearby SNe showed that the peak brightness is correlated with the time scale of their brightening and fading: slow decliners tend to be brighter than rapid ones. There are also some correlations with spectral properties. Using these correlations it became possible to reduce the remaining intrinsic dispersion, at least in the average, to $\simeq 0.15\ mag$. (For the various methods in use, and how they compare, see [50, 56], and references therein.) Other corrections, such as Galactic extinction, have been applied, resulting for each supernova in a corrected (rest-frame) magnitude. The redshift dependence of this quantity is compared with the theoretical expectation given by (41) and (39).

5.3 Results

After the classic papers [51–53] on the Hubble diagram for high-redshift type Ia supernovas, published by the SCP and HZT teams, significant progress has been made (for reviews, see [54] and [55]). I discuss first the main results presented in [56]. These are based on additional new data for $z > 1$, obtained in conjunction with the Great Observatories Origins Deep Survey (GOODS) Treasury program, conducted with the Advanced Camera for Surveys (ACS) aboard the Hubble Space Telescope (HST).

The quality of the data and some of the main results of the analysis are shown in Fig. 1. The data points in the top panel are the distance moduli relative to an empty uniformly expanding universe, $\Delta(m - M)$, and the redshifts

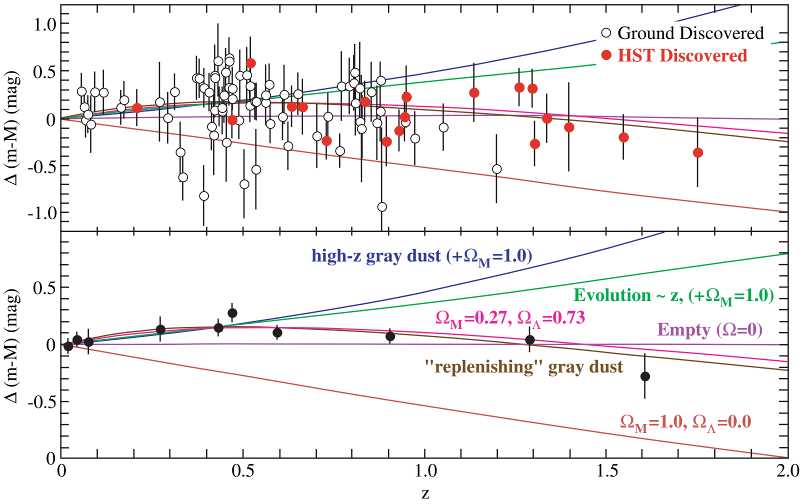

Fig. 1. Distance moduli relative to an empty uniformly expanding universe (residual Hubble diagram) for SNe Ia; see text for further explanations (Adapted from [56], Fig. 7.)

of a "gold" set of 157 SNe Ia. In this "reduced" Hubble diagram the filled symbols are the HST-discovered SNe Ia. The bottom panel shows weighted averages in fixed redshift bins.

These data are consistent with the "cosmic concordance" model ($\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$), with $\chi^2_{dof} = 1.06$. For a flat universe with a cosmological constant, the fit gives $\Omega_M = 0.29\pm^{0.13}_{0.19}$ (equivalently, $\Omega_\Lambda = 0.71$). The other model curves will be discussed below. Likelihood regions in the $(\Omega_M, \Omega_\Lambda)$-plane, keeping only these parameters in (39) and averaging H_0, are shown in Fig. 2. To demonstrate the progress, old results from 1998 are also included. It will turn out that this information is largely complementary to the restrictions we shall obtain from the CMB anisotropies.

In the meantime new results have been published. Perhaps the best high-z SN Ia compilation to date are the results from the Supernova Legacy Survey (SNLS) of the first year [57]. The other main research group has also published new data at about the same time [58].

5.4 Systematic Uncertainties

Possible systematic uncertainties due to astrophysical effects have been discussed extensively in the literature. The most serious ones are (i) *dimming* by intergalactic dust, and (ii) *evolution* of SNe Ia over cosmic time, due to changes in progenitor mass, metallicity, and C/O ratio. I discuss these concerns only briefly (see also [54, 56]).

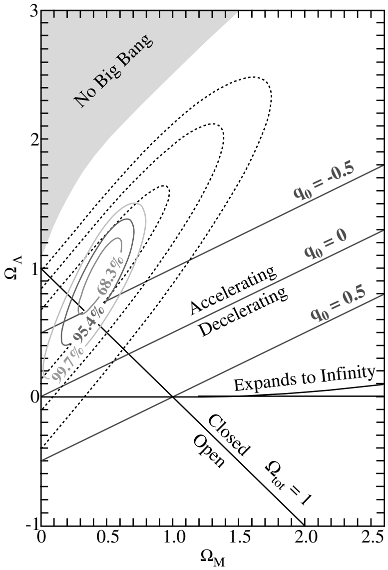

Fig. 2. Likelihood regions in the $(\Omega_M, \Omega_\Lambda)$-plane. The dotted contours are old results from 1998. (Adapted from [56], Fig. 8.)

Concerning extinction, detailed studies show that high-redshift SN Ia suffer little reddening; their B-V colors at maximum brightness are normal. However, it can a priori not be excluded that we see distant SNe through a grey dust with grain sizes large enough as to not imprint the reddening signature of typical interstellar extinction. One argument against this hypothesis is that this would also imply a larger dispersion than is observed. In Fig. 1 the expectation of a simple grey dust model is also shown. The new high-redshift data reject this monotonic model of astrophysical dimming. Equation (46) shows that at redshifts $z \geq (2\Omega_\Lambda/\Omega_M)^{1/3} - 1 \simeq 1.2$ the Universe is *decelerating*, and this provides an almost unambiguous signature for Λ, or some effective equivalent. There is now strong evidence for a transition from a deceleration to acceleration at a redshift $z = 0.46 \pm 0.13$.

The same data provide also some evidence against a simple luminosity evolution that could mimic an accelerating Universe. Other empirical constraints are obtained by comparing subsamples of low-redshift SN Ia believed to arise from old and young progenitors. It turns out that there is no difference within the measuring errors, *after* the correction based on the light-curve shape has been applied. Moreover, spectra of high-redshift SNe appear remarkably similar to those at low redshift. This is very reassuring. On the other hand, there seems to be a trend that more distant supernovas are bluer. It would, of course, be helpful if evolution could be predicted theoretically, but in view of what has been said earlier, this is not (yet) possible.

In conclusion, none of the investigated systematic errors appear to reconcile the data with $\Omega_\Lambda = 0$ and $q_0 \geq 0$. But further work is necessary before we can declare this as a really established fact.

To improve the observational situation a satellite mission called SNAP ("Supernovas Acceleration Probe") has been proposed [59]. According to the plans this satellite would observe about 2000 SNe within a year and much more detailed studies could then be performed. For the time being some scepticism with regard to the results that have been obtained is still not out of place, but the situation is steadily improving.

Finally, I mention a more theoretical complication. In the analysis of the data the luminosity distance for an ideal Friedmann universe was always used. But the data were taken in the real inhomogeneous Universe. This may perhaps not be good enough, especially for high-redshift standard candles. The simplest way to take this into account is to introduce a filling parameter which, roughly speaking, represents matter that exists in galaxies but not in the intergalactic medium. For a constant filling parameter one can determine the luminosity distance by solving the Dyer–Roeder equation. But now one has an additional parameter in fitting the data. For a flat universe this was investigated in [60]. We shall come back to this issue in Sect. 8.2.

6 Microwave Background Anisotropies

Investigations of the cosmic microwave background have presumably contributed most to the remarkable progress in cosmology during recent years (For a review, see [61]). Beside its spectrum, which is Planckian to an incredible degree, we also can study the temperature fluctuations over the "cosmic photosphere" at a redshift $z \approx 1100$. Through these we get access to crucial cosmological information (primordial density spectrum, cosmological parameters, etc.). A major reason for why this is possible relies on the fortunate circumstance that the fluctuations are tiny ($\sim 10^{-5}$) at the time of recombination. This allows us to treat the deviations from homogeneity and isotropy for an extended period of time perturbatively, i.e., by linearizing the Einstein and matter equations about solutions of the idealized Friedmann–Lemaître models. Since the physics is effectively *linear*, we can accurately work out the *evolution* of the perturbations during the early phases of the Universe, given a set of cosmological parameters. Confronting this with observations tells us a lot about the cosmological parameters as well as the initial conditions, and thus about the physics of the very early Universe. Through this window to the earliest phases of cosmic evolution we can, for instance, test general ideas and specific models of inflation.

6.1 Qualitative Remarks

Let me begin with some qualitative remarks, before I go into more technical details. Long before recombination (at temperatures $T > 6000$ K, say) photons, electrons, and baryons were so strongly coupled that these components may be treated together as a single fluid. In addition to this there is also a dark matter component. For all practical purposes the two interact only gravitationally. The investigation of such a two-component fluid for small deviations from an idealized Friedmann behavior is a well-studied application of cosmological perturbation theory (see, e.g., [63]).

At a later stage, when decoupling is approached, this approximate treatment breaks down because the mean free path of the photons becomes longer (and finally "infinite" after recombination). While the electrons and baryons can still be treated as a single fluid, the photons and their coupling to the electrons have to be described by the general relativistic Boltzmann equation. The latter is, of course, again linearized about the idealized Friedmann solution. Together with the linearized fluid equations (for baryons and cold dark matter, say) and the linearized Einstein equations one arrives at a complete system of equations for the various perturbation amplitudes of the metric and matter variables. There exist widely used codes, e.g. CMBFAST [62], that provide the CMB anisotropies – for given initial conditions – to a precision of about 1%. A lot of qualitative and semi-quantitative insight into the relevant physics can, however, be gained by looking at various approximations of the basic dynamical system.

Let us first discuss the temperature fluctuations. What is observed is the temperature autocorrelation:

$$C(\vartheta) := \left\langle \frac{\Delta T(\mathbf{n})}{T} \cdot \frac{\Delta T(\mathbf{n}')}{T} \right\rangle = \sum_{l=2}^{\infty} \frac{2l+1}{4\pi} C_l P_l(\cos\vartheta) \,, \tag{49}$$

where ϑ is the angle between the two directions of observation $\mathbf{n}, \mathbf{n}'$, and the average is taken ideally over all sky. The *angular power spectrum* is by definition $\frac{l(l+1)}{2\pi} C_l$ *versus* l $(\vartheta \simeq \pi/l)$.

A characteristic scale, which is reflected in the observed CMB anisotropies, is the sound horizon at last scattering, i.e., the distance over which a pressure wave can propagate until decoupling. This can be computed within the unperturbed model and subtends about half a degree on the sky for typical cosmological parameters. For scales larger than this sound horizon the fluctuations have been laid down in the very early Universe. These have been detected by the COBE satellite. The (gauge invariant brightness) temperature perturbation $\Theta = \Delta T/T$ is dominated by the combination of the intrinsic temperature fluctuations and gravitational redshift or blueshift effects. For example, photons that have to climb out of potential wells for high-density regions are redshifted. One can show that these effects combine for adiabatic initial conditions to $\frac{1}{3}\Psi$, where Ψ is one of the two gravitational Bardeen potentials. The latter, in turn, is directly related to the density perturbations. For scale-free initial perturbations and almost vanishing spatial curvature the corresponding angular power spectrum of the temperature fluctuations turns out to be nearly flat (Sachs–Wolfe plateau in Fig. 3).

On the other hand, inside the sound horizon before decoupling, acoustic, Doppler, gravitational redshift, and photon diffusion effects combine to the spectrum of small angle anisotropies shown in Fig. 3. These result from gravitationally driven synchronized acoustic oscillations of the photon–baryon fluid, which are damped by photon diffusion.

A particular realization of $\Theta(\mathbf{n})$, such as the one accessible to us (all sky map from our location), cannot be predicted. Theoretically, Θ is a random field $\Theta(\mathbf{x}, \eta, \mathbf{n})$, depending on the conformal time η, the spatial coordinates, and the observing direction $\mathbf{n}$. Its correlation functions should be rotationally invariant in $\mathbf{n}$, and respect the symmetries of the background time slices. If we expand Θ in terms of spherical harmonics,

$$\Theta(\mathbf{n}) = \sum_{lm} a_{lm} Y_{lm}(\mathbf{n}) \;, \tag{50}$$

the random variables a_{lm} have to satisfy

$$\langle a_{lm} \rangle = 0, \quad \langle a^\star_{lm} a_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l(\eta) \;, \tag{51}$$

where the $C_l(\eta)$ depend only on η. Hence the correlation function at the present time η_0 is given by (49), where $C_l = C_l(\eta_0)$, and the bracket now denotes the statistical average. Thus,

$$C_l = \frac{1}{2l+1} \left\langle \sum_{m=-l}^{l} a^\star_{lm} a_{lm} \right\rangle \;. \tag{52}$$

The standard deviations $\sigma(C_l)$ measure a fundamental uncertainty in the knowledge we can get about the C_l's. These are called *cosmic variances*, and are most pronounced for low l. In simple inflationary models the a_{lm} are Gaussian distributed, hence

$$\frac{\sigma(C_l)}{C_l} = \sqrt{\frac{2}{2l+1}} \;. \tag{53}$$

Therefore, the limitation imposed on us (only one sky in one universe) is small for large l.

6.2 Boltzmann Hierarchy

The brightness temperature fluctuation can be obtained from the perturbation of the photon distribution function by integrating over the magnitude of the photon momenta. The linearized Botzmann equation can then be translated into an equation for Θ, which we now regard as a function of η, x^i, and γ^j, where the γ^j are the directional cosines of the momentum vector relative to an orthonormal triad field of the unperturbed spatial metric with curvature

K. Next one performs a harmonic decomposition of Θ, which reads for the spatially flat case ($K = 0$)

$$\Theta(\eta, \boldsymbol{x}, \boldsymbol{\gamma}) = (2\pi)^{-3/2} \int d^3k \sum_l \theta_l(\eta, k) G_l(\boldsymbol{x}, \boldsymbol{\gamma}; \boldsymbol{k}) \,, \tag{54}$$

where

$$G_l(\boldsymbol{x}, \boldsymbol{\gamma}; \boldsymbol{k}) = (-i)^l P_l(\hat{\boldsymbol{k}} \cdot \boldsymbol{\gamma}) \exp(i\boldsymbol{k} \cdot \boldsymbol{x}) \,. \tag{55}$$

The dynamical variables $\theta_l(\eta)$ are the *brightness moments*, and should be regarded as random variables. Boltzmann's equation implies the following hierarchy of ordinary differential equations for the brightness moments[7] $\theta_l(\eta)$ (if polarization effects are neglected):

$$\theta_0' = -\frac{1}{3} k\theta_1 - \Phi' \,, \tag{56}$$

$$\theta_1' = k\Big(\theta_0 + \Psi - \frac{2}{5}\theta_2\Big) - \dot{\tau}(\theta_1 - V_b) \,, \tag{57}$$

$$\theta_2' = k\Big(\frac{2}{3}\theta_1 - \frac{3}{7}\theta_3\Big) - \dot{\tau}\frac{9}{10}\theta_2 \,, \tag{58}$$

$$\theta_l' = k\Big(\frac{l}{2l-1}\theta_{l-1} - \frac{l+1}{2l+3}\theta_{l+1}\Big) \,, \quad l > 2 \,. \tag{59}$$

Here, V_b is the gauge invariant scalar velocity perturbation of the baryons, $\dot{\tau} = x_e n_e \sigma_T a/a_0$, where a is the scale factor, $x_e n_e$ the unperturbed free electron density (x_e = ionization fraction), and σ_T the Thomson cross section. Moreover, Φ and Ψ denote the Bardeen potentials. (For further details, see, e.g., Sect. 6 of [3] or [63], where cosmological perturbation theory is developed in great detail.)

The C_l are determined by an integral over k, involving a primordial power spectrum (of curvature perturbations) and the $|\theta_l(\eta)|^2$, for the corresponding initial conditions (their transfer functions).

This system of equations is completed by the linearized fluid and Einstein equations. Various approximations for the Boltzmann hierarchy provide already a lot of insight. In particular, one can very nicely understand how damped acoustic oscillations are generated, and in which way they are influenced by the baryon fraction (again, see [3] or [63]). A typical theoretical CMB spectrum is shown in Fig. 3. (Beside the scalar contribution in the sense of cosmological perturbation theory, considered so far, the tensor contribution due to gravity waves is also shown there.)

6.3 Polarization

A polarization map of the CMB radiation provides important additional information to that obtainable from the temperature anisotropies. For example,

[7] In the literature the normalization of the θ_l is sometimes chosen differently: $\theta_l \to (2l+1)\theta_l$.

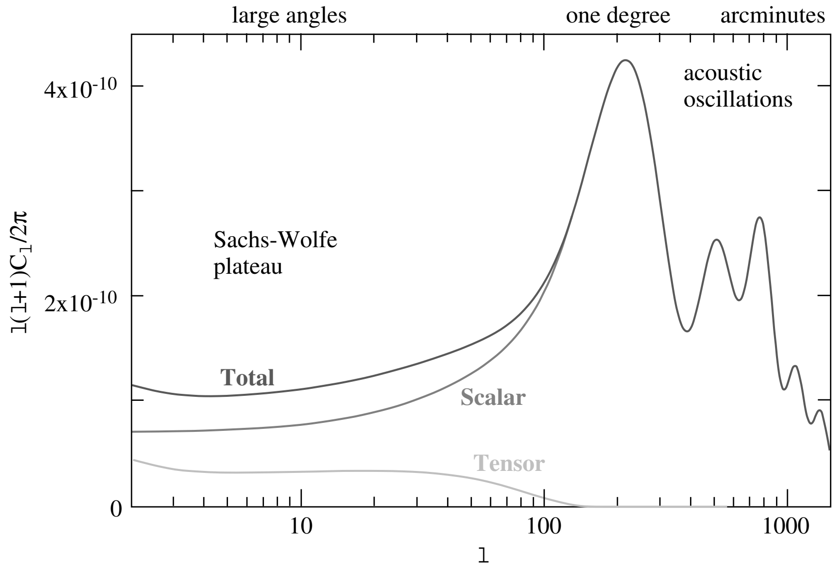

Fig. 3. Theoretical angular temperature–temperature (TT) power spectrum for adiabatic initial perturbations and typical cosmological parameters. The scalar and tensor contributions to the anisotropies are also shown

we can get constraints about the epoch of reionization. Most importantly, future polarization observations may reveal a stochastic background of gravity waves, generated in the very early Universe. In this section we give a brief introduction to the study of CMB polarization.

The mechanism which partially polarizes the CMB radiation is similar to that for the scattered light from the sky. Consider first scattering at a single electron of unpolarized radiation coming in from all directions. Due to the familiar polarization dependence of the differential Thomson cross section, the scattered radiation is, in general, polarized. It is easy to compute the corresponding Stokes parameters. Not surprisingly, they are not all equal to zero if and only if the intensity distribution of the incoming radiation has a non-vanishing quadrupole moment. The Stokes parameters Q and U are proportional to the overlap integral with the combinations $Y_{2,2} \pm Y_{2,-2}$ of the spherical harmonics, while V vanishes. This is basically the reason why a CMB polarization map traces (in the tight coupling limit) the quadrupole temperature distribution on the last scattering surface.

The polarization tensor of an all sky map of the CMB radiation can be parametrized in temperature fluctuation units, relative to the orthonormal basis $\{d\vartheta, \sin\vartheta \;\, d\varphi\}$ of the two sphere, in terms of the Pauli matrices as $\Theta \cdot 1 + Q\sigma_3 + U\sigma_1 + V\sigma_2$. The Stokes parameter V vanishes (no circular polarization). Therefore, the polarization properties can be described by the following symmetric trace-free tensor on S^2:

$$(\mathcal{P}_{ab}) = \begin{pmatrix} Q & U \\ U & -Q \end{pmatrix} . \tag{60}$$

As for gravity waves, the components Q and U transform under a rotation of the 2-bein by an angle α as

$$Q \pm iU \to e^{\pm 2i\alpha}(Q \pm iU) , \tag{61}$$

and are thus of spin-weight 2. $\mathcal{P}_{ab}$ can be decomposed uniquely into *electric* and *magnetic* parts:

$$\mathcal{P}_{ab} = E_{;ab} - \frac{1}{2} g_{ab} \Delta E + \frac{1}{2}(\varepsilon_a{}^c B_{;bc} + \varepsilon_b{}^c B_{;ac}) . \tag{62}$$

Expanding here the scalar functions E and B in terms of spherical harmonics, we obtain an expansion of the form

$$\mathcal{P}_{ab} = \sum_{l=2}^{\infty} \sum_m \left[a^E_{(lm)} Y^E_{(lm)ab} + a^B_{(lm)} Y^B_{(lm)ab} \right] \tag{63}$$

in terms of the tensor harmonics:

$$Y^E_{(lm)ab} := N_l (Y_{(lm);ab} - \frac{1}{2} g_{ab} Y_{(lm);c}{}^c), \quad Y^B_{(lm)ab} := \frac{1}{2} N_l (Y_{(lm);ac} \varepsilon^c{}_b + a \leftrightarrow b) , \tag{64}$$

where $l \geq 2$ and

$$N_l \equiv \left(\frac{2(l-2)!}{(l+2)!} \right)^{1/2} .$$

Equivalently, one can write this as

$$Q + iU = \sqrt{2} \sum_{l=2}^{\infty} \sum_m \left[a^E_{(lm)} + i a^B_{(lm)} \right] {}_2Y_l^m , \tag{65}$$

where ${}_sY_l^m$ are the spin-s harmonics.

As in (50) the multipole moments $a^E_{(lm)}$ and $a^B_{(lm)}$ are random variables, and we have equations analogous to (52):

$$C_l^{TE} = \frac{1}{2l+1} \sum_m \langle a^{\Theta\star}_{lm} a^E_{lm} \rangle, \quad \text{etc} . \tag{66}$$

(We have now put the superscript Θ on the a_{lm} of the temperature fluctuations.) The C_l's determine the various angular correlation functions. For example, one easily finds

$$\langle \Theta(\boldsymbol{n}) Q(\boldsymbol{n'}) \rangle = \sum_l C_l^{TE} \frac{2l+1}{4\pi} N_l P_l^2(\cos\vartheta) . \tag{67}$$

For the spacetime-dependent Stokes parameters Q and U of the radiation field we can perform a normal mode decomposition analogous to (54). If, for simplicity, we again consider only scalar perturbations this reads

$$Q \pm iU = (2\pi)^{-3/2} \int d^3k \sum_l (E_l \pm iB_l) \, {}_{\pm 2}G_l^0 \;, \tag{68}$$

where

$${}_sG_l^m(\boldsymbol{x}, \boldsymbol{\gamma}; \boldsymbol{k}) = (-i)^l \left(\frac{2l+1}{4\pi}\right)^{1/2} {}_sY_l^m(\boldsymbol{\gamma}) \exp(i\boldsymbol{k} \cdot \boldsymbol{x}) \;, \tag{69}$$

if the mode vector $\mathbf{k}$ is chosen as the polar axis. (Note that G_l in (55) is equal to ${}_0G_l^0$.)

The Boltzmann equation implies a coupled hierarchy for the moments θ_l, E_l, and B_l [64, 65]. It turns out that the B_l vanish for scalar perturbations. Non-vanishing magnetic multipoles would be a unique signature for a spectrum of gravity waves. In a sudden decoupling approximation, the present electric multipole moments can be expressed in terms of the brightness quadrupole moment on the last scattering surface and spherical Bessel functions as

$$\frac{E_l(\eta_0, k)}{2l+1} \simeq \frac{3}{8} \theta_2(\eta_{dec}, k) \frac{l^2 j_l(k\eta_0)}{(k\eta_o)^2} \;. \tag{70}$$

Here one sees how the observable E_l's trace the quadrupole temperature anisotropy on the last scattering surface. In the tight coupling approximation the latter is proportional to the dipole moment θ_1.

7 Observational Results and Cosmological Parameters

In recent years several experiments gave clear evidence for multiple peaks in the angular temperature power spectrum at positions expected on the basis of the simplest inflationary models and big bang nucleosynthesis [66]. These results have been confirmed and substantially improved by the first-year WMAP data [67, 68, 72]. Fortunately, the improved data after three years of integration are now available [69]. Below we give a brief summary of some of the most important results.

Figure 4 shows the 3-year data of WMAP for the TT angular power spectrum, and the best fit (power law) ΛCDM model. The latter is a spatially flat model and involves the following six parameters: $\Omega_b h_0^2$, $\Omega_M h_0^2$, H_0, amplitude of fluctuations,[8] σ_8, optical depth τ, and the spectral index, n_s, of the primordial scalar power spectrum (see Appendix C.7). Figure 5 shows in addition the TE polarization data [70]. There are now also EE data that lead to a further reduction of the allowed parameter space. The first column in Table 1 shows the best fit values of the six parameters, using only the WMAP data.

[8] σ_8^2 is the variance of mass fluctuations in spheres of radius 8 h_0^{-1} Mpc. (For a precise definition, see, e.g., Appendix A of [63].)

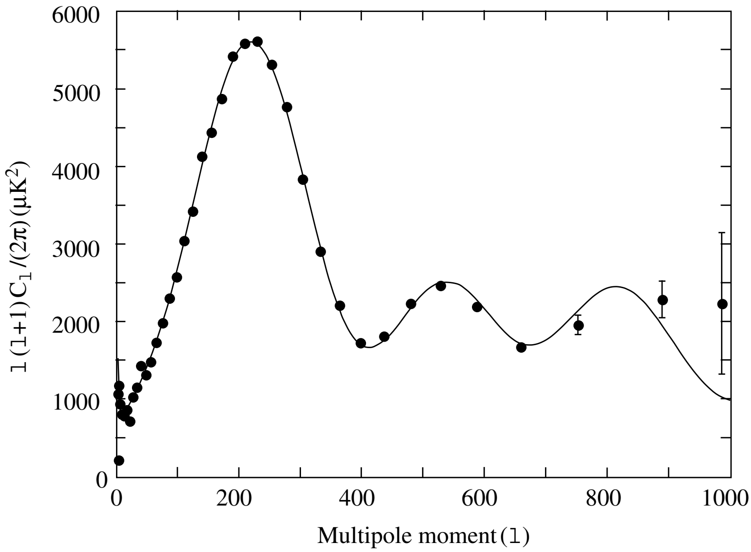

Fig. 4. Three-year WMAP data for the temperature–temperature (TT) power spectrum. The black line is the best fit ΛCDM model for the 3-year WMAP data. (Adapted from Fig. 2 of [69])

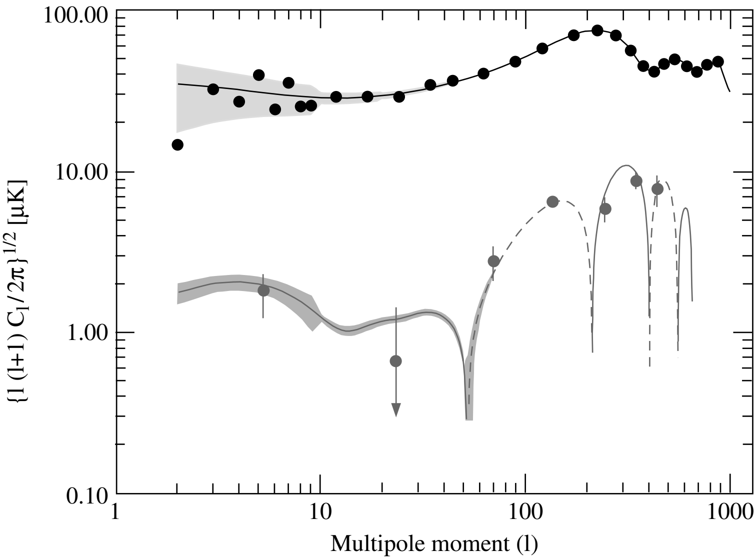

Fig. 5. WMAP data for the temperature-polarization TE power from Fig. 25 of [70])

Table 1.

Parameter	WMAP alone	WMAP + 2dFGRS
$100\Omega_b h_0^2$	$2.233^{+0.072}_{-0.0.091}$	$2.223^{+0.066}_{-0.083}$
$\Omega_M h_0^2$	$0.1268^{+0.0073}_{-0.0128}$	$0.1262^{+0.0050}_{-0.0103}$
h_0	$0.734^{+0.028}_{-0.038}$	$0.732^{+0.018}_{-0.025}$
Ω_M	$0.238^{+0.027}_{-0.045}$	$0.236^{+0.016}_{-0.029}$
σ_8	$0.744^{+0.050}_{-0.060}$	$0.737^{+0.033}_{-0.045}$
τ	$0.088^{+0.028}_{-0.034}$	$0.083^{+0.027}_{-0.031}$
n_s	$0.951^{+0.015}_{-0.019}$	$0.948^{+0.014}_{-0.018}$

Figure 6 shows the prediction of the model for the luminosity-redshift relation, together with the SLNS data [57] mentioned in Sect. 5.3. For other predictions and corresponding data sets, see [69].

Combining the WMAP results with other astronomical data reduces the uncertainties for some of the six parameters. This is illustrated in the second column which shows the 68% confidence ranges of a joint likelihood analysis when the power spectrum from the completed 2dFGRS [73] is added. In [69] other joint constraints are listed (see their Tables 5, 6). In Fig. 7 we reproduce one of many plots in [69] that shows the joint marginalized contours in the (Ω_M, h_0)-plane.

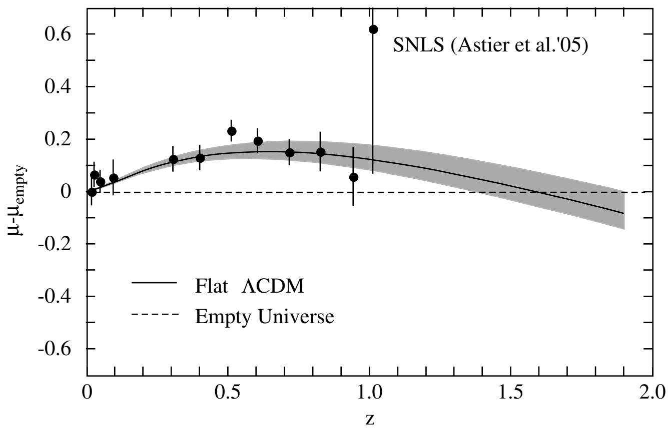

Fig. 6. Prediction for the luminosity-redshift relation from the ΛCDM model model fit to the WMAP data only. The ordinate is the deviation of the distance modulous from the empty universe model. The prediction is compared to the SNLS data [57]. (From Fig. 8 of [69])

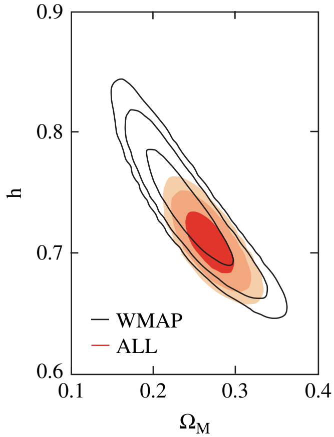

Fig. 7. Joint marginalized contours (68% and 95% confidence levels) in the (Ω_M, h_0)-plane for WMAP only (*solid lines*) and additional data (*filled red*) for the power-law ΛCDM model. (From Fig. 10 in [69])

The parameter space of the cosmological model can be extended in various ways. Because of intrinsic degeneracies, the CMB data alone are no more sufficient to determine unambiguously the cosmological model parameters. We illustrate this for non-flat models. For these the WMAP data (in particular, the position of the first acoustic peak) restricts the curvature parameter Ω_K to a narrow region around the degeneracy line $\Omega_K = -0.3040 + 0.4067$, $\Omega_\Lambda = 0.758^{+0.035}_{-0.058}$. This does not exclude models with $\Omega_\Lambda = 0$. However, when, for instance, the Hubble constant is restricted to an acceptable range, the universe must be nearly flat. For example, the restriction $h_0 = 0.72 \pm 0.08$ implies that $\Omega_K = -0.003^{+0.013}_{-0.017}$. Other strong limits are given in Table 11 of [69], assuming that $w = -1$. But even when this is relaxed, the combined data constrain Ω_K and w significantly (see Fig. 17 of [69]). The marginalized best fit values are $w = -1.062^{+0.128}_{-0.079}$, $\Omega_K = -0.024^{+0.016}_{-0.013}$ at the 68% confidence level.

The restrictions on w – assumed to have no z-dependence – for a flat model are illustrated in Fig. 8.

Another interesting result is that reionization of the Universe has set in at a redshift of $z_r = 10.9^{+2.7}_{-2.3}$. Later we shall add some remarks on what has been learnt about the primordial power spectrum.

It is most remarkable that a six parameter cosmological model is able to fit such a rich body of astronomical observations. There seems to be little room for significant modifications of the successful ΛCDM model. In spite of this we discuss in the next section some proposed attempts to explain the observations without dark energy.

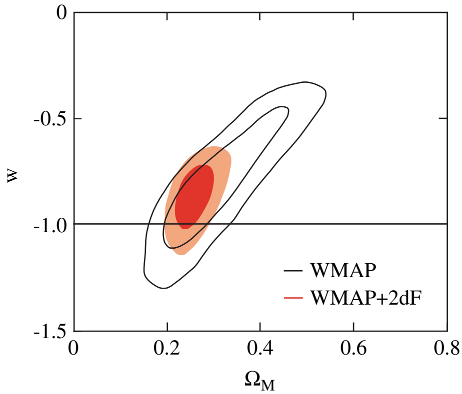

Fig. 8. Constraints on the equation of state parameter w in a flat universe model when WMAP data are combined with the 2dFGRS data. (From Fig. 15 in [69])

8 Alternatives to Dark Energy

In the previous two sections we have discussed some of the wide range of astronomical data that support the following 'concordance model': The Universe is spatially flat and dominated by a dark energy component and weakly interacting cold dark matter. Furthermore, the primordial fluctuations are adiabatic, nearly scale invariant and Gaussian, as predicted in simple inflationary models (see Sect. C.7). It is very likely that the present concordance model will survive phenomenologically.

A dominant dark energy component with density parameter $\simeq 0.7$ is so surprising that it should be examined whether this conclusion is really unavoidable. In what follows I shall briefly discuss some alternatives that have been proposed.

8.1 Changes in the Initial Conditions

Since we do not have a tested theory predicting the spectrum of primordial fluctuations, it appears reasonable to consider a wider range of possibilities than simple power laws. An instructive attempt in this direction was made some time ago [74], by constructing an Einstein–de Sitter model with $\Omega_\Lambda = 0$, fitting the CMB data as well as the power spectrum of 2dFGRS. In this the Hubble constant is, however, required to be rather low: $H_0 \simeq 46$ km/s/Mpc. The authors argued that this cannot definitely be excluded, because 'physical' methods lead mostly to relatively low values of H_0. In order to be consistent with matter fluctuations on cluster scales they added relic neutrinos with degenerate masses of order eV or a small contribution of quintessence with zero pressure ($w = 0$). In addition, they ignored the direct evidence for an accelerating Universe from the Hubble-diagram for distant Type Ia supernovae, on the basis of remaining systematic uncertainties. In the meantime, significant improvements in astronomical data sets have been made. In particular, the analysis of the 3-year WMAP data showed that there are no significant features in the primordial curvature fluctuation spectrum (see Sect. 5 of [69]). With the larger samples of high redshift supernovae and more precise information on large-scale galaxy clustering, such models with vanishing dark energy are no more possible [75].

8.2 Inhomogeneous Models

Backreaction

It has recently been suggested [76, 77] that perturbations on scales larger than the Hubble length, likely generated in the context of inflation, could mimic dark energy and cause acceleration. This suggestion caused a lot of discussion, and several papers addressed the question whether this is really possible. We repeat below a simple general argument given in [78] that the originally proposed mechanism cannot lead to acceleration, under the assumptions made in the cited papers. These include that the 4-velocity field u^μ of the CDM particles is geodesic and has zero vorticity $\omega_{\mu\nu}$. It is easy to see that these assumptions imply that the 1-form $\mathbf{u}$, belonging to the velocity field, has a vanishing exterior derivative. Hence we have locally $\mathbf{u} = dt$, thus u^μ is perpendicular to the slices $\{t = const\}$. Moreover the metric and the velocity have the form

$$g = -dt^2 + \bar{g}_t, \quad u = \partial_t \ , \tag{71}$$

where $\bar{g}_t$ is a t-dependent metric on slices of constant time t.

For such an inhomogeneous cosmological model one can introduce various definitions of the deceleration parameter which reduce to the familiar one for Friedmann models. We adopt here the one used in [77]. To motivate this, consider for some initial time t_{in} a spatial domain D and let this evolve according to the flow of u. If ω_t denotes the volume form belonging to $\bar{g}_t$, then we have for the volume $|D_t|$ and its time derivatives

$$|D_t| = \int \omega_t, \quad |\dot{D}_t| = \int \theta\omega_t, \quad |\ddot{D}_t| = \int (\dot{\theta} + \theta^2)\omega_t \ , \tag{72}$$

where $\theta = \nabla \cdot u$ denotes the expansion. If $l := |D_t|^{1/3}$, a natural definition of the deceleration parameter is $q = -(l\ddot{l})/\dot{l}^2$. This can be expressed as follows

$$\frac{1}{3}\frac{(|\dot{D}_t|)^2}{|D_t|^2} q = -\left(\frac{|\ddot{D}_t|}{|D_t|} - \frac{2}{3}\frac{(|\dot{D}_t|)^2}{|D_t|^2}\right) . \tag{73}$$

For an *infinitesimal* $|D_t|$ we obtain from the previous equations

$$\frac{1}{3}\theta^2 q = -(\dot{\theta} + \frac{1}{3}\theta^2) \ . \tag{74}$$

For the right-hand side we can now use the *Raychaudhuri equation*

$$\dot{\theta} + \frac{1}{3}\theta^2 = -\sigma_{\mu\nu}\sigma^{\mu\nu} + \omega_{\mu\nu}\omega^{\mu\nu} - R_{\mu\nu}u^\mu u^\nu \ , \tag{75}$$

where $\sigma_{\mu\nu}$ is the shear. For a vanishing vorticity, and imposing the strong energy condition (assumed in [77]), we see that $q \geq 0$. In this sense there is no acceleration.

A priori, a way out proposed in [76], is to argue that q as defined above is not what is measured in SN Ia observations. To analyze these one has to generalize the redshift–luminosity distance relation to inhomogeneous models. In doing this, two possible definitions for the deceleration parameter arise. One of them (q_4 in [78]) again has to be non-negative if the strong energy condition holds. The other (q_3 in [78]) may be negative, but in this case the supernova data would have to show acceleration in certain directions and deceleration in others. This is, however, not observed.

Kolb et al. have reacted to these considerations [79]. They admit that super-Hubble modes cannot lead to an acceleration, but they maintain that sub-Hubble modes may cause a large backreaction that may imply an effective acceleration. The authors stress that for investigating the effective dynamics averaging over a volume of size comparable with the present-day Hubble volume is essential. Let me add a few remarks on this. Adopting the notation

$$\langle\theta\rangle = \frac{\int \theta\omega_t}{\int \omega_t}, \quad \text{etc} \ , \tag{76}$$

and using the Raychaudhuri equation, we can write

$$\begin{aligned} \frac{1}{3}\frac{(|D_t|^{\cdot})^2}{|D_t|^2}q &= -\langle\dot{\theta} + \theta^2\rangle + \frac{2}{3}\langle\theta\rangle^2 \\ &= -\langle\dot{\theta} + \frac{1}{3}\theta^2\rangle - \frac{2}{3}(\langle\theta^2\rangle - \langle\theta\rangle^2) \\ &= \langle\sigma_{\mu\nu}\sigma^{\mu\nu} + R_{\mu\nu}u^\mu u^\nu\rangle - \frac{2}{3}(\langle\theta^2\rangle - \langle\theta\rangle^2) \ . \end{aligned} \tag{77}$$

The first term in the last equation, is non-negative if the strong energy condition holds, while the second term is non-positive.

The authors of [79] suggest that the second term may win and make q negative. To decide on the basis of detailed calculations whether this is indeed possible is a very difficult task. From what we know about the CMB radiation it appears, however, unlikely that there are such sizable perturbations out to very large scales. We shall say more about this in the next section.

The work by Kolb et al. triggered a lot of activity. (For a review, see [80].) We add some remarks about the ongoing discussion.

Power Spectrum of the Luminosity Distance

The deceleration parameter, defined in (73), has a simple geometrical meaning, but is not a directly measurable quantity. From an observational point of view, a more satisfactory approach is to generalize the magnitude–redshift relation, and study the fluctuations of the luminosity distance.

The magnitude–redshift relation in a perturbed Friedmann model has been derived in [81], and was later used to determine the angular power spectrum of the luminosity distance (the C_l's defined in analogy to (49)) [82]. One of the numerical results was that the uncertainties in determining cosmological parameters via the magnitude–redshift relation caused by fluctuations are small compared with the intrinsic dispersion in the absolute magnitude of Type Ia supernovae.

This subject was recently taken up in [83], as part of a program to develop the tools for extracting cosmological parameters, when much extended supernovae data become available.

Exact Inhomogeneous Model Studies

Effects of inhomogeneous matter distribution on light propagation were recently studied in the Lemaître–Tolman (LT) model, in order to see whether these can mimic an accelerated expansion.

The LT model is a family of spherically symmetric dust solutions of Einstein's equations, with a metric of the form

$$g = -dt^2 + \frac{R_{,r}^2(r,t)}{1+2E(r)}dr^2 + R^2(r,t)(d\vartheta^2 + \sin^2\vartheta d\varphi^2) \ . \qquad (78)$$

The metric functions $E(r)$, $R(r,t)$, and a matter function $M(r)$ satisfy, as a consequence of Einstein's equations, the differential equations

$$M_{,r} = 4\pi\rho R^2 R_{,r}, \quad R_{,t}^2 = 2E + \frac{2GM}{R} + \frac{1}{3}\Lambda R^2 \ . \qquad (79)$$

For these models the magnitude–redshift relation can be worked out exactly.

As an example we mention [84], where it was shown that for $\Lambda = 0$ the observed behavior of supernovae brightness cannot be fitted, unless our position in the model universe is very special. In that case one has to analyze also other data, in particular the CMB angular power spectrum. At the time of writing, this has not yet been done, but is certainly underway.

8.3 Modifications of Gravity

Since no satisfactory explanation of dark energy has emerged so far, possible modifications of GR, which would change the late expansion rate of the universe, have recently come into the focus of attention. The cosmic speed-up might, for instance, be explained by sub-dominant terms (like $1/R$) that become essential at small curvature. Modified gravity models have to be devised such that to pass the stringent Solar System tests, and are compatible with the observational data that support the concordance model.

Generalizations of the Einstein–Hilbert Action

The simplest generalization consists in replacing the Ricci scalar, R, in the Einstein–Hilbert action by a function $f(R)$. Note that this gives rise to fourth-order field equations.[9] Applying a suitable conformal transformation of the metric, the action becomes equivalent to a scalar-tensor theory. In detail, if we define a new metric $\tilde{g}_{\mu\nu} = \left[\exp(2\kappa/3)^{1/2\varphi}\right] g_{\mu\nu}$, then the action becomes

$$S = \int \left[\frac{1}{2\kappa}R[\tilde{g}] - \frac{1}{2}\tilde{g}^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - V(\varphi) + L_{matter}\right]\sqrt{-\tilde{g}}d^4x \; , \tag{80}$$

where the potential V is determined by the function f. With this formulation one can, for instance, show that an arbitrary evolution of the scale factor $a(t)$ can be obtained with an appropriate choice of $f(R)$. It is also useful to check whether a particular model passes Solar System tests (acceptable Brans-Dicke parameter). One should, however, bear in mind that the two mathematically equivalent descriptions lead to physically different properties, for instance with regard to stability. These issues and the application for specific functions f to Friedmann spacetimes have recently been reviewed in [85].

A class of models that lead to cosmic acceleration is of the form $f(R) = R + \alpha/R^n, \; n > 0$. There has been a debate on whether such models (especially for $n = 1$) are consistent with Solar System tests. Some authors argued that this is the case, because they admit as a static spherically symmetric solution the Schwarzschild–de Sitter metric. This is, however, by no means sufficient. As already emphasized, this vacuum solution is far from unique. The correct one must match onto a physically acceptable solution for the interior of the star. In [86] it was shown for $n = 1$, i.e., for $f(R) = R - \mu^4/R$, that this requirement implies for the PPN parameter γ the value 1/2, in gross violation of the measured value $\gamma = 1 + (2.1 \pm 2.3) \times 10^{-5}$. This confirms an earlier claim by Chiba [87] that was based on the scalar-tensor reformulation (80).

Presumably, similar statements can be made for a large class of $f(R)$ models. Apart from their ad hoc nature, it has not yet been demonstrated that there are examples which satisfy all the constraints stressed above. The same can be said on generalizations [88], which include other curvature invariants, such as $R_{\mu\nu}R^{\mu\nu}, \; R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$. In addition, such models are in most cases *unstable*, like mechanical Lagrangian systems with higher derivatives [89].[10] An exception seem to be Lagrangians which are functions of

---

[9] Spherically symmetric vacuum solutions are, therefore, far from unique. Connected with this is that Birkhoff's theorem fails. So, on the basis of the vacuum equations the perihelion motion (for example) is no more predicted, but at best compatible with the theory. This is an enormous loss. (The reader may reflect about other drawbacks.)

[10] This paper contains a discussion of a generic instability of Lagrangian systems in mechanics with higher derivatives, which was discovered by M. Ostrogradski in 1850.

R and the Gauss–Bonnet invariant $G = R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$. By introducing two scalar fields such models can be written as an Einstein–Hilbert term plus a particular extra piece, containing a linear coupling to G. Because the Gauss–Bonnet invariant is a total divergence the corresponding field equations are of second order. This does, however, not guarantee that the theory is ghost-free. In [90] this question was studied for a class of models [88] for which there exist accelerating late-time power-law attractors and which satisfy the solar system constraints. It turned out that in a Friedmann background there are no ghosts, but there is instead *superluminal propagation* for a wide range of parameter space. This acausality is reminiscent of the Velo-Zwanziger phenomenon [92] for higher (> 1) spin fields coupled to external fields. It may very well be that it can only be avoided if very special conditions are satisfied. This issue deserves further investigations. See also [91].

First-Order Modifications of GR

The disadvantage of complicated fourth-order equations can be avoided by using the *Palatini variational principle*, in which the metric and the symmetric affine connection (the Christoffel symbols $\Gamma^{\alpha}{}_{\mu\nu}$) are considered to be independent fields.[11]

For GR the 'Palatini formulation' is equivalent to the Einstein–Hilbert variational principle, because the variational equation with respect to $\Gamma^{\alpha}{}_{\mu\nu}$ implies that the affine connection has to be the Levi–Civita connection. Things are no more that simple for $f(R)$ models:

$$S = \int \left[\frac{1}{2\kappa} f(R) + L_{matter} \right] \sqrt{-g} d^4 x \ , \tag{81}$$

where $R[g,\Gamma] = g^{\alpha\beta} R_{\alpha\beta}[\Gamma]$, $R_{\alpha\beta}[\Gamma]$ being the Ricci tensor of the independent torsionless connection Γ. The equations of motion are in obvious notation

$$f'(R) R_{(\mu\nu)}[\Gamma] - \frac{1}{2} f(R) g_{\mu\nu} = \kappa T_{\mu\nu} \ , \tag{82}$$

$$\nabla^{\Gamma}_{\alpha} \left(\sqrt{-g} f'(R) g^{\mu\nu} \right) = 0 \ . \tag{83}$$

For the second of these equations one has to assume that L_{matter} is functionally independent of Γ. (It may, however, contain metric covariant derivatives.)

Equation (83) implies that

$$\nabla^{\Gamma}_{\alpha} \left[\sqrt{-\hat{g}} \hat{g}^{\mu\nu} \right] = 0 \tag{84}$$

for the conformally equivalent metric $\hat{g}_{\mu\nu} = f'(R) g_{\mu\nu}$. Hence, the $\Gamma^{\alpha}{}_{\mu\nu}$ are equal to the Christoffel symbols for the metric $\hat{g}_{\mu\nu}$.

[11] This approach was actually first introduced by Einstein (S.B. Preuss. Akad. Wiss., 414 (1925)). This is correctly stated in Pauli's classical text, p. 215.

The trace of (82) gives

$$Rf'(R) - 2f(R) = \kappa T \ .$$

Thanks to this algebraic equation we may regard R as a function of T. In the matter-free case it is identically satisfied if $f(R)$ is proportional to R^2. In all other cases R is equal to a constant c (which is in general not unique). If $f'(c) \neq 0$, (83) implies that Γ is the Levi–Civita connection of $g_{\mu\nu}$, and (82) reduces to Einstein's vacuum equation with a cosmological constant. In general, one can rewrite the field equations in the form of Einstein gravity with non-standard matter couplings.[12] Because of this it is, for instance, straightforward to develop cosmological perturbation theory [94].

Koivisto [95] has applied this to study the resulting matter power spectrum, and showed that the comparison with observations leads to strong constraints. The allowed parameter space for a model of the form $f(R) = R - \alpha R^\beta$ ($\alpha > 0$, $\beta < 1$) is reduced to a tiny region around the ΛCDM cosmology. For a related investigation, see [96].

The literature on this type of generalized gravity models is rapidly increasing.

Brane-World Models

Certain brane-world models[13] lead to modifications of Friedmann cosmology at very large scales. An interesting example has been proposed by Dvali, Gabadadze, and Porrati (DGP), for which the theory remains four-dimensional at 'short' distances, but crosses over to higher-dimensional behavior of gravity at some very large distance [97]. This model has the same number of parameters as the successful ΛCDM cosmology, but contains no dark energy. The resulting modified Friedmann equations can give rise to universes with accelerated expansion, due to an infrared modification of gravity.

In [100] the predictions of the model have been confronted with latest supernovae data [57], and the position of the acoustic peak in the Sloan digital sky survey (SDSS) correlation function for a luminous red galaxy sample [101]. The result is that a flat DGP brane model is ruled out at 3σ. A similar analysis was more recently performed in [99], including also the CMB shift parameter that effectively determines the first acoustic peak (see Sect. 8.1). The authors arrive at the conclusion that the flat DGP models are within the 1σ contours, but that the flat ΛCDM model provides a better fit to the data. They also point out some level of uncertainty in the use of the data, and conservatively conclude that the flat DGP models are within joint 2σ contours.

[12] It is shown in [93] that if the matter action is independent of Γ, the theory is dynamically equivalent to a Brans-Dicke theory with Brans-Dicke parameter $-3/2$, plus a potentiel term.

[13] For a review, see [98].

This nicely illustrates that observational data are restricting theoretical speculations more and more.

The DGP models have, however, serious defects on a fundamental level. A detailed analysis of the excitations about the self-accelerating solution showed that there is a *ghost mode* (negative kinetic energy) [102, 103]. Furthermore, it has very recently been pointed out [104] that due to superluminal fluctuations around non-trivial backgrounds, there is *no local causal evolution*. This infrared breakdown also happens for other apparently consistent low-energy effective theories.

* * *

The previous discussion should have made it clear that it is extremely difficult to construct consistent modifications of GR that lead to an accelerated universe at late times. The dark energy problems will presumably stay with us for a long time. Understanding the nature of DE is widely considered as one of the main goals of cosmological research for the next decade and beyond.

A Essentials of Friedmann–Lemaître Models

In this Appendix those parts of the standard model of cosmology that are needed throughout the text will be briefly introduced. More extensive treatments can be found at many places, for instance in the recent textbooks on cosmology [105], [106], [107], [108], [109].

A.1 Friedmann–Lemaître Spacetimes

There is now good evidence that the (recent as well as the early) Universe[14] is – on large scales – surprisingly homogeneous and isotropic. The most impressive support for this comes from extended redshift surveys of galaxies and from the truly remarkable isotropy of the CMB. In the two degree field (2dF) galaxy redshift survey,[15] completed in 2003, the redshifts of about 250,000 galaxies have been measured. The distribution of galaxies out to 4 billion light years shows that there are huge clusters, long filaments, and empty voids measuring over 100 million light years across. But the map also shows

[14] By *Universe* I always mean that part of the world around us which is in principle accessible to observations. In my opinion the 'Universe as a whole' is not a scientific concept. When talking about *model universes*, we develop on paper or with the help of computers, I tend to use lower case letters. In this domain we are, of course, free to make extrapolations and venture into speculations, but one should always be aware that there is the danger to be drifted into a kind of 'cosmo-mythology'.

[15] Consult the Home Page: http://www.mso.anu.edu. au/2dFGRS.

that there are *no larger structures*. The more extended SDSS has already produced very similar results, and will in the end have spectra of about a million galaxies.[16]

One arrives at the Friedmann(–Lemaître–Robertson–Walker) spacetimes by postulating that for each observer, moving along an integral curve of a distinguished four-velocity field u, the Universe looks spatially isotropic. Mathematically, this means the following: Let $Iso_x(M)$ be the group of local isometries of a Lorentz manifold (M, g), with fixed point $x \in M$, and let $SO_3(u_x)$ be the group of all linear transformations of the tangent space $T_x(M)$ which leave the four-velocity u_x invariant and induce special orthogonal transformations in the subspace orthogonal to u_x, then

$$\{T_x\phi : \ \phi \in Iso_x(M), \ \phi_\star u = u\} \supseteq SO_3(u_x)$$

($\phi_\star$ denotes the push-forward belonging to ϕ; see [1], p. 550). In [110] it is shown that this requirement implies that (M, g) is a Friedmann spacetime, whose structure we now recall. Note that (M, g) is then automatically homogeneous.

A *Friedmann spacetime* (M, g) is a warped product of the form $M = I \times \Sigma$, where I is an interval of $\mathbb{R}$, and the metric g is of the form

$$g = -dt^2 + a^2(t)\gamma \ , \tag{85}$$

such that (Σ, γ) is a Riemannian space of constant curvature $k = 0, \pm 1$. The distinguished time t is the *cosmic time*, and $a(t)$ is the *scale factor* (it plays the role of the warp factor (see Appendix B of [1])). Instead of t we often use the *conformal time* η, defined by $d\eta = dt/a(t)$. The velocity field is perpendicular to the slices of constant cosmic time, $u = \partial/\partial t$.

Spaces of Constant Curvature

For the space (Σ, γ) of constant curvature[17] the curvature is given by

$$R^{(3)}(X, Y)Z = k\left[\gamma(Z, Y)X - \gamma(Z, X)Y\right] \ ; \tag{86}$$

in components

$$R^{(3)}_{ijkl} = k(\gamma_{ik}\gamma_{jl} - \gamma_{il}\gamma_{jk}) \ . \tag{87}$$

Hence, the Ricci tensor and the scalar curvature are

$$R^{(3)}_{jl} = 2k\gamma_{jl} \quad , \quad R^{(3)} = 6k \ . \tag{88}$$

[16] For a description and pictures, see the Home Page: http://www.sdss.org/sdss.html.

[17] For a detailed discussion of these spaces I refer – for readers knowing German – to [111] or [112].

For the curvature two-forms we obtain from (87) relative to an orthonormal triad $\{\theta^i\}$

$$\Omega^{(3)}_{ij} = \frac{1}{2} R^{(3)}_{ijkl}\ \theta^k \wedge \theta^l = k\ \theta_i \wedge \theta_j \tag{89}$$

($\theta_i = \gamma_{ik}\theta^k$). The simply connected constant curvature spaces are in n dimensions the (n+1)-sphere S^{n+1} ($k = 1$), the Euclidean space ($k = 0$), and the pseudo-sphere ($k = -1$). Non-simply connected constant curvature spaces are obtained from these by forming quotients with respect to discrete isometry groups. (For detailed derivations, see [111].)

Curvature of Friedmann Spacetimes

Let $\{\bar{\theta}^i\}$ be any orthonormal triad on (Σ, γ). On this Riemannian space the first-structure equations read (we use the notation in [1]; quantities referring to this three-dimensional space are indicated by bars)

$$d\bar{\theta}^i + \bar{\omega}^i{}_j \wedge \bar{\theta}^j = 0\ . \tag{90}$$

On (M, g) we introduce the following orthonormal tetrad:

$$\theta^0 = dt, \quad \theta^i = a(t)\bar{\theta}^i\ . \tag{91}$$

From this and (90) we get

$$d\theta^0 = 0, \quad d\theta^i = \frac{\dot{a}}{a}\theta^0 \wedge \theta^i - a\ \bar{\omega}^i{}_j \wedge \bar{\theta}^j\ . \tag{92}$$

Comparing this with the first-structure equation for the Friedmann manifold implies

$$\omega^0{}_i \wedge \theta^i = 0, \quad \omega^i{}_0 \wedge \theta^0 + \omega^i{}_j \wedge \theta^j = \frac{\dot{a}}{a}\theta^i \wedge \theta^0 + a\ \bar{\omega}^i{}_j \wedge \bar{\theta}^j\ , \tag{93}$$

whence

$$\boxed{\omega^0{}_i = \frac{\dot{a}}{a}\ \theta^i, \quad \omega^i{}_j = \bar{\omega}^i{}_j\ .} \tag{94}$$

The worldlines of *comoving observers* are integral curves of the four-velocity field $u = \partial_t$. We claim that these are geodesics, i.e., that

$$\nabla_u u = 0\ . \tag{95}$$

To show this (and for other purposes) we introduce the basis $\{e_\mu\}$ of vector fields dual to (91). Since $u = e_0$ we have, using the connection forms (94),

$$\nabla_u u = \nabla_{e_0} e_0 = \omega^\lambda{}_0(e_0)e_\lambda = \omega^i{}_0(e_0)e_i = 0\ .$$

A.2 Einstein Equations for Friedmann Spacetimes

Inserting the connection forms (94) into the second-structure equations we readily find for the curvature 2-forms $\Omega^{\mu}{}_{\nu}$:

$$\Omega^0{}_i = \frac{\ddot{a}}{a}\theta^0 \wedge \theta^i , \quad \Omega^i{}_j = \frac{k + \dot{a}^2}{a^2}\theta^i \wedge \theta^j \ . \tag{96}$$

A routine calculation leads to the following components of the Einstein tensor relative to the basis (91)

$$G_{00} = 3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) , \tag{97}$$

$$G_{11} = G_{22} = G_{33} = -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} - \frac{k}{a^2} , \tag{98}$$

$$G_{\mu\nu} = 0 \quad (\mu \neq \nu) \ . \tag{99}$$

In order to satisfy the field equations, the symmetries of $G_{\mu\nu}$ imply that the energy–momentum tensor *must* have the perfect fluid form (see [1], Sect. 1.4.2):

$$T^{\mu\nu} = (\rho + p)u^{\mu}u^{\nu} + pg^{\mu\nu} , \tag{100}$$

where u is the comoving velocity field introduced above.

Now, we can write down the field equations (including the cosmological term),

$$3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) = 8\pi G\rho + \Lambda , \tag{101}$$

$$-2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} - \frac{k}{a^2} = 8\pi G p - \Lambda \ . \tag{102}$$

Although the 'energy–momentum conservation' does not provide an independent equation, it is useful to work this out. As expected, the momentum 'conservation' is automatically satisfied. For the 'energy conservation' we use the general form (see (1.37) in [1])

$$\nabla_u \rho = -(\rho + p)\nabla \cdot u \ . \tag{103}$$

In our case we have for the *expansion rate*

$$\nabla \cdot u = \omega^{\lambda}{}_0(e_{\lambda})u^0 = \omega^i{}_0(e_i) ,$$

thus with (94)

$$\nabla \cdot u = 3\frac{\dot{a}}{a} \ . \tag{104}$$

Therefore, (103) becomes

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0 \ . \tag{105}$$

For a given equation of state, $p = p(\rho)$, we can use (105) in the form

$$\frac{d}{da}(\rho a^3) = -3pa^2 \tag{106}$$

to determine ρ as a function of the scale factor a. Examples: (1) For free massless particles (radiation) we have $p = \rho/3$, thus $\rho \propto a^{-4}$. (2) For dust ($p = 0$) we get $\rho \propto a^{-3}$.

With this knowledge the *Friedmann equation* (101) determines the time evolution of $a(t)$. It is easy to see that (102) follows from (101) and (105).

As an important consequence of (101) and (102) we obtain for the acceleration of the expansion

$$\ddot{a} = -\frac{4\pi G}{3}(\rho + 3p)a + \frac{1}{3}\Lambda a \; . \tag{107}$$

This shows that as long as $\rho + 3p$ is positive, the first term in (107) is decelerating, while a positive cosmological constant is repulsive. This becomes understandable if one writes the field equation as

$$G_{\mu\nu} = \kappa(T_{\mu\nu} + T^{\Lambda}_{\mu\nu}) \qquad (\kappa = 8\pi G) \; , \tag{108}$$

with

$$T^{\Lambda}_{\mu\nu} = -\frac{\Lambda}{8\pi G} g_{\mu\nu} \; . \tag{109}$$

This vacuum contribution has the form of the energy–momentum tensor of an ideal fluid, with energy density $\rho_\Lambda = \Lambda/8\pi G$ and pressure $p_\Lambda = -\rho_\Lambda$. Hence the combination $\rho_\Lambda + 3p_\Lambda$ is equal to $-2\rho_\Lambda$, and is thus negative. In what follows we shall often include in ρ and p the vacuum pieces.

A.3 Redshift

As a result of the expansion of the Universe the light of distant sources appears redshifted. The amount of redshift can be simply expressed in terms of the scale factor $a(t)$.

Consider two integral curves of the average velocity field u. We imagine that one describes the worldline of a distant comoving source and the other that of an observer at a telescope (see Fig. 9). Since light is propagating along null geodesics, we conclude from (85) that along the worldline of a light ray $dt = a(t)d\sigma$, where $d\sigma$ is the line element on the three-dimensional space (Σ, γ) of constant curvature $k = 0, \pm 1$. Hence the integral on the left of

$$\int_{t_e}^{t_o} \frac{dt}{a(t)} = \int_{source}^{obs.} d\sigma \; , \tag{110}$$

between the time of emission (t_e) and the arrival time at the observer (t_o) is independent of t_e and t_o. Therefore, if we consider a second light ray that is

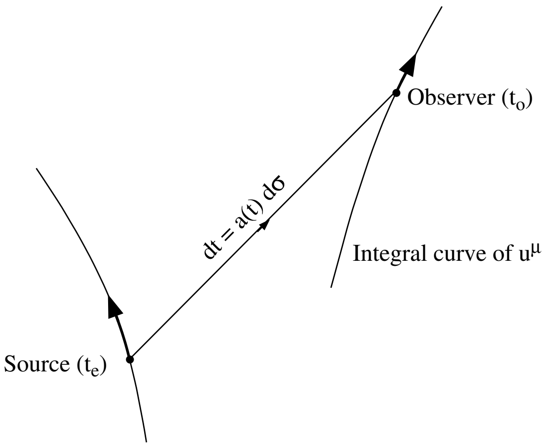

Fig. 9. Redshift for Friedmann models

emitted at the time $t_e + \Delta t_e$ and is received at the time $t_o + \Delta t_o$, we obtain from the last equation

$$\int_{t_e+\Delta t_e}^{t_o+\Delta t_o} \frac{dt}{a(t)} = \int_{t_e}^{t_o} \frac{dt}{a(t)} \; . \tag{111}$$

For a small Δt_e this gives

$$\frac{\Delta t_o}{a(t_o)} = \frac{\Delta t_e}{a(t_e)} \; .$$

The observed and the emitted frequences ν_o and ν_e, respectively, are thus related according to

$$\frac{\nu_o}{\nu_e} = \frac{\Delta t_e}{\Delta t_o} = \frac{a(t_e)}{a(t_o)} \; . \tag{112}$$

The redshift parameter z is defined by

$$z := \frac{\nu_e - \nu_o}{\nu_o} \; , \tag{113}$$

and is given by the key equation

$$\boxed{1 + z = \frac{a(t_o)}{a(t_e)}.} \tag{114}$$

One can also express this by the equation $\nu \cdot a = const$ along a null geodesic.

A.4 Cosmic Distance Measures

We now introduce a further important tool, namely operational definitions of three different distance measures, and show that they are related by simple redshift factors.

If D is the physical (proper) extension of a distant object and δ is its angle subtended, then the *angular diameter distance* D_A is defined by

$$D_A := D/\delta \,. \tag{115}$$

If the object is moving with the proper transversal velocity $V_\perp$ and with an apparent angular motion $d\delta/dt_0$, then the *proper-motion distance* is by definition

$$D_M := \frac{V_\perp}{d\delta/dt_0} \,. \tag{116}$$

Finally, if the object has the intrinsic luminosity $\mathcal{L}$ and $\mathcal{F}$ is the received energy flux then the *luminosity distance* is naturally defined as

$$D_L := (\mathcal{L}/4\pi\mathcal{F})^{1/2} \,. \tag{117}$$

Below we show that these three distances are related as follows

$$\boxed{D_L = (1+z)D_M = (1+z)^2 D_A.} \tag{118}$$

It will be useful to introduce on (Σ, γ) 'polar' coordinates (r, ϑ, φ), such that

$$\gamma = \frac{dr^2}{1-kr^2} + r^2 d\Omega^2, \quad d\Omega^2 = d\vartheta^2 + \sin^2\vartheta d\varphi^2 \,. \tag{119}$$

One easily verifies that the curvature forms of this metric satisfy (89). (This follows without doing any work by using in [1] the curvature forms (3.9) in the ansatz (3.3) for the Schwarzschild metric.)

To prove (118) we show that the three distances can be expressed as follows, if r_e denotes the comoving radial coordinate (in (119)) of the distant object and the observer is (without loss of generality) at $r = 0$.

$$D_A = r_e a(t_e), \quad D_M = r_e a(t_0), \quad D_L = r_e a(t_0)\frac{a(t_0)}{a(t_e)} \,. \tag{120}$$

Once this is established, (118) follows from (114).

From Fig. 10 and (119) we see that

$$D = a(t_e) r_e \delta \,, \tag{121}$$

hence the first equation in (120) holds.

To prove the second one we note that the source moves in a time dt_0 a proper transversal distance

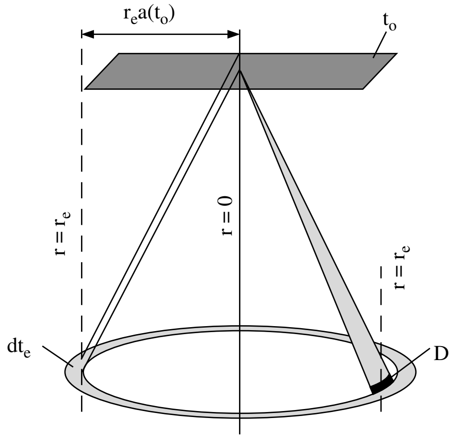

Fig. 10. Spacetime diagram for cosmic distance measures. The angular diameter distance $D_{ang} \equiv D_A$ and the luminosity distance $D_{lum} \equiv D_L$ have been introduced in this Appendix. The other two will be introduced in the Appendix C

$$dD = V_\perp dt_e = V_\perp dt_0 \frac{a(t_e)}{a(t_0)} \ .$$

Using again the metric (119) we see that the apparent angular motion is

$$d\delta = \frac{dD}{a(t_e) r_e} = \frac{V_\perp dt_0}{a(t_0) r_e} \ .$$

Inserting this into the definition (116) shows that the second equation in (120) holds. For the third equation we have to consider the observed energy flux. In a time dt_e the source emits an energy $\mathcal{L} dt_e$. This energy is redshifted to the present by a factor $a(t_e)/a(t_0)$, and is now distributed by (119) over a sphere with proper area $4\pi(r_e a(t_0))^2$ (see Fig. 10). Hence the received flux (*apparent luminosity*) is

$$\mathcal{F} = \mathcal{L} dt_e \frac{a(t_e)}{a(t_0)} \frac{1}{4\pi(r_e a(t_0))^2} \frac{1}{dt_0} \ ,$$

thus

$$\mathcal{F} = \frac{\mathcal{L} a^2(t_e)}{4\pi a^4(t_0) r_e^2} \ .$$

Inserting this into the definition (117) establishes the third equation in (120). For later applications we write the last equation in the more transparent form

$$\boxed{\mathcal{F} = \frac{\mathcal{L}}{4\pi(r_e a(t_0))^2} \frac{1}{(1+z)^2}.} \qquad (122)$$

The last factor is due to redshift effects.

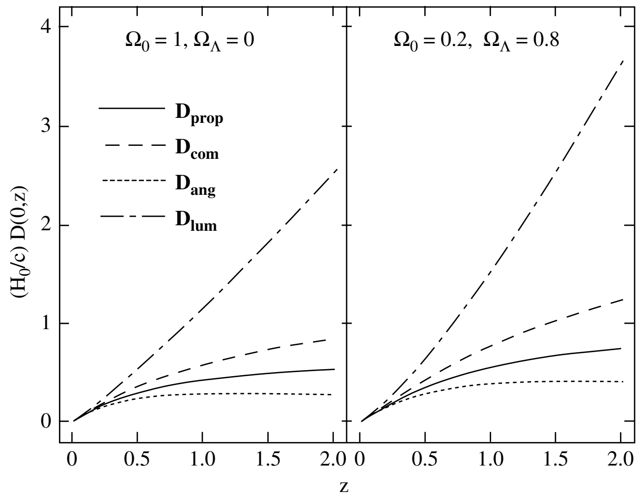

Fig. 11. Cosmological distance measures as a function of source redshift for two cosmological models

Two of the discussed distances as a function of z are shown in Fig. 11 for two Friedmann models with different cosmological parameters. The other two distance measures will be introduced in Appendix C.

B Thermal History below 100 MeV

B.1 Overview

Below the transition at about 200 MeV from a quark-gluon plasma to the confinement phase, the Universe was initially dominated by a complicated dense hadron soup. The abundance of pions, for example, was so high that they nearly overlapped. The pions, kaons, and other hadrons soon began to decay and most of the nucleons and antinucleons annihilated, leaving only a tiny baryon asymmetry. The energy density is then almost completely dominated by radiation and the stable leptons ($e^{\pm}$, the three neutrino flavors, and their antiparticles). For some time all these particles are in thermodynamic equilibrium. For this reason, only a few initial conditions have to be imposed. The Universe was never as simple as in this lepton era. (At this stage it is almost inconceivable that the complex world around us would eventually emerge.)

The first particles which freeze out of this equilibrium are the weakly interacting neutrinos. Let us estimate when this happened. The coupling of the neutrinos in the lepton era is dominated by the reactions:

$$e^- + e^+ \leftrightarrow \nu + \bar{\nu}\ , \quad e^{\pm} + \nu \to e^{\pm} + \nu\ , \quad e^{\pm} + \bar{\nu} \to e^{\pm} + \bar{\nu}\ .$$

For dimensional reasons, the cross sections are all of magnitude

$$\sigma \simeq G_F^2 T^2 \ , \tag{123}$$

where G_F is the Fermi coupling constant ($\hbar = c = k_B = 1$). Numerically, $G_F m_p^2 \simeq 10^{-5}$. On the other hand, the electron and neutrino densities n_e, n_ν are about T^3. For this reason, the reaction rates Γ for ν-scattering and ν-production per electron are of magnitude $c \cdot v \cdot n_e \simeq G_F^2 T^5$. This has to be compared with the expansion rate of the Universe

$$H = \frac{\dot{a}}{a} \simeq (G\rho)^{1/2} \ .$$

Since $\rho \simeq T^4$ we get

$$H \simeq G^{1/2} T^2 \ , \tag{124}$$

and thus

$$\frac{\Gamma}{H} \simeq G^{-1/2} G_F^2 T^3 \simeq (T/10^{10}\ K)^3 \ . \tag{125}$$

This ration is larger than 1 for $T > 10^{10}\ K \simeq 1$ MeV, and the neutrinos thus remain in thermodynamic equilibrium until the temperature has decreased to about 1 MeV. But even below this temperature the neutrinos remain Fermi distributed,

$$n_\nu(p)dp = \frac{1}{2\pi^2} \frac{1}{e^{p/T_\nu} + 1} p^2 dp \ \ , \tag{126}$$

as long as they can be treated as massless. The reason is that the number density decreases as a^{-3} and the momenta with a^{-1}. Because of this we also see that the neutrino temperature T_ν decreases after decoupling as a^{-1}. The same is, of course, true for photons. The reader will easily find out how the distribution evolves when neutrino masses are taken into account. (Since neutrino masses are so small this is only relevant at very late times.)

B.2 Chemical Potentials of the Leptons

The equilibrium reactions below 100 MeV, say, conserve several additive quantum numbers,[18] namely the electric charge Q, the baryon number B, and the three lepton numbers L_e, L_μ, L_τ. Correspondingly, there are five independent chemical potentials. Since particles and antiparticles can annihilate to photons, their chemical potentials are oppositely equal: $\mu_{e^-} = -\mu_{e^+}$, etc. From the following reactions

$$e^- + \mu^+ \to \nu_e + \bar{\nu}_\mu, \quad e^- + p \to \nu_e + n, \quad \mu^- + p \to \nu_\mu + n$$

[18] Even if B, L_e, L_μ, L_τ should not be strictly conserved, this is not relevant within a Hubble time H_0^{-1}.

we infer the equilibrium conditions

$$\mu_{e^-} - \mu_{\nu_e} = \mu_{\mu^-} - \mu_{\nu_\mu} = \mu_n - \mu_p \,. \tag{127}$$

As independent chemical potentials we can thus choose

$$\boxed{\mu_p,\ \mu_{e^-},\ \mu_{\nu_e},\ \mu_{\nu_\mu},\ \mu_{\nu_\tau}.} \tag{128}$$

Because of local electric charge neutrality, the charge number density n_Q vanishes. From observations (see subsection E) we also know that the baryon number density n_b is much smaller than the photon number density ($\sim$ entropy density s_γ). The ratio n_B/s_γ remains constant for adiabatic expansion (both decrease with a^{-3}; see the next section). Moreover, the lepton number densities are

$$n_{L_e} = n_{e^-} + n_{\nu_e} - n_{e^+} - n_{\bar{\nu}_e}, \quad n_{L_\mu} = n_{\mu^-} + n_{\nu_\mu} - n_{\mu^+} - n_{\bar{\nu}_\mu}, \quad etc\,. \tag{129}$$

Since in the present Universe the number density of electrons is equal to that of the protons (bound or free), we know that after the disappearance of the muons $n_{e^-} \simeq n_{e^+}$ (recall $n_B \ll n_\gamma$), thus μ_{e^-} $(= -\mu_{e^+}) \simeq 0$. It is conceivable that the chemical potentials of the neutrinos and antineutrinos cannot be neglected, i.e., that n_{L_e} is not much smaller than the photon number density. In analogy to what we know about the baryon density we make the reasonable *asumption* that the lepton number densities are also much smaller than s_γ. Then we can take the chemical potentials of the neutrinos equal to zero ($|\mu_\nu|/kT \ll 1$). With what we said before, we can then put the five chemical potentials (128) equal to zero, because the charge number densities are all odd in them. Of course, n_B does not really vanish (otherwise we would not be here), but for the thermal history in the era we are considering they can be ignored.

B.3 Constancy of Entropy

Let ρ_{eq}, p_{eq} denote (in this subsection only) the total energy density and pressure of all particles in thermodynamic equilibrium. Since the chemical potentials of the leptons vanish, these quantities are only functions of the temperature T. According to the second law, the differential of the entropy $S(V,T)$ is given by

$$dS(V,T) = \frac{1}{T}[d(\rho_{eq}(T)V) + p_{eq}(T)dV]\,. \tag{130}$$

This implies

$$\begin{aligned} d(dS) = 0 &= d\left(\frac{1}{T}\right) \wedge d(\rho_{eq}(T)V) + d\left(\frac{p_{eq}(I)}{T}\right) \wedge dV \\ &= -\frac{\rho_{eq}}{T^2} dT \wedge dV + \frac{d}{dT}\left(\frac{p_{eq}(T)}{T}\right) dT \wedge dV\,, \end{aligned}$$

i.e., the Maxwell relation

$$\boxed{\frac{dp_{eq}(T)}{dT} = \frac{1}{T}[\rho_{eq}(T) + p_{eq}(T)] \ .} \tag{131}$$

If we use this in (130), we get

$$dS = d\left[\frac{V}{T}(\rho_{eq} + p_{eq})\right] \ ,$$

so the entropy density of the particles in equilibrium is

$$\boxed{s = \frac{1}{T}[\rho_{eq}(T) + p_{eq}(T)] \ .} \tag{132}$$

For an adiabatic expansion the entropy in a comoving volume remains constant:

$$S = a^3 s = \text{const} \ . \tag{133}$$

This constancy is equivalent to the energy equation (105) for the equilibrium part. Indeed, the latter can be written as

$$a^3 \frac{dp_{eq}}{dt} = \frac{d}{dt}[a^3(\rho_{eq} + p_{eq})] \ ,$$

and by (132) this is equivalent to $dS/dt = 0$.

In particular, we obtain for massless particles ($p = \rho/3$) from (131) again $\rho \propto T^4$ and from (132) that $S =$ constant implies $T \propto a^{-1}$.

Once the electrons and positrons have annihilated below $T \sim m_e$, the equilibrium components consist of photons, electrons, protons, and – after the big bang nucleosynthesis – of some light nuclei (mostly He^4). Since the charged particle number densities are much smaller than the photon number density, the photon temperature T_γ still decreases as a^{-1}. Let us show this formally. For this we consider beside the photons an ideal gas in thermodynamic equilibrium with the black body radiation. The total pressure and energy density are then (we use units with $\hbar = c = k_B = 1$; n is the number density of the non-relativistic gas particles with mass m):

$$p = nT + \frac{\pi^2}{45}T^4, \quad \rho = nm + \frac{nT}{\gamma - 1} + \frac{\pi^2}{15}T^4 \tag{134}$$

($\gamma = 5/3$ for a monoatomic gas). The conservation of the gas particles, $na^3 =$ const., together with the energy equation (106) implies, if $\sigma := s_\gamma/n$,

$$\frac{d\ln T}{d\ln a} = -\left[\frac{\sigma + 1}{\sigma + 1/3(\gamma - 1)}\right] \ .$$

For $\sigma \ll 1$ this gives the well-known relation $T \propto a^{3(\gamma-1)}$ for an adiabatic expansion of an ideal gas.

We are, however, dealing with the opposite situation $\sigma \gg 1$, and then we obtain, as expected, $a \cdot T = \text{const}$.

Let us look more closely at the famous ratio n_B/s_γ. We need

$$s_\gamma = \frac{4}{3T}\rho_\gamma = \frac{4\pi^2}{45}T^3 = 3.60 n_\gamma, \quad n_B = \rho_B/m_p = \Omega_B \rho_{crit}/m_p \,. \qquad (135)$$

From the present value of $T_\gamma \simeq 2.7\ K$ and (30), $\rho_{crit} = 1.12 \times 10^{-5}\ h_0^2 (m_p/\text{cm}^3)$, we obtain as a measure for the baryon asymmetry of the Universe

$$\boxed{\frac{n_B}{s_\gamma} = 0.75 \times 10^{-8} (\Omega_B h_0^2) \,.} \qquad (136)$$

It is one of the great challenges to explain this tiny number. So far, this has been achieved at best qualitatively in the framework of grand unified theories (GUTs).

B.4 Neutrino Temperature

During the electron–positron annihilation below $T = m_e$ the a-dependence is complicated, since the electrons can no more be treated as massless. We want to know at this point what the ratio T_γ/T_ν is after the annihilation. This can easily be obtained by using the constancy of comoving entropy for the photon–electron–positron system, which is sufficiently strongly coupled to maintain thermodynamic equilibrium.

We need the entropy for the electrons and positrons at $T \gg m_e$, long before annihilation begins. To compute this note the identity

$$\int_0^\infty \frac{x^n}{e^x - 1}dx - \int_0^\infty \frac{x^n}{e^x + 1}dx = 2\int_0^\infty \frac{x^n}{e^{2x} - 1}dx = \frac{1}{2^n}\int_0^\infty \frac{x^n}{e^x - 1}dx \,,$$

whence

$$\int_0^\infty \frac{x^n}{e^x + 1}dx = (1 - 2^{-n})\int_0^\infty \frac{x^n}{e^x - 1}dx \,. \qquad (137)$$

In particular, we obtain for the entropies s_e, s_γ the following relation

$$s_e = \frac{7}{8}s_\gamma \quad (T \gg m_e) \,. \qquad (138)$$

Equating the entropies for $T_\gamma \gg m_e$ and $T_\gamma \ll m_e$ gives

$$(T_\gamma a)^3\big|_{before}\left[1 + 2 \times \frac{7}{8}\right] = (T_\gamma a)^3\big|_{after} \times 1 \,,$$

because the neutrino entropy is conserved. Therefore, we obtain

$$(aT_\gamma)|_{after} = \left(\frac{11}{4}\right)^{1/3}(aT_\gamma)|_{before} \,. \qquad (139)$$

But $(aT_\nu)|_{after} = (aT_\nu)|_{before} = (aT_\gamma)|_{before}$, hence we obtain the important relation

$$\boxed{\left(\frac{T_\gamma}{T_\nu}\right)\bigg|_{after} = \left(\frac{11}{4}\right)^{1/3} = 1.401\,.} \tag{140}$$

B.5 Epoch of Matter–Radiation Equality

In the main parts of these lectures the epoch when radiation (photons and neutrinos) have about the same energy density as non-relativistic matter (dark matter and baryons) plays a very important role. Let us determine the redshift, z_{eq}, when there is equality.

For the three neutrino and antineutrino flavors the energy density is according to (137)

$$\rho_\nu = 3 \times \frac{7}{8} \times \left(\frac{4}{11}\right)^{4/3} \rho_\gamma\,. \tag{141}$$

Using

$$\frac{\rho_\gamma}{\rho_{crit}} = 2.47 \times 10^{-5} h_0^{-2} (1+z)^4\,, \tag{142}$$

we obtain for the total radiation energy density, ρ_r,

$$\frac{\rho_r}{\rho_{crit}} = 4.15 \times 10^{-5} h_0^{-2} (1+z)^4\,. \tag{143}$$

Equating this to

$$\frac{\rho_M}{\rho_{crit}} = \Omega_M (1+z)^3 \tag{144}$$

we obtain

$$\boxed{1 + z_{eq} = 2.4 \times 10^4 \Omega_M h_0^2\,.} \tag{145}$$

Only a small fraction of Ω_M is baryonic. There are several methods to determine the fraction Ω_B in baryons. A traditional one comes from the abundances of the light elements. This is treated in most texts on cosmology. (German-speaking readers find a detailed discussion in my lecture notes [112], which are available in the Internet.) The comparison of the straightforward theory with observation gives a value in the range $\Omega_B h_0^2 = 0.021 \pm 0.002$. Other determinations are all compatible with this value. In Sect. 8 we shall obtain Ω_B from the CMB anisotropies. The striking agreement of different methods, sensitive to different physics, strongly supports our standard big bang picture of the Universe.

C Inflation and Primordial Power Spectra

C.1 Introduction

The horizon and flatness problems of standard big bang cosmology are so serious that the proposal of a very early accelerated expansion, preceding the hot era dominated by relativistic fluids, appears quite plausible. This general qualitative aspect of 'inflation' is now widely accepted. However, when it comes to concrete model building the situation is not satisfactory. Since we do not know the fundamental physics at superhigh energies not too far from the Planck scale, models of inflation are usually of a phenomenological nature. Most models consist of a number of scalar fields, including a suitable form for their potential. Usually there is no direct link to fundamental theories, like supergravity; however, there have been many attempts in this direction. For the time being, inflationary cosmology should be regarded as an attractive scenario, and not yet as a theory.

The most important aspect of inflationary cosmology is that *the generation of perturbations on large scales from initial quantum fluctuations is unavoidable and predictable.* For a given model these fluctuations can be calculated accurately, because they are tiny and cosmological perturbation theory can be applied. And, most importantly, these predictions can be *confronted with the cosmic microwave anisotropy measurements.* We are in the fortunate position to witness rapid progress in this field. The results from various experiments, most recently from WMAP, give already strong support of the basic predictions of inflation. Further experimental progress can be expected in the coming years.

C.2 The Horizon Problem and the General Idea of Inflation

I begin by describing the famous horizon puzzle, which is a very serious causality problem of standard big bang cosmology.

Past and Future Light Cone Distances

Consider our past light cone for a Friedmann spacetime model (Fig. 12). For a radial light ray the differential relation $dt = a(t)dr/(1-kr^2)^{1/2}$ holds for the coordinates (t, r) of the metric (19). The proper radius of the past light sphere at time t (cross section of the light cone with the hypersurface $\{t = const\}$) is

$$l_p(t) = a(t)\int_0^{r(t)} \frac{dr}{\sqrt{1-kr^2}}\ , \tag{146}$$

where the coordinate radius is determined by

$$\int_0^{r(t)} \frac{dr}{\sqrt{1-kr^2}} = \int_t^{t_0} \frac{dt'}{a(t')}\ . \tag{147}$$

Hence,

$$l_p(t) = a(t)\int_t^{t_0} \frac{dt'}{a(t')}\ . \tag{148}$$

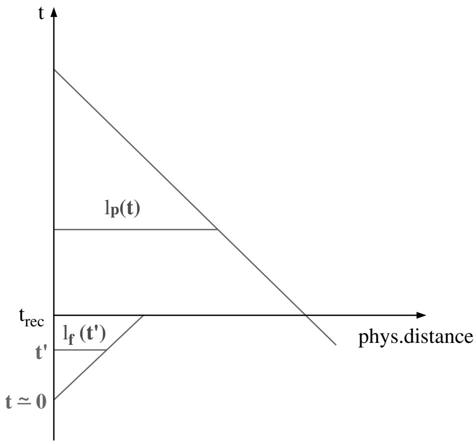

Fig. 12. Spacetime diagram illustrating the horizon problem

We rewrite this in terms of the redhift variable. From $1 + z = a_0/a$ we get $dz = -(1 + z)H dt$, so

$$\frac{dt}{dz} = -\frac{1}{H_0(1+z)E(z)}, \quad H(z) = H_0 E(z) \; .$$

Therefore,

$$l_p(z) = \frac{1}{H_0(1+z)} \int_0^z \frac{dz'}{E(z')} \; . \tag{149}$$

Similarly, the extension $l_f(t)$ of the forward light cone at time t of a very early event ($t \simeq 0, \; z \simeq \infty$) is

$$l_f(t) = a(t) \int_0^t \frac{dt'}{a(t')} = \frac{1}{H_0(1+z)} \int_z^\infty \frac{dz'}{E(z')} \; . \tag{150}$$

For the present Universe (t_0) this becomes what is called the *particle horizon distance*

$$D_{hor} = H_0^{-1} \int_0^\infty \frac{dz'}{E(z')} \; , \tag{151}$$

and gives the size of the *observable Universe*.

Analytical expressions for these distances are only available in special cases. For orientation we consider first the Einstein–de Sitter model ($K = 0, \; \Omega_\Lambda = 0, \; \Omega_M = 1$), for which $a(t) = a_0(t/t_0)^{2/3}$ and thus

$$D_{hor} = 3t_0 = 2H_0^{-1}, \quad l_f(t) = 3t \; , \quad \frac{l_p}{l_f} = \left(\frac{t_0}{t}\right)^{1/3} - 1 = \sqrt{1+z} - 1 \; . \tag{152}$$

For a flat Universe a good fitting formula for cases of interest is (Hu and White)

$$D_{hor} \simeq 2H_0^{-1} \frac{1 + 0.084 \ln \Omega_M}{\sqrt{\Omega_M}} \,. \tag{153}$$

It is often convenient to work with 'comoving distances', by rescaling distances referring to time t (like $l_p(t), l_f(t)$) with the factor $a(t_0)/a(t) = 1 + z$ to the present. We indicate this by the superscript c. For instance,

$$l_p^c(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{E(z')} \,. \tag{154}$$

This distance is plotted in Fig. 11 of Appendix A as $D_{com}(z)$. Note that for $a_0 = 1 :\ l_f^c(\eta) = \eta,\ l_p^c(\eta) = \eta_0 - \eta$. Hence (150) gives the following relation between η and z:

$$\eta = \frac{1}{H_0} \int_z^\infty \frac{dz'}{E(z')} \,.$$

The Number of Causality Distances on the Cosmic Photosphere

The number of causality distances at redshift z between two antipodal emission points is equal to $l_p(z)/l_f(z)$, and thus the ratio of the two integrals on the right of (149) and (150). We are particularly interested in this ratio at the time of last scattering with $z_{rec} \simeq 1100$. Then we can use for the numerator a flat Universe with non-relativistic matter, while for the denominator we can neglect in the standard hot big bang model Ω_K and Ω_Λ. A reasonable estimate is already obtained by using the simple expression in (152), i.e., $z_{rec}^{1/2} \approx 30$. A more accurate evaluation would increase this number to about 40. The length $l_f(z_{rec})$ subtends an angle of about 1 degree (exercise). How can it be that there is such a large number of causally disconnected regions we see on the microwave sky all having the same temperature? This is what is meant by the *horizon problem* and was a troublesome mystery before the invention of inflation.

Vacuum-Like Energy and Exponential Expansion

This causality problem is potentially avoided, if $l_f(t)$ would be increased in the very early Universe as a result of different physics. If a vacuum-like energy density would dominate, the Universe would undergo an *exponential expansion*. Indeed, in this case the Friedmann equation is

$$\left(\frac{\dot a}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\rho_{vac}, \quad \rho_{vac} \simeq \text{const} \,, \tag{155}$$

and has the solutions

$$a(t) \propto \begin{cases} \cosh\ H_{vac}t & : \quad k = 1 \\ e^{H_{vac}t} & : \quad k = 0 \\ \sinh\ H_{vac}t & : \quad k = 1\ , \end{cases} \tag{156}$$

with

$$H_{vac} = \sqrt{\frac{8\pi G}{3}\rho_{vac}}\ . \tag{157}$$

Assume that such an exponential expansion starts for some reason at time t_i and ends at the *reheating time* t_e, after which standard expansion takes over. From

$$a(t) = a(t_i)e^{H_{vac}(t-t_i)} \quad (t_i < t < t_e)\ , \tag{158}$$

for $k = 0$ we get

$$l_f^c(t_e) \simeq a_0 \int_{t_i}^{t_e} \frac{dt}{a(t)} = \frac{a_0}{H_{vac}a(t_i)}\left(1 - e^{-H_{vac}\Delta t}\right) \simeq \frac{a_0}{H_{vac}a(t_i)}\ ,$$

where $\Delta t := t_e - t_i$. We want to satisfy the condition $l_f^c(t_e) \gg l_p^c(t_e) \simeq H_0^{-1}$ (see (153)), i.e.,

$$a_i H_{vac} \ll a_0 H_0 \quad \Leftrightarrow \frac{a_i}{a_e} \ll \frac{a_0 H_0}{a_e H_{vac}} \tag{159}$$

or

$$e^{H_{vac}\Delta t} \gg \frac{a_e H_{vac}}{a_0 H_0} = \frac{H_{eq}a_{eq}}{H_0 a_0}\frac{H_{vac}a_e}{H_{eq}a_{eq}}\ .$$

Here, eq indicates the values at the time t_{eq} when the energy densities of non-relativistic and relativistic matter were equal. We now use the Friedmann equation for $k = 0$ and $w := p/\rho = \text{const}$. From (25) it follows that in this case

$$Ha \propto a^{-(1+3w)/2}\ ,$$

and hence we arrive at

$$e^{H_{vac}\Delta t} \gg \left(\frac{a_0}{a_{eq}}\right)^{1/2}\left(\frac{a_{eq}}{a_e}\right) = (1+z_{eq})^{1/2}\left(\frac{T_e}{T_{eq}}\right) = (1+z_{eq})^{-1/2}\frac{T_{Pl}}{T_0}\frac{T_e}{T_{Pl}}\ , \tag{160}$$

where we used $aT = \text{const}$. So the number of e-folding periods during the inflationary period, $\mathcal{N} = H_{vac}\Delta t$, should satisfy

$$\mathcal{N} \gg \ln\left(\frac{T_{Pl}}{T_0}\right) - \frac{1}{2}\ln z_{eq} + \ln\left(\frac{T_e}{T_{Pl}}\right) \simeq 70 + \ln\left(\frac{T_e}{T_{Pl}}\right)\ . \tag{161}$$

For a typical GUT scale, $T_e \sim 10^{14}\ GeV$, we arrive at the condition $\mathcal{N} \gg 60$.

Such an exponential expansion would also solve the *flatness problem*, another worry of standard big bang cosmology. Let me recall how this problem arises.

The Friedmann equation (101) can be written as

$$(\Omega^{-1}-1)\rho a^2 = -\frac{3k}{8\pi G} = \text{const.}\,,$$

where

$$\Omega(t) := \frac{\rho(t)}{3H^2/8\pi G} \tag{162}$$

(ρ includes vacuum energy contributions). Thus

$$\Omega^{-1}-1 = (\Omega_0^{-1}-1)\frac{\rho_0 a_0^2}{\rho a^2}\,. \tag{163}$$

Without inflation we have

$$\rho = \rho_{eq}\left(\frac{a_{eq}}{a}\right)^4 \quad (z > z_{eq})\,, \tag{164}$$

$$\rho = \rho_0\left(\frac{a_0}{a}\right)^3 \quad (z < z_{eq})\,. \tag{165}$$

According to (26) z_{eq} is given by

$$1+z_{eq} = \frac{\Omega_M}{\Omega_R} \simeq 10^4\ \Omega_0 h_0^2\,. \tag{166}$$

For $z > z_{eq}$ we obtain from (163) and (164)

$$\Omega^{-1}-1 = (\Omega_0^{-1}-1)\frac{\rho_0 a_0^2}{\rho_{eq} a_{eq}^2}\frac{\rho_{eq} a_{eq}^2}{\rho a^2} = (\Omega_0^{-1}-1)(1+z_{eq})^{-1}\left(\frac{a}{a_{eq}}\right)^2 \tag{167}$$

or

$$\Omega^{-1}-1 = (\Omega_0^{-1}-1)(1+z_{eq})^{-1}\left(\frac{T_{eq}}{T}\right)^2 \simeq 10^{-60}(\Omega_0^{-1}-1)\left(\frac{T_{Pl}}{T}\right)^2\,. \tag{168}$$

Let us apply this equation for $T = 1\,\text{MeV}$, $\Omega_0 \simeq 0.2-0.3$. Then $\mid \Omega - 1 \mid \leq 10^{-15}$, thus the Universe was already incredibly flat at modest temperatures, not much higher than at the time of nucleosynthesis.

Such a fine tuning must have a physical reason. This is naturally provided by inflation, because our observable Universe could originate from a small patch at t_e. (A tiny part of the Earth surface is also practically flat.)

Beside the horizon scale $l_f(t)$, the *Hubble length* $H^{-1}(t) = a(t)/\dot{a}(t)$ plays also an important role. One might call this the "microphysics horizon", because this is the maximal distance microphysics can operate coherently in one expansion time. It is this length scale which enters in basic evolution equations, such as the equation of motion for a scalar field (see (175) below).

We sketch in Figs. 13–15 the various length scales in inflationary models, that is for models with a period of accelerated (e.g., exponential) expansion.

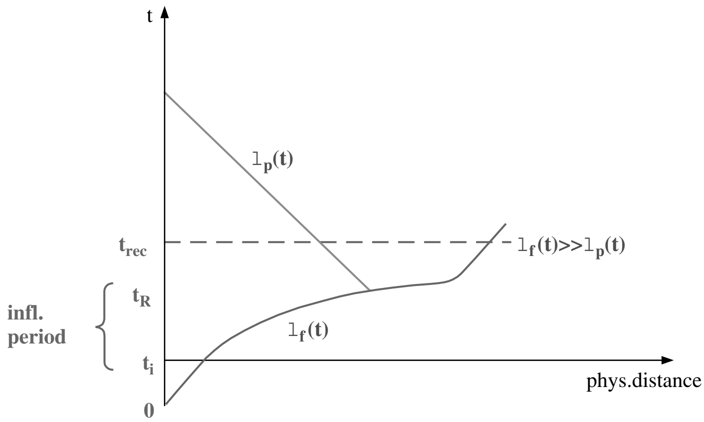

Fig. 13. Past and future light cones in models with an inflationary period

From these it is obvious that there can be – at least in principle – a *causal generation mechanism for perturbations*. This topic will be discussed in great detail in later parts of these lectures.

Exponential inflation is just an example. What we really need is an early phase during which the *comoving Hubble length decreases* (Fig. 15). This means that (for Friedmann spacetimes)

$$\boxed{\left(H^{-1}(t)/a\right)^{\cdot} < 0 \, .} \tag{169}$$

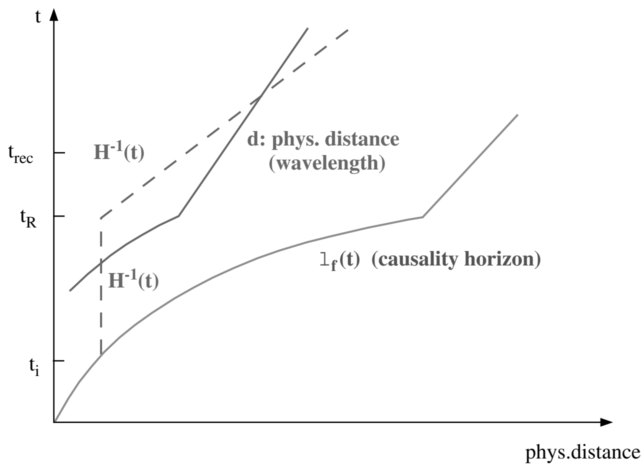

Fig. 14. Physical distance (e.g., between clusters of galaxies) and Hubble distance, and causality horizon in inflationary models

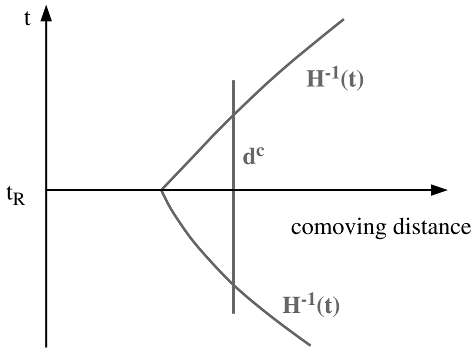

Fig. 15. Part of Fig. 14 expressed in terms of comoving distances

This is the *general definition of inflation*; equivalently, $\ddot{a} > 0$ (accelerated expansion). For a Friedmann model (107) tells us that

$$\ddot{a} > 0 \Leftrightarrow p < -\rho/3 \; . \tag{170}$$

This is, of course, not satisfied for 'ordinary' fluids.

Assume, as another example, *power-law inflation*: $a \propto t^p$. Then $\ddot{a} > 0 \Leftrightarrow p > 1$.

C.3 Scalar Field Models

Models with $p < -\rho/3$ are naturally obtained in scalar field theories. Most of the time we shall consider the simplest case of *one* neutral scalar field φ minimally coupled to gravity. Thus the Lagrangian density is assumed to be

$$\mathcal{L} = \frac{M_{pl}^2}{16\pi} R[g] - \frac{1}{2} \nabla_\mu \varphi \nabla^\mu \varphi - V(\varphi \; , \tag{171}$$

where $R[g]$ is the Ricci scalar for the metric g. The scalar field equation is

$$\Box \varphi = V_{,\varphi} \; , \tag{172}$$

and the energy–momentum tensor in the Einstein equation

$$G_{\mu\nu} = \frac{8\pi}{M_{Pl}^2} T_{\mu\nu} \tag{173}$$

is

$$T_{\mu\nu} = \nabla_\mu \varphi \nabla_\nu \varphi + g_{\mu\nu} \mathcal{L}_\varphi \tag{174}$$

($\mathcal{L}_\varphi$ is the scalar field part of (171)).

We consider first Friedmann spacetimes. Using previous notation, we obtain from (85)

$$\sqrt{-g} = a^3\sqrt{\gamma}, \quad \Box\varphi = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\varphi) = -\frac{1}{a^3}(a^3\dot{\varphi})^{\cdot} + \frac{1}{a^2}\triangle_\gamma\varphi \,.$$

The field equation (172) becomes

$$\boxed{\ddot{\varphi} + 3H\dot{\varphi} - \frac{1}{a^2}\triangle_\gamma\varphi = -V_{,\varphi}(\varphi)\,.} \tag{175}$$

Note that the expansion of the Universe induces a 'friction' term. In this basic equation one also sees the appearance of the Hubble length. From (174) we obtain for the energy density and the pressure of the scalar field

$$\rho_\varphi = T_{00} = \frac{1}{2}\dot{\varphi}^2 + V + \frac{1}{2a^2}(\nabla\varphi)^2\,, \tag{176}$$

$$p_\varphi = \frac{1}{3}T^i{}_i = \frac{1}{2}\dot{\varphi}^2 - V - \frac{1}{6a^2}(\nabla\varphi)^2\,. \tag{177}$$

(Here, $(\nabla\varphi)^2$ denotes the squared gradient on the 3-space (Σ, γ).)

Suppose the gradient terms can be neglected, and that φ is during a certain phase slowly varying in time, then we get

$$\rho_\varphi \approx V, \quad p_\varphi \approx -V\,. \tag{178}$$

Thus $p_\varphi \approx -\rho_\varphi$, as for a cosmological term.

Let us ignore for the time being the spatial inhomogeneities in the previous equations. Then these reduce to

$$\ddot{\varphi} + 3H\dot{\varphi} + V_{,\varphi}(\varphi) = 0\,; \tag{179}$$

$$\rho_\varphi = \frac{1}{2}\dot{\varphi}^2 + V, \quad p_\varphi = \frac{1}{2}\dot{\varphi}^2 - V\,. \tag{180}$$

Beside (179) the other dynamical equation is the Friedmann equation

$$\boxed{H^2 + \frac{K}{a^2} = \frac{8\pi}{3M_{Pl}^2}\left[\frac{1}{2}\dot{\varphi}^2 + V(\varphi)\right]\,.} \tag{181}$$

Equations (179) and (181) define a non-linear dynamical system for the dynamical variables $a(t), \varphi(t)$, which can be studied in detail (see, e.g., [113]).

Let us ignore the curvature term K/a^2 in (181). Differentiating this equation and using (179) shows that

$$\dot{H} = -\frac{4\pi}{M_{Pl}^2}\dot{\varphi}^2\,. \tag{182}$$

Regard H as a function of φ, then

$$\frac{dH}{d\varphi} = -\frac{4\pi}{M_{Pl}^2}\dot{\varphi}\,. \tag{183}$$

This allows us to write the Friedmann equation as

$$\left(\frac{dH}{d\varphi}\right)^2 - \frac{12\pi}{M_{Pl}^2} H^2(\varphi) = -\frac{32\pi^2}{M_{Pl}^4} V(\varphi) \; . \tag{184}$$

For a given potential $V(\varphi)$ this is a differential equation for $H(\varphi)$. Once this function is known, we obtain $\varphi(t)$ from (183) and $a(t)$ from (182).

C.4 Power-Law Inflation

We now proceed in the reverse order, assuming that $a(t)$ follows a power law

$$a(t) = \text{const.}\ t^p \; . \tag{185}$$

Then $H = p/t$, so by (182)

$$\dot{\varphi} = \sqrt{\frac{p}{4\pi}} M_{Pl} \frac{1}{t}, \quad \varphi(t) = \sqrt{\frac{p}{4\pi}} M_{Pl} \ln(t) + \text{const.} \; ,$$

hence

$$H \propto \exp\left(-\sqrt{\frac{4\pi}{p}} \frac{\varphi}{M_{Pl}}\right) \; . \tag{186}$$

Using this in (184) leads to an exponential potential

$$V(\varphi) = V_0 \exp\left(-4\sqrt{\frac{\pi}{p}} \frac{\varphi}{M_{Pl}}\right) \; . \tag{187}$$

C.5 Slow-Roll Approximation

An important class of solutions is obtained in the slow-roll approximation (SLA), in which the basic (179) and (181) can be replaced by

$$H^2 = \frac{8\pi}{3M_{Pl}^2} V(\varphi) \; , \tag{188}$$

$$3H\dot{\varphi} = -V_{,\varphi} \; . \tag{189}$$

A necessary condition for their validity is that the *slow-roll parameters*

$$\varepsilon_V(\varphi) := \frac{M_{Pl}^2}{16\pi} \left(\frac{V_{,\varphi}}{V}\right)^2 \; , \tag{190}$$

$$\eta_V(\varphi) := \frac{M_{Pl}^2}{8\pi} \frac{V_{,\varphi\varphi}}{V} \tag{191}$$

are small:

$$\varepsilon_V \ll 1, \qquad | \, \eta_V \, | \ll 1 \; . \tag{192}$$

These conditions, which guarantee that the potential is flat, are, however, not sufficient.

The simplified system (188) and (189) implies

$$\boxed{\dot{\varphi}^2 = \frac{M_{Pl}^2}{24\pi}\frac{1}{V}\left(V_{,\varphi}\right)^2 \ .} \tag{193}$$

This is a differential equation for $\varphi(t)$.

Let us consider potentials of the form

$$V(\varphi) = \frac{\lambda}{n}\varphi^n \ . \tag{194}$$

Then (193) becomes

$$\boxed{\dot{\varphi}^2 = \frac{n^2 M_{Pl}^2}{24\pi}\frac{1}{\varphi^2}V \ .} \tag{195}$$

Hence, (188) implies

$$\frac{\dot{a}}{a} = -\frac{4\pi}{nM_{Pl}^2}\left(\varphi^2\right)^{\cdot} \ ,$$

and so

$$\boxed{a(t) = a_0 \exp\left[\frac{4\pi}{nM_{Pl}^2}\left(\varphi_0^2 - \varphi^2(t)\right)\right] \ .} \tag{196}$$

We see from (195) that $\frac{1}{2}\dot{\varphi}^2 \ll V(\varphi)$ for

$$\varphi \gg \frac{n}{4\sqrt{3\pi}}M_{Pl} \ . \tag{197}$$

Consider first the example $n = 4$. Then (195) implies

$$\frac{\dot{\varphi}}{\varphi} = \sqrt{\frac{\lambda}{6\pi}}M_{Pl} \ \Rightarrow \varphi(t) = \varphi_0 \exp\left(-\sqrt{\frac{\lambda}{6\pi}}M_{Pl}\ t\right) \ . \tag{198}$$

For $n \neq 4$:

$$\varphi(t)^{2-n/2} = \varphi_0^{2-n/2} + t\left(2 - \frac{n}{2}\right)\sqrt{\frac{n\lambda}{24\pi}}M_{Pl}^{3-n/2} \ . \tag{199}$$

For the special case $n = 2$ this gives, using the notation $V = \frac{1}{2}m^2\varphi^2$, the simple result

$$\varphi(t) = \varphi_0 - \frac{mM_{Pl}}{2\sqrt{3\pi}}t \ . \tag{200}$$

Inserting this into (196) provides the time dependence of $a(t)$.

C.6 Why Did Inflation Start?

Attempts to answer this and related questions are *very speculative* indeed. A reasonable direction is to imagine random initial conditions and try to understand how inflation can emerge, perhaps generically, from such a state of matter. A. Linde first discussed a scenario along these lines which he called *chaotic inflation*. In the context of a single scalar field model he argued that typical initial conditions correspond to $\frac{1}{2}\dot{\varphi}^2 \sim \frac{1}{2}(\partial_i\varphi)^2 \sim V(\varphi) \sim 1$ (in Planckian units). The chance that the potential energy dominates in some domain of size $> \mathcal{O}(1)$ is presumably not very small. In this situation inflation could begin and $V(\varphi)$ would rapidly become even more dominant, which ensures continuation of inflation. Linde concluded from such considerations that chaotic inflation occurs under rather natural initial conditions. For this to happen, the form of the potential $V(\varphi)$ can even be a simple power law of the form (194). Many questions remain, however, open.

The chaotic inflationary Universe will look on very large scales – much larger than the present Hubble radius – extremely inhomogeneous. For a review of this scenario I refer to [114]. A much more extended discussion of inflationary models, including references, can be found in [107].

C.7 Inflation and Primordial Power Spectra

For a detailed derivation of the primordial power spectra that are generated as a result of quantum fluctuations during an inflationary period, I refer to my Combo-lectures [63].

The main steps are quite straightforward. First, one studies classical perturbations of the scalar field and the metric. For the scalar field one can reduce the problem to a Klein–Gordon equation with a time-dependent mass for a suitable gauge invariant perturbation amplitude. The quantization of this field follows standard rules. The quantization of the scalar part of the metric (Bardeen potentials) is then also fixed. Of particular interest is the power spectrum, $P_{\mathcal{R}}(k)$, of the so-called "curvature perturbation amplitude" $\mathcal{R}$. This is proportional to the Fourier transform of the two-point correlation function. More precisely, if

$$\mathcal{R}(\eta, \mathbf{x}) = (2\pi)^{-3/2} \int \mathcal{R}_{\mathbf{k}}(\eta) e^{i\mathbf{k}\cdot\mathbf{x}} d^3k \ ,$$

then

$$\langle 0|\mathcal{R}_{\mathbf{k}}\mathcal{R}^{\dagger}_{\mathbf{k}'}|0\rangle =: \frac{2\pi^2}{k^3} P_{\mathcal{R}}(k)\delta^{(3)}(\mathbf{k}-\mathbf{k}') \ .$$

In the slow-roll approximation, this can be worked out explicitly, with the result

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{4}{M_{Pl}^4} \left. \frac{H^4}{(dH/d\varphi)^2} \right|_{k=aH} \tag{201}$$

$$\simeq \frac{128\pi}{3} \frac{1}{M_{Pl}^6} \left. \frac{U^3}{(U_{,\varphi})^2} \right|_{k=aH} \ . \tag{202}$$

The expression on the right is evaluated at *horizon crossing* $k = aH$.

It is even simpler to determine the *power spectrum of gravitational waves* (tensor modes). In the same approximation one finds

$$\mathcal{P}_g(k) = \frac{16}{\pi} \left.\frac{H^2}{M_{Pl}^2}\right|_{k=aH}, \quad H^2 \simeq \frac{8\pi}{3M_{Pl}^2} U \ . \tag{203}$$

For a given inflationary model, the power spectra are uniquely determined.

There is one delicate question, namely why we have chosen in the definition of the power spectrum the Fock state, relative to modes that at very short distances ($k/aH \to \infty$) approach the plane waves of the gravity free case with positive frequences. A priori, the initial state could contain all kinds of excitations. These would, however, be redshifted away by the enormous inflationary expansion, and the final power spectrum on interesting scales, much larger than the Hubble length, should be largely independent of possible initial excitations.

For a comparison with observations the power index, n_s, for scalar perturbations, defined by

$$n_s - 1 := \frac{d \ln P_{\mathcal{R}(k)}}{d \ln k} \tag{204}$$

is of particular interest. In terms of the slow-roll parameters (190) and (191) it is given by

$$n_s - 1 = -6\varepsilon_U + 2\eta_U \ , \tag{205}$$

whence the spectrum is nearly scale-free. For the ratio r of the amplitudes of P_g and $P_{\mathcal{R}}$ one finds $r = 16\varepsilon_U$. The WMAP data match the basic inflationary predictions, and are even well fit by the simplest model $U \propto \varphi^2$.

D Quintessence Models

In quintessence models the exotic missing energy with negative pressure is again described by a scalar field, whose potential is chosen such that the energy density of the homogeneous scalar field adjusts itself to be comparable to the matter density today for quite generic initial conditions, and is dominated by the potential energy. This ensures that the pressure becomes sufficiently negative. It is not simple to implement this general idea such that the model is phenomenologically viable. For instance, the success of BBN should not be spoiled. CMB and large-scale structure impose other constraints. One also would like to understand why cosmological acceleration started at about $z \sim 1$, and not much earlier or in the far future. There have been attempts to connect this with some characteristic events in the post-recombination Universe. On a fundamental level, the origin of a quintessence field that must be extremely weakly coupled to ordinary matter remains in the dark.

Let me briefly describe a simple model of this kind [115]. For the dynamics of the scalar field ϕ we adopt an exponential potential

$$V = V_0 \, e^{-\lambda\phi/M_P} \ .$$

Such potentials often arise in Kaluza–Klein and string theories. Matter is described by a fluid with a baryotropic equation of state: $p_f = (\gamma - 1)\rho_f$.

For a Friedmann model with zero space-curvature, one can cast the basic equations into an autonomous two-dimensional dynamical system for the quantities

$$x(\tau) = \frac{\kappa\dot{\phi}}{\sqrt{6}H}, \quad y(\tau) = \frac{\kappa\sqrt{V}}{\sqrt{3}H} \ ,$$

where

$$H = \dot{a}/a, \quad \tau = \log a, \quad \kappa^2 = 8\pi G$$

($a(t)$ is the scalar factor). This system of autonomous differential equations has the form

$$\frac{dx}{d\tau} = f(x, y; \lambda, \gamma), \quad \frac{dy}{d\tau} = g(x, y; \lambda, \gamma) \ ,$$

where f and g are polynomials in x and y of third degree, which depend parametrically on λ and γ. The density parameters Ω_ϕ and Ω_f for the field ϕ and the fluid are given by

$$\Omega_\phi = x^2 + y^2, \quad \Omega_\phi + \Omega_f = 1 \ .$$

The interesting fact is that, for a large domain of the parameters λ, γ, the phase portrait has qualitatively the shape of Fig. 16. Therefore, under generic

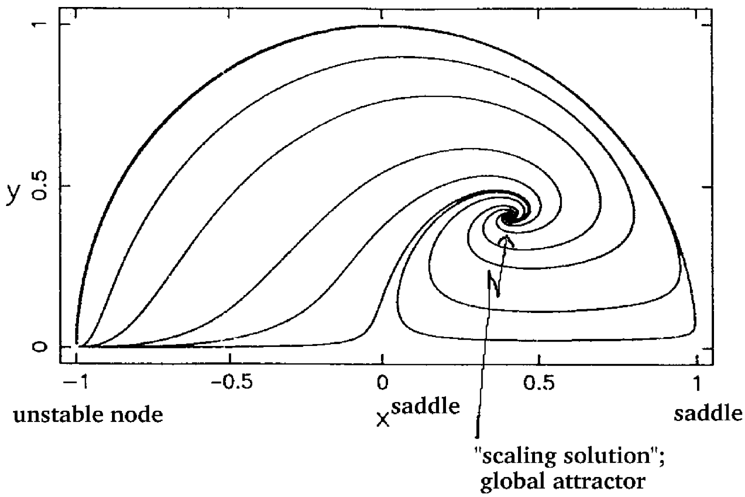

Fig. 16. Phase plane for $\gamma = 1$, $\lambda = 3$. The late-time attractor is the scaling solution with $x = y = 1/\sqrt{6}$ (from [115])

initial conditions, there is a global attractor (a node or a spiral) for which $\Omega_\phi = 3\gamma/\lambda^2$. For this "scaling solution" Ω_ϕ/Ω_f remains fixed, and for any other solution this ration is finally approached.

Unfortunately, if we set $p_\phi = (\gamma_\phi - 1)$ we find that $\gamma_\phi = 2x^2/(x^2+y^2)$, and this is equal to γ for the scaling solution. Thus this does *not* correspond to a quintessence solution. Moreover, the condition that ρ_ϕ should be subdominant during nucleosynthesis implies a small value for Ω_ϕ.

A more successful example of a so-called "tracker potential", with the property that the scalar field approaches a common evolutionary path from a wide range of initial conditions, has the form of an inverse power law, $V(\phi) = V_0/\phi^\alpha$ [117]. There is an extended literature on the subject. References [116]–[121] give a small selection of important early papers. For a recent review that describes also other scalar field models, see [122]. I emphasize once more that on the basis of the vacuum energy problem we would expect a huge additive constant for the quintessence potential that would destroy the whole picture. Thus, assuming for instance that the potential approaches zero as the scalar field goes to infinity has (so far) no basis. Apart of this and other fine tuning problems, I doubt that this kind of phenomenological models – with no natural field theoretical justification – will lead to an understanding of dark energy at a deeper level.

Fortunately, future more precise observations will allow us to decide whether the presently dominating exotic energy density satisfies $p/\rho = -1$ or whether this ratio is somewhere between -1 and $-1/3$. Recent studies (see [71, 72], and references therein), which make use of existing cosmological data, do not yet support quintessence. The restrictions for a possible redshift dependence are, so far, rather weak.

If convincing evidence for such a dependence should be established, we will not be able to predict the distant future of the Universe. Eventually, the dark energy density may perhaps become negative. This illustrates that we may *never be able to predict* the asymptotic behavior of the most grandiose of all dynamical systems. Other conclusions are left to the reader.

References

1. N. Straumann, *General Relativity, With Applications to Astrophysics*, Texts and Monographs in Physics, Springer Verlag, 2004.
2. G.E. Lemaître, in *Albert Einstein: Philosopher-Scientist*, P.A. Schilpp, ed., Illinois: The Library of Living Philosophers (1949).
3. N. Straumann, *On the Cosmological Constant Problems and the Astronomical Evidence for a Homogeneous Energy Density with Negative Pressure*, in *Poincarè Seminar 2002, Vacuum Energy – Renormalization*, B. Duplantier, and V. Rivasseau, eds.; Birkhäuser-Verlag 2003, pp. 7–51; astro-ph/0203330.
4. N. Straumann, *The History of the Cosmological Constant Problem*, in *On the Nature of Dark Energy*, IAP Astrophysics Colloquium 2002, Frontier Group, 2003, p. 17; gr-qc/0208027.

5. A. Einstein, *Sitzungsber. Preuss. Akad. Wiss.* phys.-math. Klasse VI, 142 (1917). See also: [6], Vol. 6, p. 540, Doc. 43.
6. A. Einstein, *The Collected Papers of Albert Einstein*, Vols. 1–9, Princeton University Press, 1987–. See also: [http:www. einstein.caltech.edu/].
7. A. Einstein, *On the Foundations of the General Theory of Relativity.* Reference [4], Vol. 7, Doc. 4.
8. A. Pais, *'Subtle is the Lord...': The Science and the Life of Albert Einstein.* Oxford University Press (1982). See especially Sect. 15e.
9. W. de Sitter, Proc. Acad. Sci., **19**, 1217 (1917); and **20**, 229 (1917).
10. A.S. Eddington, *The Mathematical Theory of Relativity.*Chelsea Publishing Company (1924). Third (unaltered) Edition (1975). See especially Sect. 70.
11. Letter from Hermann Weyl to Felix Klein, 7 February 1919; see also [5], Vol. 8, Part B, Doc. 567.
12. H. Weyl, Phys. Zeits. **24**, 230, (1923); Phil. Mag. **9**, 923 (1930).
13. C. Lanczos, Phys. Zeits. **23**, 539 (1922).
14. C. Lanczos, Zeits. f. Physik **17**, 168 (1923).
15. A. Friedmann, Z.Phys. **10**, 377 (1922); **21**, 326 (1924).
16. G.E. Lemaître, Ann. Soc. Sci. Brux. A **47**, 49 (1927).
17. G.E. Lemaître, Monthly Not. Roy. Astron. Soc. **91**, 483 (1931).
18. A. Einstein, S.B. Preuss. Akad. Wiss. (1931), 235.
19. A. Einstein, Appendix to the 2nd edn. of *The Meaning of Relativity*, (1945); reprinted in all later editions.
20. W. Pauli, *Theory of Relativity.* Pergamon Press (1958); Supplementary Note **19**.
21. O. Heckmann, *Theorien der Kosmologie*, berichtigter Nachdruck, Springer-Verlag (1968).
22. V. Petrosian, E.E. Salpeter, and P. Szekeres, Astrophys. J. **147**, 1222 (1967).
23. W. Pauli, *Pauli Lectures on Physics*; Ed. C.P. Enz. MIT Press (1973); Vol. 4, especially Sect. 20.
24. C.P. Enz, and A. Thellung, Helv. Phys. Acta **33**, 839 (1960).
25. W. Pauli, *Die allgemeinen Prinzipien der Wellenmechanik.* Handbuch der Physik, Vol. XXIV (1933). New edition by N. Straumann, Springer-Verlag (1990); see Appendix III, p. 202.
26. W. Heisenberg and H. Euler, Z. Phys. **38**, 714 (1936).
27. V.S. Weisskopf, Kongelige Danske Videnskabernes Selskab, Mathematisk-fysiske Meddelelser XIV, No.6 (1936).
28. M. Bordag, U. Mohideen, and V.M. Mostepanenko *New Developments in the Casimir Effect*, quant-ph/0106045.
29. C. Beck and M.C. Mackey, Phys. Lett. **B 605**, 295 (2005); astro-ph/0406504.
30. Ph. Jetzer and N. Straumann, Phys. Lett. **B 606**, 77 (2005); astro-ph/0411034.
31. C. Beck and M. C. Mackey, astro-ph/0603397.
32. Ph. Jetzer and N. Straumann, Phys. Lett. **B 639**, 57 (2006); [astro-ph/0604522].
33. Y.B. Zel'dovich, JETP letters **6**, 316 (1967); Soviet Physics Uspekhi **11**, 381 (1968).
34. C.G. Callan, S. Coleman, and R. Jackiw, Ann. Phys. **59**, 42 (1970).
35. T. Schäfer and E.V. Shuryak, Rev. Mod. Phys. **70**, 323 (1998).
36. R.D. Peccei and H. Quinn, Phys. Rev. Lett, **38**, 1440 (1977); Phys. Rev. **D16**, 1791 (1977).

37. E. Witten, hep-ph/0002297.
38. S.M. Caroll, *Living Reviews in Relativity*, astro-ph/0004075.
39. N. Straumann, in *Dark Matter in Astro- and Particle Physics*, Edited by H.V.Klapdor-Kleingrothaus, Springer (2001), p. 110.
40. S.E. Rugh and H. Zinkernagel, hep-th/0012253.
41. T. Padmanabhan, Phys. Repts. **380**, 235 (2003).
42. W. Baade, Astrophys. J. **88**, 285 (1938).
43. A. Sandage, G.A. Tammann, A. Saha, B. Reindl, F.D. Machetto and Panagia, astro-ph/0603647; A. Saha, F. Thim, G.A. Tammann, B. Reindl, A. Sandage, astro-ph/0602572.
44. G.A. Tammann. In *Astronomical Uses of the Space Telescope*, eds. F. Macchetto, F. Pacini and M. Tarenghi, p. 329. Garching: ESO.
45. S. Colgate, Astrophys. J. **232**,404 (1979).
46. SCP-Homepage: http://www-supernova.LBL.gov
47. HZT-Homepage: http://cfa-www.harvard.edu/cfa/oir/Research/supernova/HighZ.html
48. *Cosmic Explosions*, eds. S. Holt and W. Zhang, AIP Conference Proceedings 522, American Institute of Physics (New York) (2000).
49. W. Hillebrandt and J.C. Niemeyer, Ann. Rev. Astron. Astrophys. **38**, 191–230 (2000).
50. B. Leibundgut, Astron. Astrophys. **10**, 179 (2000).
51. S. Perlmutter, et al., Astrophys. J. **517**, 565 (1999).
52. B. Schmidt, et al., Astrophys. J. **507**, 46 (1998).
53. A.G. Riess, et al., Astron. J. **116**, 1009 (1998).
54. B. Leibundgut, Ann. Rev. Astron. Astrophys. **39**, 67 (2001).
55. A.V. Filippenko, *Measuring and Modeling the Universe*, ed. W.L. Freedman, Cambridge University Press, (2004); astro-ph/0307139.
56. A.G. Riess, et al., Astrophys. J. **607**, 665 (2004); astro-ph/0402512.
57. P. Astier, et al., Astron. Astrophys. **447**, 31 (2006) [astro-ph/0510447].
58. A. Clocchiatti, et al., Astron. Astrophys. **447**, 31 (2006); [astro-ph/0510155].
59. Snap-Homepage: http://snap.lbl.gov
60. R. Kantowski and R.C. Thomas, Astrophys. J. **561**, 491 (2001); astro-ph/0011176.
61. W. Hu and S. Dodelson, Annu. Rev. Astron. Astrophys. **40**, 171–216 (2002).
62. U. Seljak, and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996). (See also *http://www.sns.ias.edu/matiasz/CMBFAST/cmbfast.html*)
63. N. Straumann, *From primordial quantum fluctuations to the anisotropies of the cosmic microwave background radiation*, Ann. Phys. (Leipzig) **15**, No. 10–11, 701–847 (2006); [hep-ph/0505249].
64. W. Hu and M. White, Phys. Rev. D **56**, 596(1997).
65. W. Hu, U. Seljak, M. White, and M. Zaldarriaga, Phys. Rev. D **57**, 3290 (1998).
66. G. Steigman, Int.J.Mod.Phys. **E15**, 1 (2006); astro-ph/0511534.
67. C.L. Bennett, et al., ApJS **148**, 1 (2003); ApJS **148**, 97 (2003).
68. D.N. Spergel, et al., ApJS **148** 175 (2003).
69. D.N. Spergel, et al., astro-ph/0603449.
70. L. Page et al., astro-ph/0603450.
71. M. Tegmark et al., Phys. Rev. **D69**, 103501 (2004); astro-ph/0310723.
72. U. Seljak, et al., Phys. Rev. **D71**, 103515 (2005); astro-ph/0407372; astro-ph/0604335.

73. S. Cole, et al., MNRAS, **362**, 505 (2005).
74. A. Blanchard, M. Douspis, M. Rowan-Robinson, and S. Sarkar, Astron.Astrophys. **412**, 35 (2003); astro-ph/0304237.
75. A. Blanchard, M. Douspis, M. Rowan-Robinson, and S. Sarkar, astro-ph/0512085.
76. E. Barausse, S. Matarrese, and A. Riotto, Phys. Rev. **D71**,0635537 (2005); astro-ph/0501152.
77. E. W. Kolb, S. Matarrese, A. Notari, and A. Riotto, hep-th/0503117.
78. Ch. M. Hirata and U. Seljak, Phys. Rev. **D72**, 083501 (2005); astro-ph/0503582.
79. E. W. Kolb, S. Matarrese, and A. Riotto, astro-ph/0506534.
80. S. Räsänen, astro-ph/0607626.
81. M. Sasaki, Mon. Not. R. Astron. Soc. **228**, 653 (1987).
82. N. Sigiura, N. Sugiyama, and M. Sasaki, Prog.Theo. Phys. **101**, 903 (1999).
83. C. Bonvin, R. Durrer and M.A. Gasparini, Phys. Rev. **D73**, 023523 (2006); astro-ph/0511183.
84. K. Bolejko, astro-ph/0512103.
85. S. Nojiri and S.D. Odintsov, hep-th/0601213.
86. A.L. Erickcek, T.L. Smith, and M. Kamiokowski, astro-ph/0610483.
87. T. Chiba, Phys. Lett. **B575**, 1 (2003); astro-ph/0307338.
88. S. Caroll, A. De Felice, V. Duvvuri, D. Easson, M. Trodden and M. Turner, Phys. Rev. **D70**, 063513 (2005).
89. R.P. Woodard, astro-ph/0601672.
90. A. De Felice, M. Hindmarsh, and M. Trodden, astro-ph/0604154.
91. G. Calcagni, B. de Carlos and A. De Felice, hep-th/0604201.
92. G. Velo and D. Zwanziger, Phys. Rev. **186**, 1337–41 (1969) ; Phys. Rev. **188**, 2218–22 (1969).
93. T.P. Sotiriou, gr-qc/0604028.
94. T. Koivisto and H. Kurki-Suonio, astro-ph/0509422.
95. T. Koivisto, astro-ph/0602031.
96. M. Amarzguioui, O. Elgaroy, D.F. Mota and T. Multmaki, astro-ph/0510519.
97. G.R. Dvali, G. Gabadadze and M. Porrati, Phys. Lett. B **485**, 208 (2000); hep-th/0005016.
98. R. Maartens, *Living Reviews*, gr-qc/0312059.
99. R. Maartens and E. Majerotto, astro-ph/0603353.
100. M. Fairbairn and A. Goobar, astro-ph/0511029.
101. D. J. Eisenstein et al., *Astrophys. J.* **633**, 560 (2005).
102. D. Gorbunov, K. Koyama, and S. Sibiryakov, Phys. Rev. **D73**, 044016 (2006).
103. Ch. Charmousis, R. Gregory, N. Kaloper and A. Padilla, hep-th/0604086.
104. A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, hep-th/0602178.
105. P.J.E. Peebles, *Principles of Physical Cosmology.* Princeton University Press 1993.
106. J.A. Peacock, *Cosmological Physics.* Cambridge University Press 1999.
107. A.R. Liddle and D.H. Lyth, *Cosmological Inflation and Large Scale Structure.* Cambridge University Press 2000.
108. S. Dodelson, *Modern Cosmology.* Academic Press 2003.
109. G. Börner, *The Early Universe.* Springer-Verlag 2003 (4th edition).
110. N. Straumann, Helv. Phys. Acta **45**, 1089 (1972).

111. N. Straumann, *Allgemeine Relativitätstheorie und Relativistische Astrophysik*, 2. Auflage, Lecture Notes in Physics, Vol. 150, Springer-Verlag (1988); Kap. IX.
112. N. Straumann, *Kosmologie I*, Vorlesungsskript, http://web.unispital.ch/neurologie/vest/homepages/straumann/norbert
113. V.A. Belinsky, L.P. Grishchuk, I.M. Khalatnikov, and Ya.B. Zeldovich, Phys. Lett. **155B**, 232 (1985).
114. A. Linde, *Lectures on Inflationary Cosmology*, hep-th/9410082.
115. E.J. Copeland, A.R. Liddle, and D. Wands, Phys. Rev. **D57**, 4686 (1998).
116. C.Wetterich, Nucl. Phys. **B302**, 668 (1988).
117. B. Ratra and P.J.E. Peebles, Astrophys. J. Lett. **325**, L17 (1988); Phys. Rev. **D37**, 3406 (1988).
118. R.R. Caldwell, R. Dave and P.J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998).
119. P.J. Steinhardt, L. Wang and I. Zlatev, Phys. Rev. Lett. **82**, 896 (1999); Phys. Rev. **D59**, 123504 (1999).
120. C. Armendariz-Picon, V. Mukhanov and P.J. Steinhardt, astro-ph/0004134 and astro-ph/0006373.
121. P. Binétruy, Int. J. Theor. Phys. **39**, 1859 (2000); hep-ph/0005037.
122. E. J. Copeland, M. Sami, and S. Tsujikawa, hep-th/0603057.