Elements of String Cosmology

Maurizio Gasperini

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521868754

This page intentionally left blank

ELEMENTS OF STRING COSMOLOGY

The standard cosmological picture of our Universe emerging from a "big bang" leaves open many fundamental questions: Is the big bang a true physical singularity? What happens to the Universe at ultra-high energy densities when even gravity should be quantized? Has our cosmological history a finite or infinite past extension? Do we live in more than four space-time dimensions? String theory, a unified theory of all forces of nature, should be able to answer these questions.

This book contains a pedagogical introduction to the basic notions of string theory and cosmology. It describes the new possible scenarios suggested by string theory for the primordial evolution of our Universe. It discusses the main phenomenological consequences of these scenarios, stresses their differences from each other, and compares them with the more conventional models of inflation.

The first book dedicated to string cosmology, it summarizes over 15 years of research in this field and introduces current advances. The book is self-contained so it can be read by astrophysicists with no knowledge of string theory, and high-energy physicists with little understanding of cosmology. Detailed and explicit derivations of all the results presented provide a deeper appreciation of the subject.

Maurizio Gasperini is Professor of Theoretical Physics in the Physics Department of the University of Bari, Italy. He has authored several publications on gravitational theory, high-energy physics and cosmology, and has twice received one of the Awards for Essays on Gravitation from the Gravity Research Foundation (1996 and 1998).

ELEMENTS OF STRING COSMOLOGY

MAURIZIO GASPERINI

University of Bari, Italy

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521868754

First published in print format 2007

ISBN-13 978-0-511-33229-6 eBook (Adobe Reader)
ISBN-10 0-511-33229-7 eBook (Adobe Reader)

ISBN-13 978-0-521-86875-4 hardback
ISBN-10 0-521-86875-0 hardback

To Patty
with gratitude

Contents

Preface

The aim of this book is to provide an elementary, but detailed, introduction to the possible impact of string theory on the basic aspects of primordial cosmology. The content of the book includes a discussion of the new models of the Universe obtained by solving the string theory equations, as well as a systematic analysis of their phenomenological consequences, for a close comparison with more conventional inflationary scenarios based on the Einstein equations.

The book is primarily intended for graduate students, not necessarily equipped with a background knowledge of cosmology and string theory; but any reader in possession of the basic notions of general relativity and quantum field theory should be able to benefit from the use of this book (or, at least, of a great part of it). Some chapters (in particular, Chapters 1, 7 and 8) could also be used as a "soft" introduction to modern cosmology for string theorists, while other chapters (in particular, Chapters 2 and 3) as a soft introduction to string theory for astrophysicists; however, all readers are strongly advised to refer to other, more specialized books for a rigorous (independent) study of cosmology and string theory. It should be stressed, also, that this book *is not* aimed as a comprehensive and up-to-date review of all research work available in a string cosmology context: it only provides a pedagogic introduction to the basic ideas and theoretical tools, hopefully useful to the interested reader as a starting point towards more advanced research topics currently in progress in this field.

This book grew out of lectures given in May 2001 at the *First International Ph.D. Course on "Gravitational Physics and Astrophysics"*, jointly organized by the Universities of Berlin, Portsmouth, Potsdam, Salerno and Zurich. The style is that of class lectures: I have tried to be self-contained as much as possible, and I have not hesitated to insert many computational details and explanations, which may even appear to be trivial to the expert reader, but which may result in being of crucial importance for many students, as I have personally verified during the lectures. Besides organizing known material in a form appropriate to a pedagogic presentation, the book also presents explicit calculations never seen in the literature; in addition, it contains new results obtained through simple generalizations of previous studies. In particular, all topics are discussed (whenever possible) in the general context of a $(d+1)$-dimensional space-time manifold: known results in $d=3$ are thus extended (some of them for the first time) to a generic number d of spatial dimensions.

A possible objection concerning the explicit absence of exercises and problems can be preempted by noting that the main text of the various chapters is literally "filled" with *solved* exercises, in the sense that all computations are displayed in full details, including all the explicit passages required for a reader's easy understanding. In view of such a large "equation density" in all sections and appendices, the inclusion of additional exercises seemed to be inappropriate.

Another warning concerns the appendices. In contrast with the common use of presenting technical details and computations (and with the exception of Appendix 2A), here the appendices are devoted to a self-contained discussion of specific topics which are closely related to the subject of the chapter, but which are not essential for the understanding of other chapters, and can be skipped in a first reading.

It should be explained, finally, why some chapters are characterized by a list of references much longer than others. The reason is that in some cases (for instance in Chapters 2 and 3) one can conveniently refer to existing books, which provide an excellent discussion of the subject; in other cases (for instance in Chapters 7, 8 and 10), no such book is presently available, and one has to resort to a more detailed bibliography with explicit references to the original papers on the subject.

Acknowledgements

It is a pleasure to thank all my collaborators from almost fifteen years of fun with string cosmology. In alphabetical order, they are: Luca Amendola, Nicola Bonasia, Valerio Bozza, Ram Brustein, Alessandra Buonanno, Cyril Cartier, Marco Cavaglià, Eugenio Coccia, Edmund Copeland, Giuseppe De Risi, Ruth Durrer, Massimo Giovannini, Michele Maggiore, Jnan Maharana, Kris Meissner, Slava Mukhanov, Stefano Nicotri, Federico Piazza, Roberto Ricci, Mairi Sakellariadou, Norma Sanchez, Carlo Ungarelli and Gabriele Veneziano. I would like to thank, in particular, Massimo Giovannini and Gabriele Veneziano for the numberless hours of stimulating discussions during which some of the topics discussed in this book literally grew “out of nothing”.

Special thanks are due to Augusto Sagnotti, for important clarifying discussions; to Carlo Angelantonj, Carlo Ungarelli and my student Carla Coppola for their careful reading of some parts of the manuscript; to Giuseppe De Risi for his help in the preparation of the chapter on the branes; to my colleagues of the Physics Department of the University of Bari, and in particular to Leonardo Angelini, Antonio Marrone and Egidio Scrimieri who helped me improve some plots and figures. I also wish to thank all the staff of the Theory Unit at CERN for support and warm hospitality during the last stages of preparation of the final manuscript.

But, above all, I am greatly indebted to Gabriele Veneziano for his friendship and invaluable help and advice over all the years, for the generous sharing of his mastery of string theory, and for his crucial role in encouraging me to concentrate on the study of string cosmology scenarios. It is fair to say that this book would not exist without his basic input.

Finally, I would like to thank Simon Capelin (Publishing Director at Cambridge) for his kind advice and encouragement, Lindsay Barnes (Assistant Editor) for her prompt assistance, Margaret Patterson (Copy Editor) for her very careful correction of the original manuscript, and Jeanette Alfoldi (Production Editor) for her kind and efficient help during the final stages of the publishing process.

Notation, units and conventions

Unless otherwise stated, we adopt the following conventions:

– spatial indices: $i, j, k, \ldots = 1, \ldots, d$;
– space-time indices: $\mu, \nu, \alpha, \ldots = 0, 1, \ldots, d$;
– metric signature: $g_{\mu\nu} = \mathrm{diag}(+, -, -, -, \cdots)$;
– Riemann tensor: $R_{\mu\nu\alpha}{}^{\beta} = \partial_\mu \Gamma_{\nu\alpha}{}^{\beta} + \Gamma_{\mu\rho}{}^{\beta}\Gamma_{\nu\alpha}{}^{\rho} - (\mu \leftrightarrow \nu)$;
– Ricci tensor: $R_{\nu\alpha} = R_{\mu\nu\alpha}{}^{\mu}$;
– covariant derivatives: $\nabla_\mu V^\alpha = \partial_\mu V^\alpha + \Gamma_{\mu\beta}{}^{\alpha} V^\beta$; $\quad \nabla_\mu V_\alpha = \partial_\mu V_\alpha - \Gamma_{\mu\alpha}{}^{\beta} V_\beta$.

Covariant objects are referred to the symmetric, metric-compatible Christoffel connection,

$$\Gamma_{\alpha\beta}{}^{\mu} = \frac{1}{2} g^{\mu\nu} \left(\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\alpha\nu} - \partial_\nu g_{\alpha\beta} \right), \tag{1}$$

satisfying $\nabla_\alpha g_{\mu\nu} = 0$. We use natural units $\hbar = c = k_{\mathrm{B}} = 1$, where k_{B} is the Boltzmann constant. The fundamental string mass, M_{s}, and string length, λ_{s}, are thus related to the string tension $T = (2\pi\alpha')^{-1}$ by

$$M_{\mathrm{s}}^2 = \lambda_{\mathrm{s}}^{-2} = (2\pi\alpha')^{-1}. \tag{2}$$

The four-dimensional (reduced) Planck mass M_{P}, and the Planck length λ_{P}, are related to the Newton constant G (in $d = 3$ spatial dimensions) by

$$M_{\mathrm{P}}^2 = \lambda_{\mathrm{P}}^{-2} = (8\pi G)^{-1}. \tag{3}$$

The current experimental value $G \simeq 6.709 \times 10^{-39}\,\mathrm{GeV}^{-2}$ [1] then leads to

$$M_{\mathrm{P}} = (8\pi G)^{-1/2} \simeq 2.43 \times 10^{18}\,\mathrm{GeV} \tag{4}$$

(note the difference from an alternative – often used – definition, $M_{\mathrm{P}} = G^{-1/2} \simeq 1.22 \times 10^{19}\,\mathrm{GeV}$). In a manifold with $D = d + 1$ space-time dimensions Eq. (3) becomes

$$M_{\mathrm{P}}^{d-1} = \lambda_{\mathrm{P}}^{1-d} = (8\pi G_D)^{-1}, \tag{5}$$

where G_D is the D-dimensional gravitational coupling constant, and $M_{\rm P}$, $\lambda_{\rm P}$ are gravitational scales, possibly different (in principle) from the numerical value (4) determined by four-dimensional phenomenology. If the geometry of the higher-dimensional manifold has a factorized, Kaluza–Klein structure, then G_D is related to the four-dimensional Newton constant G through the proper volume of the internal space $\mathcal{M}_{d-3}$ as follows:

$$(8\pi G_D)^{-1} V_{d-3} = (8\pi G)^{-1}, \qquad V_{d-3} = \int_{\mathcal{M}_{d-3}} \sqrt{|g|}\, \mathrm{d}^{d-3}x. \tag{6}$$

The relative strength of $M_{\rm s}$ and $M_{\rm P}$ is controlled by the scalar dilaton field ϕ, defined in such a way that, at the tree-level, and in d spatial dimensions,

$$\left(\frac{M_{\rm s}}{M_{\rm P}}\right)^{d-1} = \mathrm{e}^{\phi}. \tag{7}$$

Masses, energies and temperatures are usually expressed in eV (or multiples of eV), and distances in cm (or eV^{-1}), using the equivalence relations:

$$(1\,\mathrm{eV})^{-1} \simeq 1.97 \times 10^{-5}\ \mathrm{cm} \simeq 6.59 \times 10^{-16}\,\mathrm{s} \simeq 8.6 \times 10^{-5}\,\mathrm{kelvin}^{-1}. \tag{8}$$

The Planck length, defined as in Eq. (3), corresponds to

$$\lambda_{\rm P} = (8\pi G)^{-1/2} \simeq 8.1 \times 10^{-33}\ \mathrm{cm}. \tag{9}$$

The curvature scale of the cosmological manifolds, parametrized by the Hubble parameter H, is often expressed in Planck units, and the energy densities in units of critical density $\rho_{\rm c} = 3H^2/8\pi G$. For the present Universe, in particular,

$$H_0 = 3.2h \times 10^{-18}\,\mathrm{s}^{-1} \simeq 8.7h \times 10^{-61} M_{\rm P}, \tag{10}$$

where $h = H_0/(100\ \mathrm{km\ s^{-1}\,Mpc^{-1}})$. Recent observations suggest

$$h = 0.73^{+0.04}_{-0.03} \tag{11}$$

as the current standard [1]. The corresponding critical density is

$$\rho_{\rm c}(t_0) = \frac{3H_0^2}{8\pi G} = 3H_0^2 M_{\rm P}^2 \simeq 1.88h^2 \times 10^{-29}\,\mathrm{g\ cm^{-3}} \simeq 2.25h^2 \times 10^{-120} M_{\rm P}^4. \tag{12}$$

Reference

[1] Particle Data Group webpage at pdg.lbl.gov/

1

A short review of standard and inflationary cosmology

In this chapter we will recall some basic notions of standard and inflationary cosmology that will be used later, in a string cosmology context. We will assume that the reader is already familiar with the geometric formalism of the theory of general relativity, and with the main observational aspects of large-scale astronomy and astrophysics. We will discuss, in particular, the various assumptions of the so-called standard cosmological model, the problems associated with its initial conditions, and the basic aspects of its "inflationary" completion driven by the potential energy of a cosmic scalar field (further details on the inflationary scenario will be supplied in Chapter 8). This presentation aims at a self-contained study of the early cosmological dynamics: for a more detailed introduction, and a deeper analysis of the topics discussed in this chapter, we refer the interested reader to [1, 2, 3] for the standard cosmological model, and to [4, 5, 6] for the inflationary scenario.

1.1 The standard cosmological model

The standard cosmological model, developed during the second half of the last century, was inspired by two fundamental observational results: the recession of galaxies, discovered by Hubble [7], and the presence of the Cosmic Microwave Background (CMB), discovered by Penzias and Wilson [8]. The model relies upon a number of hypotheses – also motivated by direct and indirect observations – that we now list, with some illustrative discussion.

1.1.1 Einstein equations

The first assumption is that the gravitational interaction, on cosmological scales of distance, is well described by the classical theory of general relativity, and in particular by the equations derived from the effective four-dimensional action

$$S = -\frac{1}{16\pi G}\int \mathrm{d}^4x\sqrt{-g}R + S_\Sigma + \int \mathrm{d}^4x\sqrt{-g}\,\mathcal{L}_m. \tag{1.1}$$

Here S_Σ is the Gibbons–Hawking boundary term [9], required in order to reproduce the standard Einstein equations, and $\mathcal{L}_m$ is the Lagrangian density of the matter fields, acting as gravitational sources. The variation of the action (1.1) with respect to the metric $g_{\mu\nu}$ yields (see Chapter 2 for an explicit derivation)

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G\, T_{\mu\nu}, \tag{1.2}$$

where $G_{\mu\nu}$ is the so-called Einstein tensor, and $T_{\mu\nu}$ is the (dynamical) energy-momentum tensor of the matter sources, defined by the variation (or functional differentiation) of the matter action as

$$\delta_g\left(\sqrt{-g}\,\mathcal{L}_m\right) = \frac{1}{2}\sqrt{-g}\,T_{\mu\nu}\,\delta g^{\mu\nu}. \tag{1.3}$$

The right-hand side of Eq. (1.2) represents all the sources gravitationally coupled to the metric, and therefore includes the possible contribution of the vacuum energy density associated with a cosmological constant Λ, and described by the effective energy-momentum tensor $T_{\mu\nu} = \Lambda g_{\mu\nu}$.

1.1.2 Homogeneity and isotropy

A second assumption is that the spatial sections of the Universe, on large enough scales of distance, can be described as homogeneous and isotropic (three-dimensional) Riemann manifolds, geometrically represented by maximally symmetric spaces where rotations and translations form a six-parameter isometry group.

It may be noted that, on scales much smaller than the Hubble radius $H_0^{-1} \simeq 0.9h^{-1} \times 10^{28}$ cm, the distribution of visible matter seems to follow a "fractal" distribution (see for instance [10]), and that it is not very clear, at present, at which scale the (averaged) matter distribution becomes really homogeneous and isotropic. The hypothesis of homogeneity and isotropy refers, however, to the full set of cosmic gravitational sources (including, as we shall see, radiation, dark matter, dark energy, ...), and is quite powerful, since it allows a simplified cosmological description in which the space-time geometry can be parametrized by the so-called "comoving" chart (or set of coordinates). In that case, the fundamental space-time interval reduces to

$$\mathrm{d}s^2 = b^2(t)\mathrm{d}t^2 - a^2(t)\mathrm{d}\sigma^2(\vec{r}), \tag{1.4}$$

where $a(t)$, $b(t)$ are generic functions of the time coordinate, and $d\sigma^2$ is the line-element of a three-dimensional space with constant (positive, negative or zero) curvature K. Using a set of stereographic coordinates $\{x_1, x_2, x_3\}$, the metric of such a maximally symmetric space can be parametrized as [1]

$$d\sigma^2 = dx_i\, dx^i + K \frac{(x_i\, dx^i)^2}{1 - K x_i x^i}, \tag{1.5}$$

where scalar products are performed with the Euclidean metric δ_{ij}.

An important property of the comoving chart is the fact that *static* observers, with four-velocity $u^\mu = (u^0, \vec{0})$, are also *geodesic* observers. The normalization condition $g_{\mu\nu} u^\mu u^\nu = 1$, with the metric (1.4), gives indeed $u^0 = b^{-1}(t)$ and

$$\frac{du^0}{d\tau} = -\frac{\dot{b}}{b^3}, \qquad {\Gamma_{00}}^0 (u^0)^2 = \frac{\dot{b}}{b^3}, \tag{1.6}$$

which implies that the field u^0 satisfies the geodesic equation

$$\frac{du^0}{d\tau} + {\Gamma_{00}}^0 (u^0)^2 = 0. \tag{1.7}$$

Here τ is the proper time (related to the coordinate time t by $d\tau = \sqrt{g_{00}}\, dt = b(t) dt$), and the dot denotes differentiation with respect to t. In addition, if $u^i = 0$, then

$$\frac{du^i}{d\tau} = -\Gamma^i_{00} (u^0)^2 = -\frac{1}{2b^2} g^{ij} \left(2\partial_0 g_{j0} - \partial_j g_{00}\right) \equiv 0. \tag{1.8}$$

Thus, in the absence of non-gravitational forces, static observers are always at rest with respect to comoving coordinates, even if the geometry is time dependent.

The existence of such observers provides a natural reference frame for synchronizing clocks, and suggests the use of a convenient time coordinate, the so-called *cosmic time*, which corresponds to the proper time of the static observers. The choice of this time coordinate leads to the *synchronous gauge*, defined by the condition $g_{00} = 1$. It is also convenient to parametrize the maximally symmetric space of Eq. (1.5) with spherical coordinates $\{r, \theta, \varphi\}$. By setting $x_1 = r \sin\theta \cos\varphi$, $x_2 = r \sin\theta \sin\varphi$, $x_3 = r\cos\theta$, and differentiating to compute $d\sigma^2$, in the synchronous gauge of the comoving chart, one finally arrives at the well-known Robertson–Walker metric, defined by

$$ds^2 = dt^2 - a^2(t) \left[\frac{dr^2}{1 - Kr^2} + r^2 (d\theta^2 + \sin^2\theta\, d\varphi^2)\right]. \tag{1.9}$$

Here t is the cosmic-time coordinate, and the constant K (with dimensions L^{-2}) controls the intrinsic curvature of the space-like $t = \text{const}$ hypersurfaces, representing three-dimensional sections of the space-time manifold. With our conventions the function $a(t)$, called the "scale factor", is dimensionless, while the comoving radial coordinate r has conventional dimensions of length.

Another choice of time coordinate (often used in this book) is the so-called *conformal gauge*, defined by the condition $g_{00} = a^2$. The time parameter of this gauge, usually denoted by η, is thus related to the cosmic time t by $\mathrm{d}t = a\,\mathrm{d}\eta$. The choice of the conformal gauge is particularly convenient for spatially flat manifolds ($K = 0$), whose metric can then be written in conformally flat form, using cartesian coordinates, as

$$\mathrm{d}s^2 = a^2(\eta)\left(\mathrm{d}\eta^2 - \mathrm{d}x_i\,\mathrm{d}x^i\right). \tag{1.10}$$

A space-time described by the Robertson–Walker metric is characterized by a number of interesting kinematical properties concerning the motion of test bodies and the propagation of signals (see for instance [1]). For the purposes of this book it will be enough to recall two effects.

The first effect concerns the spectral shift of a periodic signal, a shift originating from the well-known temporal slow-down produced by gravity. Indeed, at any given time t, all points of the three-dimensional spatial sections at constant curvature will be affected by exactly the same gravitational field, so that any local process will be equally slowed-down with respect to the same process occurring in the flat Minkowski space, quite independently of its spatial position. However, if the scale factor $a(t)$ varies with time, then the curvature radius of the spatial sections (and the associated intensity of the local effective gravitational field) will also vary with time. This will produce a difference in the local gravitational field (and in the local "slow-down") between the time t_{em} of emission of a periodic signal of pulsation ω_{em}, and the time $t_{\mathrm{obs}} > t_{\mathrm{em}}$ when the same signal is observed with pulsation ω_{obs}. The ratio of the two pulsations will be clearly proportional to the ratio of the local gravitational intensities at t_{em} and t_{obs}, and thus inversely proportional to the spatial curvature radius.

For a more precise computation of the spectral shift $\omega_{\mathrm{em}}/\omega_{\mathrm{obs}}$ we may consider a photon of four-momentum p^μ, traveling along a null geodesic of a spatially flat Robertson–Walker metric. In the cosmic-time gauge such a null path has differential equation $\mathrm{d}t = a\widehat{n}_i\,\mathrm{d}x^i$, where $\widehat{n}$ is a unit vector ($|\widehat{n}| = 1$) specifying the photon direction; the null photon momentum is, in this gauge, $p^\mu = p^0(1, \widehat{n}^{\,i}/a)$, with $g_{\mu\nu}p^\mu p^\nu = 0$. The momentum is parallelly transported along the geodesic, and for the energy p^0 we have, in particular,

$$\begin{aligned} \mathrm{d}p^0 &= -\Gamma_{\alpha\beta}{}^0\,\mathrm{d}x^\alpha p^\beta = \Gamma_{ij}^0\,\mathrm{d}x^i p^j \\ &= -\dot{a}p^0\widehat{n}_i\,\mathrm{d}x^i = -\frac{\dot{a}}{a}p^0\,\mathrm{d}t. \end{aligned} \tag{1.11}$$

The integration gives $p^0 = \overline{\omega}/a(t)$, where the integration constant $\overline{\omega}$ represents the proper frequency of the photon in the Minkowski space locally tangent to the given cosmological manifold.

The local frequency measured by a static, comoving observer u^μ is thus time dependent, being determined by the projection $p^\mu u_\mu = \overline{\omega}/a(t)$. A photon emitted at $t = t_{\rm em}$ and received at $t = t_{\rm obs}$, even in the absence of a (possible) Doppler effect due to the relative motion of source and emitter, will be characterized by the spectral shift

$$\frac{\omega_{\rm em}}{\omega_{\rm obs}} = \frac{(p^\mu u_\mu)_{\rm em}}{(p^\mu u_\mu)_{\rm obs}} = \frac{a_{\rm obs}}{a_{\rm em}} \tag{1.12}$$

(see also Eqs. (8.172)–(8.173), and the discussion of Section 8.2). If the Universe is expanding, then $a_{\rm obs} > a_{\rm em}$ for $t_{\rm obs} > t_{\rm em}$, and the Robertson–Walker metric produces an effective redshift of the signals received from distant sources, i.e. $\omega_{\rm obs} < \omega_{\rm em}$. In particular, since observations are carried out at the present time, $t_{\rm obs} = t_0$, it may be useful to introduce a redshift parameter $z(t)$ defined as

$$1 + z(t) = \frac{a(t_0)}{a(t)} \equiv \frac{a_0}{a(t)}, \tag{1.13}$$

which controls the relative "stretching" of the wavelengths of the received radiation,

$$z = \frac{\Delta\lambda}{\lambda} = \frac{\lambda_{\rm obs} - \lambda_{\rm em}}{\lambda_{\rm em}}. \tag{1.14}$$

A second important feature of the Robertson–Walker kinematics, which we recall here for later applications, is the possible existence of "horizons", i.e. of surfaces with relevant causal properties. For any given observer we may consider, in particular, the *particle horizon*, which divides the portion of space-time already observed from the one yet to be observed, and the *event horizon*, which divides the observable portion of space-time from the one causally disconnected [11]. For their precise definition we must refer to the limiting times $t_{\rm m}$ and $t_{\rm M}$ corresponding, respectively, to the maximum *past* extension and *future* extension of the time coordinate on the given cosmological manifold.

Let us consider a signal propagating towards the origin along a null radial geodesic of the metric (1.9) ($ds^2 = 0$, $d\theta = 0 = d\varphi$), satisfying the equation $dt/a = dr/\sqrt{1-Kr^2}$, and received by a comoving observer at rest at the origin of the polar coordinate system. A signal emitted from a radial position $r = r_1$, at a time $t = t_1$, will be received at $r = 0$ at a time $t = t_0 > t_1$, such that

$$\int_0^{r_1} \frac{dr}{\sqrt{1-Kr^2}} = \int_{t_1}^{t_0} \frac{dt}{a(t)}. \tag{1.15}$$

The considered signal was emitted at a *proper* distance $d(t)$ from the origin which, at time t_0, is determined by

$$d(t_0) = a(t_0) \int_0^{r_1} \frac{\mathrm{d}r}{\sqrt{1-Kr^2}} = a(t_0) \int_{t_1}^{t_0} \frac{\mathrm{d}t}{a(t)}. \tag{1.16}$$

In the limit $t_1 \to t_\mathrm{m}$ we then define the "**particle horizon**", for the given observer at time t_0, as the spherical surface centered at the origin $r = 0$ with proper radius

$$d_\mathrm{p}(t_0) = a(t_0) \int_{t_\mathrm{m}}^{t_0} \frac{\mathrm{d}t}{a(t)}. \tag{1.17}$$

This surface encloses the maximal portion of space physically accessible to direct observation from the origin of the coordinate system at the time t_0. Points located at a proper spatial distance $d > d_\mathrm{p}(t_0)$ *cannot* be causally connected with the given observer at the given time t_0 (they may become causally connected at later times, at least in principle).

Consider now a radial signal emitted towards the origin at time t_0, from a point located at a comoving position r_2, and received at the origin at a time $t_2 > t_0$. The proper distance of the emitter from the origin, at time t_0, is then

$$d(t_0) = a(t_0) \int_0^{r_2} \frac{\mathrm{d}r}{\sqrt{1-Kr^2}} = a(t_0) \int_{t_0}^{t_2} \frac{\mathrm{d}t}{a(t)}. \tag{1.18}$$

In the limit $t_2 \to t_\mathrm{M}$ we can then define the "**event horizon**", at the time t_0, as the spherical surface centered at the origin with proper radius

$$d_\mathrm{e}(t_0) = a(t_0) \int_{t_0}^{t_\mathrm{M}} \frac{\mathrm{d}t}{a(t)}. \tag{1.19}$$

Signals emitted from points located at a proper distance $d > d_\mathrm{e}(t_0)$ will *never* be able to reach the origin. In other words, points with spatial separations $d > d_\mathrm{e}$ will never become causally connected, even extending the time coordinate to the extremal future limit allowed by the given cosmological manifold.

The above horizons exist if the integrals of Eqs. (1.17) and (1.19) are convergent, of course. Consider, for instance, a cosmological solution describing a Universe expanding for ever from an initial singularity, and parametrized in cosmic time by the power-law scale factor $a(t) = t^\alpha$, with $\alpha > 0$, and $0 \le t \le \infty$: it can be easily checked that the particle horizon exists if $0 < \alpha < 1$, while the event horizon exists if $\alpha > 1$. For $\alpha = 1$ neither the particle horizon nor the event horizon exists. The definitions of horizon given here will be used in the following chapters, and will be applied in particular in Section 5.3 to illustrate some important differences between standard and string cosmology models of inflation.

1.1.3 Perfect fluid sources

A third assumption (or set of assumptions) of the standard cosmological model refers to the gravitational sources that we need to specify in order to solve the Einsten equations. According to the standard model the sources of the cosmological gravitational field on large scales, after averaging over possible spatial fluctuations, can be represented as a barotropic, perfect fluid with energy-momentum tensor

$$T_\mu^\nu = (\rho + p) u_\mu u^\nu - p\delta_\mu^\nu, \tag{1.20}$$

where the energy density ρ and pressure p depend only on time, and are related by the equation of state

$$\frac{p}{\rho} = \gamma = \text{const.} \tag{1.21}$$

In addition, the fluid is assumed to be at rest in the comoving frame. Thus, in the synchronous gauge, $u^\mu = (1, \vec{0})$ and T_μ^ν becomes diagonal,

$$T_0^0 = \rho(t), \qquad T_i^j = -p(t)\delta_i^j. \tag{1.22}$$

With the given sources we are now able to write explicitly the Einstein equations (1.2), using the following (more convenient, but equivalent) form:

$$R_\mu^\nu = 8\pi G \left(T_\mu^\nu - \frac{1}{2} T \delta_\mu^\nu \right). \tag{1.23}$$

For the Robertson–Walker metric (1.9) the non-zero components of the Ricci tensor, in mixed form, depend only on time, and are given by

$$\begin{aligned} R_1^1 = R_2^2 = R_3^3 &= -\frac{\ddot{a}}{a} - 2\left(H^2 + \frac{K}{a^2} \right), \\ R_0^0 &= -3\frac{\ddot{a}}{a}, \end{aligned} \tag{1.24}$$

where $H = \dot{a}/a$ (the dot indicates the derivative with respect to cosmic time). The time and spatial components of Eqs. (1.23) then provide, respectively, the following independent equations:

$$\begin{aligned} \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + 3p), \\ \frac{\ddot{a}}{a} + 2\left(H^2 + \frac{K}{a^2} \right) &= 4\pi G(\rho - p). \end{aligned} \tag{1.25}$$

Combining them in order to eliminate $\ddot{a}/a$, and differentiating the energy density ρ with respect to time, leads to the system of first-order differential equations:

$$H^2 + \frac{K}{a^2} = \frac{8\pi G}{3}\rho\,, \tag{1.26}$$

$$\dot{\rho} + 3H(\rho + p) = 0\,. \tag{1.27}$$

The last equation can also be directly obtained from the covariant conservation of the energy-momentum tensor, $\nabla_\nu T_\mu^\nu = 0$, which is a consequence of the contracted Bianchi identity $\nabla_\nu G_\mu^\nu = 0$ (see Eq. (1.2)).

In order to solve the above system of equations for the three unknown functions $a(t)$, $\rho(t)$, $p(t)$, it is necessary to use the equation of state $p = p(\rho)$, which in our case corresponds to the barotropic condition (1.21). In general, the gravitational sources of the standard cosmological model can be represented as a mixture of barotropic perfect fluids,

$$\rho = \sum_n \rho_n, \qquad p = \sum_n p_n, \qquad p_n = \gamma_n \rho_n, \tag{1.28}$$

with no energy transfer between the different fluid components, so that the energy-momentum tensor of each fluid is separately conserved. Equation (1.27) then yields, for each component,

$$\rho_n(t) = \rho_n(t_0)\left(\frac{a}{a_0}\right)^{-3(1+\gamma_n)}, \tag{1.29}$$

where $\rho_n(t_0)$ is an integration constant. Since the energy density of the different components has a different time behavior, the evolution of the Universe will then be characterized by different phases, each of them dominated by different fluid components.

In each cosmological phase the time evolution of the scale factor can be obtained by substituting Eq. (1.29) into (1.26), and solving the corresponding differential equation for $a(t)$. If, in particular, we are interested in the very early time evolution we can neglect the spatial curvature term (see below), and we obtain the scale factor

$$a_n(t) = \left(\frac{t}{t_0}\right)^{2/3(1+\gamma_n)}, \qquad \gamma_n \neq -1, \tag{1.30}$$

where t_0 is an integration constant. The case $\gamma_n = -1$ corresponds to the energy-momentum tensor of a cosmological constant

$$T_\mu^\nu = \Lambda \delta_\mu^\nu, \tag{1.31}$$

which describes an effective fluid with equation of state $p_n = -\rho_n = -\Lambda = \text{const}$ (see Eq. (1.22)). In this case Eq. (1.29) is still valid, and the integration of Eq. (1.26) (with $K = 0$) gives the exponential solution

$$a_n(t) = \exp[H(t - t_0)], \qquad H = \left(\frac{8\pi G\Lambda}{3}\right)^{1/2} = \text{const}. \tag{1.32}$$

The standard cosmological model, in its original formulation [1], assumes that the cosmic fluid consists of two fundamental components: incoherent matter (ρ_m) with zero pressure $p_m = 0$, and radiation (ρ_r) with pressure $p_r = \rho_r/3$. The radiation component of the cosmic fluid represents the contribution of all massless (or very light) relativistic particles (photons, gravitons, neutrinos, ...), while the pressureless matter component takes into account the large-scale contribution of the macroscopic gravitational sources (galaxies, clusters, interstellar gas, ...), and the contribution of cosmic backgrounds of heavy, non-relativistic particles (baryons, as well as other, more exotic, possible dark-matter components). As we shall see later in more detail (see Eq. (1.39)), the present energy density of incoherent matter is roughly of the same order of magnitude as the critical density, $\rho_m(t_0) \sim \rho_c(t_0)$, where [12]

$$\rho_c(t_0) = \frac{3\,H_0^2}{8\pi G} = 3\,H_0^2 M_P^2 \simeq 2.25 h^2 \times 10^{-120} M_P^4, \tag{1.33}$$

and is thus much greater than the radiation energy density today, since [12]

$$\rho_r(t_0) \simeq 4.15 h^{-2} \times 10^{-5} \rho_c(t_0). \tag{1.34}$$

Therefore, according to the standard cosmological model, the present scale factor (assuming negligible spatial curvature) should evolve in time as $a(t) \sim t^{2/3}$.

As the Universe expands, however, the energy density of the matter component decreases in time as the inverse of the proper volume, $\rho_m \sim a^{-3}$, i.e. more slowly than the radiation component, $\rho_r \sim a^{-4}$ (see Eq. (1.29)). Going backwards in time one thus necessarily reaches the so-called *equality* time, $t = t_{eq}$, characterized by the same amount of matter and radiation energy density, $\rho_m(t_{eq}) = \rho_r(t_{eq})$. At earlier times, $t < t_{eq}$, the standard model then predicts the existence of a primordial phase where the radiation is the dominant component of the total energy density, and the scale factor evolves with different kinematics, $a(t) \sim t^{1/2}$, according to Eq. (1.30).

It is worth stressing that both the matter-dominated and the radiation-dominated regimes, according to the standard model, correspond to a phase of expansion which is *decelerated* and has *decreasing curvature*, i.e. satisfies

$$\dot{a} > 0\,, \qquad \ddot{a} < 0\,, \qquad \dot{H} < 0\,, \tag{1.35}$$

as one can easily verify by differentiating Eq. (1.30) for $\gamma = 0$ and $\gamma = 1/3$ (with a power-law scale factor, we can take H as a good indicator of the time behavior of the space-time curvature scale). However, the recent large-scale observations concerning both the Hubble diagram of Type Ia Supernovae [13, 14] and the harmonic analysis of the CMB anisotropies [15, 16, 17] seem to indicate, with a growing level of precision and confidence [18, 19, 20], that the present Universe is undergoing a phase of accelerated expansion, $\ddot{a} > 0$.

Such observations are thus compatible with the first of Eqs. (1.25) only if the sources of cosmic gravity are presently dominated by a component with negative enough pressure (i.e. $\rho + 3p < 0$), so as to produce a kind of "cosmic repulsion" on large scales. Adding explicitly this new source ρ_q (dubbed "*quintessence*", or "*dark energy*") to the usual dust matter sources ρ_m, Eq. (1.26) becomes

$$H^2 + \frac{K}{a^2} = \frac{8\pi G}{3}(\rho_m + \rho_q), \tag{1.36}$$

where $\rho_q > \rho_m$, and $p_q/\rho_q \equiv \gamma_q < -1/3$. Dividing by H^2 we can then obtain a relation between the various components of the cosmic fluid in critical units, i.e.

$$1 = \Omega_m + \Omega_q + \Omega_K, \tag{1.37}$$

where

$$\Omega_m = \frac{\rho_m}{\rho_c}, \qquad \Omega_q = \frac{\rho_q}{\rho_c}, \qquad \Omega_K = -\frac{K}{a^2 H^2}. \tag{1.38}$$

The simplest model of dark energy is a cosmological constant, $\rho_q = \Lambda = \text{const}$ (which corresponds to $\gamma_q = -1$). In this case, replacing Ω_q with $\Omega_\Lambda = \Lambda/\rho_c$, the results of present observations can be summarized as follows [12]:

$$\Omega_m = 0.24^{+0.03}_{-0.04}, \qquad \Omega_\Lambda = 0.76^{+0.04}_{-0.06}. \tag{1.39}$$

These results refer to the particular case $K = 0$, but can be consistently applied to the present cosmological state where the allowed deviations of $\Omega_m + \Omega_\Lambda$ from 1 are very small: indeed,

$$\Omega_K = -0.015^{+0.020}_{-0.016} \tag{1.40}$$

according to a recent combination of supernovae and CMB data [20].

The experimental results are not very different from those of Eq. (1.39) even if ρ_q does not correspond to a cosmological constant, but represents the contribution of some weakly coupled, time-dependent field, as will be discussed in Section 9.3. In such a case, the effective equation of state $\gamma_q = p_q/\rho_q$ of the dark-energy component is presently constrained by the limits

$$\gamma_q = -0.97^{+0.07}_{-0.09}, \tag{1.41}$$

obtained by combining supernovae and CMB data [20] (and assuming $K = 0$). In any case, it may be noted that the dilution of the dark-energy density due to the expansion of the Universe, $\rho_q \sim a^{-3(1+\gamma_q)}$, is much slower than the corresponding dilution of the matter component, $\rho_m \sim a^{-3}$. Therefore, going backward in time, the dominance of ρ_q and the associated cosmic acceleration tend to disappear quickly. In a decelerated Universe, on the other hand, the contribution of the spatial curvature decreases as $\Omega_K(t) \sim (aH)^{-2}$, going backward in time. Considering the present limits (1.40) we are thus fully entitled to neglect the spatial curvature during the early stages of the standard cosmological evolution.

It is also worth mentioning that the addition of ρ_q (with negative pressure) to the Einstein equation (1.36) may drastically change the conventional, well-known picture (see for instance [1]) where a Universe with positive spatial curvature $\Omega_K < 0$ (also called a "closed" Universe) will collapse in a finite time with a future "big crunch", while a Universe with $\Omega_K > 0$ (also called an "open" Universe) will expand forever. If $\Omega_q \neq 0$ there are indeed closed models with $\Omega_m + \Omega_q > 1$ which are of the "hyperbolic" type, and evergrowing, and open models with $\Omega_m + \Omega_q < 1$ which are of the "elliptic" type, and recollapsing. This possibility can be easily explored by assuming for instance $\rho_q = \Lambda$, and performing the numerical integration of Eq. (1.36) for various different initial values of ρ_m and ρ_q (see for instance [21]). If, in addition, $\gamma_q \neq -1$, and/or γ_q is time dependent, we can find different types of singularities eventually characterizing the future configuration of our Universe: *"big rip"* singularities [22] and *"sudden"* singularities [23].

In order to obtain experimental information on the parameters characterizing our present cosmological state, such as $\Omega_m(t_0)$, $\Omega_q(t_0)$, γ_q, H_0, we can use two important quantities which can be directly confronted with observations: (1) the so-called "**age of the Universe**", t_0, and (2) the **luminosity distance**, $d_L(t_0)$.

(1) The first parameter t_0 simply (and more properly) represents the time scale of our present cosmological state, and can be defined starting from Eq. (1.36). Expressing the scale factor in terms of the redshift parameter (1.13), i.e. $a(t) = a_0(1+z)^{-1}$, and using the explicit time evolution (1.29) of the ρ_m, ρ_q components, Eq. (1.36) can then be recast in the form

$$(1+z)^{-2}\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2 = H^2(z), \tag{1.42}$$

where

$$\begin{aligned} H(z) &= H_0\left[\Omega_m(t_0)(1+z)^3 + \Omega_q(t_0)(1+z)^{3+3\gamma_q} + \Omega_K(t_0)(1+z)^2\right]^{1/2} \\ &= H_0(1+z)\left\{1 + z\Omega_m(t_0) + \Omega_q(t_0)\left[(1+z)^{1+3\gamma_q} - 1\right]\right\}^{1/2}. \end{aligned} \tag{1.43}$$

(We have used the definitions (1.38) and, in the second line of the equation, we have eliminated Ω_K through Eq. (1.37).) Let us now integrate Eq. (1.42) from $t=0$ to $t=t_0$, assuming that $z\to\infty$ for $t\to 0$ (in other words, we are extrapolating the standard model up to the so-called "big bang" singularity $a=0$ at $t=0$). We obtain

$$t_0=\int_0^\infty \frac{\mathrm{d}z}{(1+z)H(z)}, \tag{1.44}$$

which defines t_0 as a function of the four parameters H_0, $\Omega_m(t_0)$, $\Omega_q(t_0)$ and γ_q.

The precision of this definition can be improved by inserting into Eq. (1.43) the contribution of the radiation energy density, which scales as $\rho_r\sim(1+z)^4$, and becomes important at earlier epochs than equality (i.e. for $z\gtrsim 10^4$, see below). In any case, because of our ignorance about the very early cosmological evolution, one cannot determine t_0 from any direct observation; however, given the age of some component of our present Universe, one can put lower limits on t_0, and then derive indirect constraints on the dark-matter and dark-energy parameters.

(2) What can be directly measured is the correlation between the luminosity and the redshift of signals received from very distant sources. To obtain such a correlation we can consider a signal propagating towards the origin along a null radial geodesic, and satisfying the differential condition

$$\frac{\mathrm{d}r}{\sqrt{1-Kr^2}}=\frac{\mathrm{d}r}{\left[1+a_0^2H_0^2\Omega_K(t_0)r^2\right]^{1/2}}=\frac{\mathrm{d}t}{a}=\frac{(1+z)}{a_0}\mathrm{d}t=\frac{\mathrm{d}z}{a_0H(z)} \tag{1.45}$$

(we have used the definitions of Ω_K and of z, and Eq. (1.42) for $\mathrm{d}t/\mathrm{d}z$). Suppose that the signal was emitted at a distance r from the origin: by integrating between 0 and r the first term of the above equation, taking into account the intrinsic sign of Ω_K, and using the elementary results

$$\int\frac{\mathrm{d}x}{\sqrt{1+\alpha x^2}}=\begin{cases}x, & \alpha=0,\\ \alpha^{-1/2}\sinh^{-1}(\sqrt{\alpha}\,x), & \alpha>0,\\ \alpha^{-1/2}\sin^{-1}(\sqrt{\alpha}\,x), & \alpha<0,\end{cases} \tag{1.46}$$

we can then obtain from Eq. (1.45) the comoving distance of the source as a function of the redshift of the received signal, $r(z)$, as follows:

$$a_0\,r(z)=\begin{cases}\int_0^z \frac{\mathrm{d}z'}{H(z')}, & K=0,\\ H_0^{-1}\,|\Omega_K|^{-1/2}\sinh\left[H_0\,|\Omega_K|^{1/2}\int_0^z\frac{\mathrm{d}z'}{H(z')}\right], & K<0,\\ H_0^{-1}\,|\Omega_K|^{-1/2}\sin\left[H_0\,|\Omega_K|^{1/2}\int_0^z\frac{\mathrm{d}z'}{H(z')}\right], & K>0,\end{cases} \tag{1.47}$$

where $H(z)$ is defined by Eq. (1.43).

The above relation for $r(z)$ cannot be directly applied to observations, however, because we do not know the comoving radial distance of the various astrophysical sources. For pratical applications we must use, instead of the radial distance, the notion of apparent and absolute magnitude, commonly used to set the distance scales of astronomical observations. Let us consider, for this purpose, a source of massless radiation located at a distance $r_{\rm em}$ from the origin, with absolute emitting power (or luminosity) $L_{\rm em} = (\mathrm{d}E/\mathrm{d}t)_{\rm em}$. The energy flux received at $r = 0$, per unit of time and surface, at time $t = t_0$, is then given by

$$L_0 = \frac{1}{4\pi d_0^2}\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_0, \tag{1.48}$$

where

$$d_0 \equiv d(t_0) = a(t_0)\int_0^{r_{\rm em}} \frac{\mathrm{d}r}{\sqrt{1-Kr^2}} \equiv a_0\, r_{\rm em}(z) \tag{1.49}$$

is the proper distance of the source at time t_0, expressed as in Eq. (1.47). Because of the frequency shift (1.12), the received energy will be shifted by the factor $(\mathrm{d}E)_0 = (\mathrm{d}E)_{\rm em}(a_{\rm em}/a_0)$. The time intervals will also be shifted, for the same reason, as $(\mathrm{d}t)_0 = (\mathrm{d}t)_{\rm em}(a_0/a_{\rm em})$. Taking into account the total shift of the received power, we can thus express the apparent luminosity L_0, for a source at a distance r, at the time t_0, as

$$L_0 = \frac{L_{\rm em}}{4\pi d_0^2}\left(\frac{a_{\rm em}}{a_0}\right)^2 = \frac{L_{\rm em}}{4\pi a_0^2 r_{\rm em}^2 (1+z_{\rm em})^2}. \tag{1.50}$$

We can now introduce the so-called "luminosity distance" of the source, defined as the proper distance $d_L(z)$ such that $L_{\rm em}/L_0 = 4\pi d_L^2(z)$. For a source located at a distance $r(z)$ we obtain, using Eq. (1.47),

$$d_L(z) = a_0\, r(z)(1+z) = \frac{(1+z)}{H_0\,|\Omega_K|^{1/2}}\, \mathcal{F}\left[H_0\,|\Omega_K|^{1/2}\int_0^z \frac{\mathrm{d}z'}{H(z')}\right], \tag{1.51}$$

where the function $\mathcal{F}$ is defined as $\mathcal{F}(x) = \sinh x$ if $\Omega_K > 0$, $\mathcal{F}(x) = \sin x$ if $\Omega_K < 0$, and $\mathcal{F}(x) = x$ if $\Omega_K = 0$. The conventional astronomical unit of luminosity is the *apparent magnitude* m, defined by

$$m = -2.5\log_{10} L_0 + \text{const}, \tag{1.52}$$

where the constant is conventionally fixed by defining the apparent magnitude of the pole star to be $m = 2.15$. Comparing Eqs. (1.50) and (1.52) we finally obtain

$$m(z) = 5\log_{10} d_L(z) + c_M, \tag{1.53}$$

where c_M is a z-independent quantity related to the absolute magnitude M (i.e. to the absolute luminosity) of the source (see also [1]).

Fitting the experimental data for $m(z)$ in terms of the curves generated by the theoretical predictions (1.51) and (1.53) it becomes possible, in principle, to determine the parameters H_0, $\Omega_m(t_0)$, $\Omega_q(t_0)$ and γ_q contained in $H(z)$. This analysis is usually performed using the luminosity distance of the various models, computing the so-called "distance modulus", i.e. the difference between the apparent and absolute magnitude $m-M$, and finally plotting the difference $\Delta(m-M)$ between the distance modulus of a given model and the distance modulus of the hyperbolic empty model with $\Omega_m = 0 = \Omega_q$.

This last special model is characterized by $\Omega_K = 1$, and corresponds to the well-known Milne parametrization of the globally flat Minkowski space, represented in Robertson–Walker form by a linear (cosmic time) evolution, $a = t/t_0$, and by spatial sections with constant negative curvature $K = -1/t_0^2$. The luminosity distance for such a model, according to Eqs. (1.51) and (1.43), is given by

$$d_L^0(z) = \frac{(1+z)}{H_0}\sinh\ln(1+z) = \frac{z(2+z)}{2H_0}. \tag{1.54}$$

A convenient phenomenological representation of the distance–redshift relation is then obtained through the variable

$$\begin{aligned}\Delta(m-M) &= 5\log_{10} d_L(z) - 5\log_{10} d_L^0(z)\\ &= 5\log_{10}\left\{\frac{2(1+z)}{z(2+z)\,|\Omega_K|^{1/2}}\mathcal{F}\left[H_0\,|\Omega_K|^{1/2}\int_0^z \frac{dz'}{H(z')}\right]\right\},\end{aligned} \tag{1.55}$$

where $H(z)$ is given by Eq. (1.43). Note that such a relation can be easily extended to a generic model containing an arbitrary number of sources $\rho_n(t)$, evolving independently according to Eq. (1.29), provided we replace $H(z)$ with the more general expression

$$H(z) = H_0\left[\Omega_K(t_0)(1+z)^2 + \sum_n \Omega_n(t_0)(1+z)^{3+3\gamma_n}\right]^{1/2}. \tag{1.56}$$

1.1.4 Thermal equilibrium

Another important assumption of the standard cosmological model concerns the spectral distribution of the radiation fluid. Following present observational evidence, the radiation is assumed to be in a state of thermodynamic equilibrium at a proper temperature T, with a Planck or Fermi–Dirac distribution according to the bosonic or fermionic character of the various radiation components. The energy distribution, per unit volume and per unit logarithmic frequency, can then be written in the form

$$\frac{\mathrm{d}\rho(\omega,t)}{\mathrm{d}\log\omega} \equiv \omega\frac{\mathrm{d}\rho}{\mathrm{d}\omega} = \frac{N}{2\pi^2}\frac{\omega^4}{e^{\omega/T}\pm 1} \tag{1.57}$$

(see [1, 5]), where the + and − signs correspond to the fermionic and bosonic cases, respectively, and N is the number of independent polarization states (for instance, $N = 2$ for photons, ultrarelativistic electrons and positrons, $N = 1$ for any relativistic neutrino/antineutrino species). Integrating over all modes we obtain the total energy density, which is given by

$$\rho_b(t) = \frac{\pi^2}{30} N_b T_b^4 \tag{1.58}$$

for pure bosonic radiation, and by

$$\rho_f(t) = \frac{7}{8}\frac{\pi^2}{30} N_f T_f^4 \tag{1.59}$$

for pure fermionic radiation. For a thermal mixture of N_b bosonic and N_f fermionic states the total energy density can then be written as

$$\rho_r = \frac{\pi^2}{30} N_\star T^4 , \tag{1.60}$$

where

$$N_\star = \sum_b N_b \left(\frac{T_b}{T}\right)^4 + \frac{7}{8}\sum_f N_f \left(\frac{T_f}{T}\right)^4 \tag{1.61}$$

is the total effective number of degrees of freedom in thermal equilibrium at temperature T.

It may be useful, for later applications, to recall that the entropy S of a system in thermal equilibrium at temperature T, characterized by proper volume $V \sim a^3$, pressure p, and energy density ρ, must satisfy the differential thermodynamic condition

$$\mathrm{d}S = \frac{1}{T}\left[\mathrm{d}(\rho V) + p\,\mathrm{d}V\right]. \tag{1.62}$$

It follows that the thermal entropy S is exactly conserved during the standard cosmological evolution, thanks to the conservation equation (1.27), which (multiplied by V) can be rewritten in differential form as

$$V\,\mathrm{d}\rho = -3\frac{\mathrm{d}a}{a}(\rho+p)V = -(\rho+p)\mathrm{d}V. \tag{1.63}$$

Substituting into Eq. (1.62) this condition leads in fact to $\mathrm{d}S = 0$, which implies an exact adiabatic evolution for each decoupled component of the cosmic fluid in thermal equilibrium.

Using T and V as independent variables, differentiating Eq. (1.62) twice, and imposing the integrability condition $\partial^2 S/\partial V\, \partial T = \partial^2 S/\partial T\, \partial V$, one also obtains [1]

$$T\,\mathrm{d}p = (\rho + p)\mathrm{d}T\,, \tag{1.64}$$

which allows one to rewrite Eq. (1.62) as

$$\mathrm{d}S = \frac{1}{T}\left[\mathrm{d}(\rho V) + \mathrm{d}(pV) - V\,\mathrm{d}p\right] = \mathrm{d}\left[\frac{V}{T}(\rho + p)\right]. \tag{1.65}$$

The integration provides the entropy density σ, for a generic equation of state $p = \gamma\rho$,

$$\sigma = \frac{S(T,V)}{V} = \frac{1+\gamma}{T}\rho. \tag{1.66}$$

For a radiation fluid, in particular, $\gamma = 1/3$ and $\sigma = 4\rho/3\,T$, a result valid for bosons as well as for fermions. For a thermal mixture, with N_b bosonic and N_f fermionic states, we can use Eqs. (1.58) and (1.59) to obtain

$$\sigma_r(t) = \frac{2\pi^2}{45} g_\star T^3\,, \tag{1.67}$$

where

$$g_\star = \sum_b N_b \left(\frac{T_b}{T}\right)^3 + \frac{7}{8}\sum_f N_f \left(\frac{T_f}{T}\right)^3 \tag{1.68}$$

is the effective number of degrees of freedom contributing to the thermal entropy density at a given time t. It is important to stress that this number, as well as the number $N_\star$ of Eq. (1.61), is in general time dependent in a cosmological context, and that a change in $g_\star$ (due for instance to the disappearance of some degrees of freedom from the thermal mixture) must necessarily be accompanied by a corresponding variation of the temperature, in order for the total entropy to be conserved.

The direct integration of Eq. (1.64) provides another important relation between ρ and T for a fluid in thermal equilibrium, namely

$$T \sim \rho^{\gamma/(1+\gamma)}. \tag{1.69}$$

The time evolution of ρ, on the other hand, is determined by the solution (1.29) of the conservation equation. The combination of these two results implies that the proper temperature of the thermal mixture is not a constant in the Robertson–Walker geometry, but evolves in time as

$$T(t) \sim a^{-3\gamma}. \tag{1.70}$$

A radiation fluid, in particular, has $\gamma = 1/3$ and $T(t) \sim a^{-1}(t)$. It follows that the radiation temperature is redshifted by the cosmological expansion exactly as the proper frequency $\omega(t)$ (see Eq. (1.12)). The same conclusion can be reached by combining Eqs. (1.60) and (1.29).

Thus, although the radiation becomes colder because of the expansion, the ratio ω/T is constant, and *the shape* (1.57) of the (bosonic and fermionic) spectral distributions does not change in time (even if the overall height of the peak of the distributions decreases, being controlled by T^4). This means that the condition of thermal equilibrium is preserved in the course of the standard cosmological expansion: as a consequence, one can take the CMB temperature T_γ as a useful evolution parameter – like the cosmic time, or the space-time curvature radius – to which to refer the various phases of the history of our Universe. Using as a reference the present value of the CMB temperature [12],

$$T_0 \equiv T_\gamma(t_0) = 2.725 \pm 0.001\,\mathrm{K} \sim 2.3 \times 10^{-4}\,\mathrm{eV}, \tag{1.71}$$

one can compute, for instance, the temperature at the epoch of matter–radiation equality, $t = t_{\rm eq}$, when $\rho_m = \rho_r$. From the Einstein equations we have $\rho_r/\rho_m \sim a^{-1}(t)$, hence,

$$\frac{\rho_r(t_0)/\rho_m(t_0)}{\rho_r(t_{\rm eq})/\rho_m(t_{\rm eq})} = \frac{a_{\rm eq}}{a_0} = \frac{1}{1+z_{\rm eq}}. \tag{1.72}$$

From the condition of thermodynamic equilibrium, on the other hand, $(1+z_{\rm eq})^{-1} = T_0/T_{\rm eq}$. We can therefore write

$$\begin{aligned} T_{\rm eq} = T_0\,(1+z_{\rm eq}) &= T_0\,\frac{\rho_m(t_0)}{\rho_r(t_0)} \\ &\simeq 0.7 \times 10^4 T_0 h^2 \left(\frac{\Omega_m}{0.3}\right) \simeq 1.6\,h^2 \left(\frac{\Omega_m}{0.3}\right) \mathrm{eV} \\ &\simeq 2 \times 10^4 h^2 \left(\frac{\Omega_m}{0.3}\right) \mathrm{K}, \end{aligned} \tag{1.73}$$

where we have used Eqs. (1.34) and (1.71), and a typical value of the matter density suggested by the present data (see Eq. (1.39)).

During the radiation-dominated epoch the temperature is directly related to another important evolution parameter, the curvature scale $H(t)$. Indeed, using Eqs. (1.26) and (1.60), and neglecting the contribution of the spatial curvature, one obtains

$$H(t) = \left(\frac{\pi^2 N_\star}{90}\right)^{1/2} \frac{T^2(t)}{M_{\rm P}}. \tag{1.74}$$

As a simple application, also useful for future discussions, this equation can be used to estimate the curvature scale at the epoch of matter–radiation equality, when the radiation temperature is given by Eq. (1.73).

For a precise computation of $N_\star$ we need to take into account that the cosmic radiation fluid at $t = t_{\rm eq}$, according to the standard model of particle interactions, should contain two bosonic degrees of freedom, associated with the polarization states of the photon, and six fermionic degrees of freedom, associated with the three neutrino flavors and the corresponding antineutrinos (we are neglecting other, possibly sub-leading, contributions such as that of gravitons, see Chapter 7). However, neutrinos are slightly colder than photons, since the photon gas has been heated up by the annihilation of the electron/positron pairs taking place well before $t_{\rm eq}$, at a temperature of about 0.43 MeV [1]. Indeed, at the epoch of electron annihilation, the conservation of the entropy associated with the thermal mixture of photons (γ) and electron/positron ($\mathrm{e}^\pm$) pairs has caused a jump in the electromagnetic temperature, from an initial value identical to the neutrino temperature, $T = T_\nu$, to a new value $T = T_\gamma$ such that $\sigma(\gamma) = \sigma(\gamma, \mathrm{e}^\pm)$. Taking into account that the $\mathrm{e}^\pm$ pairs contribute to the fermionic degrees of freedom before annihilation, and using Eqs. (1.67) and (1.68), one obtains from entropy conservation

$$g_\star(\gamma)T_\gamma^3 = 2\,T_\gamma^3 = g_\star(\gamma, \mathrm{e}^\pm)T_\nu^3 = \left(2 + \frac{7}{8}\times 4\right)T_\nu^3, \tag{1.75}$$

from which

$$T_\gamma = \left(\frac{11}{4}\right)^{1/3} T_\nu. \tag{1.76}$$

Let us now consider the total radiation energy density (1.60), where we take T_γ as the reference temperature. After the $\mathrm{e}^\pm$ annihilation epoch we have $N_b = 2$, $N_f = 6$, so that

$$N_\star = \sum_b N_b + \frac{7}{8}\sum_f N_f \left(\frac{T_\nu}{T_\gamma}\right)^4 = 2 + \frac{42}{8}\left(\frac{4}{11}\right)^{4/3} \simeq 3.36\,. \tag{1.77}$$

At the time of matter–radiation equality, neglecting the spatial curvature and a possible dark-energy contribution, the total energy density is $\rho(t_{\rm eq}) = \rho_m + \rho_r = 2\,\rho_r(t_{\rm eq})$. Using Eqs. (1.60), (1.73) and (1.77) one finally obtains

$$H_{\rm eq} = \left(\frac{3.36\pi^2}{45}\right)^{1/2} \frac{T_{\rm eq}^2}{M_{\rm P}} \simeq 3.7\times 10^{-55} h^4 \left(\frac{\Omega_m}{0.3}\right)^2 M_{\rm P}. \tag{1.78}$$

This value is still a very tiny fraction of the Planck mass, but is nevertheless much greater than the present curvature scale [12]:

$$H_0 \simeq 8.7h\times 10^{-61} M_{\rm P} \simeq 2.35\times 10^{-6} H_{\rm eq}\left(\frac{0.3}{\Omega_m}\right)^2 h^{-3}. \tag{1.79}$$

The two curvatures H_0 and $H_{\rm eq}$ will be used as convenient reference scales in the following chapters.

The standard cosmological model provides a detailed thermal history of the Universe [1, 5], and suggests an evolution scenario where an initially hot, dense and highly curved configuration expands, becoming cooler and flatter. This scenario is in excellent agreement with important observational data referring to our present cosmological state, such as those concerning the galactic recession velocities and the relic background of microwave radiation. It is also consistent with the primordial mechanisms of nucleosynthesis and baryogenesis, which can only take place in the presence of a sufficiently high temperature.

The cosmological solutions of the standard model, however, cannot be extended indefinitely backward in time. In the radiation-dominated solution, for instance, the energy density and the temperature diverge at a fixed instant of time, conventionally chosen to coincide with $t = 0$:

$$t \to 0 \quad \Rightarrow \quad \rho \sim T^4 \sim a^{-4} \sim t^{-2} \to \infty\,. \tag{1.80}$$

At the same instant of time the curvature invariants also diverge:

$$t \to 0 \quad \Rightarrow \quad \left(R_{\mu\nu}R^{\mu\nu}\right)^{1/2} \sim \left(R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta}\right)^{1/2} \sim H^2 \sim t^{-2} \to \infty\,. \tag{1.81}$$

This singularity is a consequence of rigorous theorems formulated within the theory of general relativity (see for instance [24]), and cannot be removed even by abandoning the symmetry hypotheses underlying the Robertson–Walker metric (see for instance the discussion of the Kasner solution in [2]). It can be shown, in particular, that a geodesic time-like curve of the standard model, evolved backward in time from any given finite epoch t_0, reaches the singularity at $t = 0$ in a *finite* value of its affine parameter (i.e. in a finite proper time interval). At the classical level, the space-time cannot be extended beyond a singuarity, and this implies that the time $t = 0$ (the so-called "big-bang" singularity) should be interpreted, within the standard cosmological model, as the *beginning* of space-time, and as the *birth* of the Universe itself.

This conclusion could be avoided if some drastic modification of the standard scenario were to take place before reaching the initial singularity. After all, the standard cosmological model is based on general relativity, i.e. on a classical theory which is not guaranteed to be valid when the space-time curvature becomes large in Planck units, and the Universe enters the quantum gravity regime.

On the other hand, as already stressed in this section, recent observations indicate that the standard cosmological model has to be modified even at the present low-curvature scales, in order to account for the "cosmic repulsion" producing an accelerated space-time expansion. In addition, as we shall see in the next section, there are many valid reasons (other than the existence of a singularity) why the standard cosmological model should be modified when the curvature reaches high enough scales (i.e. at early enough cosmological epochs). These primordial modifications lead to the introduction of the so-called inflationary scenario, which will be illustrated in the next section.

1.2 The inflationary cosmological model

The present observed values of the cosmological parameters, if taken as initial conditions to evolve our Universe backward in time according to the standard picture, lead us to a primordial state characterized by somewhat "unnatural" properties, even without reaching the initial singularity. A dynamical explanation of such peculiar properties of the primordial Universe is possible, provided the epoch of standard (decelerated) expansion is preceded by an appropriate phase of accelerated evolution, dubbed "inflation" [25, 26, 27]. We start this section by presenting simple arguments that motivate the introduction of such a phase.

1.2.1 Standard kinematic problems

A first argument is based on the so-called "flatness problem". As already pointed out in Section 1.1, the spatial curvature today provides only a small, sub-dominant contribution to the total space-time curvature (see Eq. (1.40)). This contribution, however, is not a constant,

$$|\Omega_K| \sim (aH)^{-2} \equiv r^2(t), \tag{1.82}$$

being controlled by the parameter $r(t)$ which is a monotonically increasing function of time in the standard cosmological model. For an expanding, power-law scale factor, $a(t) \sim t^\alpha$, with $0 < \alpha < 1$ and $t \to \infty$, one obtains

$$t \to \infty \quad \Rightarrow \quad r(t) = (aH)^{-1} = \dot{a}^{-1} \sim t^{1-\alpha} \to \infty, \tag{1.83}$$

so that $r(t)$ is increasing in both the matter-dominated ($\alpha = 2/3$) and radiation-dominated ($\alpha = 1/2$) eras. This implies that the contribution of the spatial curvature becomes less and less significant as we go back in time, according to the standard-model equations.

Let us consider, for instance, the Planck epoch $t = t_P$, defined as the time when $H = M_P$ and $\rho = \rho_c = 3M_P^4$. Using the kinematic properties of the standard model ($a \sim H^{-2/3}$, $a \sim H^{-1/2}$ for the matter- and radiation-dominated eras, respectively), we can rescale the parameter $r(t)$ to the time t_P as follows:

$$\frac{r_P}{r_0} \equiv \frac{r(t_P)}{r(t_0)} = \frac{(aH)_0}{(aH)_{eq}} \frac{(aH)_{eq}}{(aH)_P} = \left(\frac{H_0}{H_{eq}}\right)^{1/3} \left(\frac{H_{eq}}{M_P}\right)^{1/2}. \tag{1.84}$$

Using Eqs. (1.78) and (1.79) for the values of H_0 and H_{eq}, and adopting the conservative constraint $|\Omega_K(t_0)| < 0.1$, we obtain

$$|\Omega_K(t_P)| = |\Omega_K(t_0)| \left(\frac{r_P}{r_0}\right)^2 \lesssim 10^{-60}. \tag{1.85}$$

Such a large suppression of the spatial curvature with respect to the space-time curvature represents a rather unnatural initial condition, and requires a significant amount of fine-tuning. Also, what makes the problem even more serious is the fact that a violation of the above upper limit by only a few orders of magnitude would be enough to forbid the formation of our present cosmological configuration. In that case the Universe would enter (much before the present epoch) a curvature-dominated phase ($|\Omega_K| \sim 1$), which would lead to a subsequent collapse if $K > 0$, and which would not be appropriate to sustain the formation of large-scale structures if $K < 0$.

The initial condition (1.85) can be dynamically explained, however, if the phase of standard evolution is preceded by a primordial phase during which the function $r(t)$ *decreases* in time (instead of growing, as in Eq. (1.83)). It then becomes possible to start from a "natural" set of initial conditions for this primordial epoch, characterized by $r \sim 1$, provided that the decrease of r during such an epoch is large enough to compensate the subsequent growth produced by the standard evolution.

As an example of this non-standard phase let us consider again a power-law scale factor, $a(t) \sim t^\beta$, with $t \to \infty$. If β is large enough (in particular, $\beta > 1$), then the function $r(t)$ decreases in time:

$$t \to \infty \quad \Rightarrow \quad r(t) = \dot{a}^{-1} \sim t^{1-\beta} \to 0. \tag{1.86}$$

It is straightforward to check that such a scale factor describes accelerated expansion,

$$\frac{\dot{a}}{a} = \frac{\beta}{t} > 0, \qquad \frac{\ddot{a}}{a} = \frac{\beta(\beta-1)}{t^2} > 0, \tag{1.87}$$

hence the term *inflation* used to denote this phase, complementary to the decelerated evolution of the standard cosmological model (see Chapter 5 for a general classification of the various classes of inflationary kinematics). As will be shown later, this kind of accelerated expansion can be obtained from the Einstein equation using, as gravitational source, a scalar field with an appropriate exponential potential (see Eqs. (1.120) and (1.121)).

The presence of such an inflationary epoch, besides solving the flatness problem, also provides a solution to the so-called *horizon problem* of the standard cosmological model. The standard cosmological evolution ($a \sim t^\alpha,\ \alpha < 1$) is characterized, in fact, by the existence of particle horizons (see Section 1.1), which control, at any given instant of time, the maximum size of the spatial regions within which causal interactions take place. The proper size of the particle horizon is of the order of the Hubble radius H^{-1}, and grows linearly with the cosmic time, according to Eq. (1.17):

$$d_{\mathrm{p}}(t) = a(t)\int_0^t \frac{\mathrm{d}t'}{a(t')} = \frac{t}{1-\alpha} = \frac{\alpha}{1-\alpha}H^{-1}(t). \tag{1.88}$$

Let us consider the spatial section of the Universe included within our current particle horizon, namely the portion of space of typical size H_0^{-1}, currently accessible to our direct observation. Going backward in time, the proper volume of this spatial region decreases as a^3, and therefore its proper radius decreases as $a \sim t^{\alpha}$. The radius of the particle horizon, i.e. of the causally connected portion of space, also decreases going backward in time, but goes linearly with the cosmic time (according to Eq. (1.88)), and thus faster than the scale factor (recall that $\alpha < 1$). As a consequence, the portion of space that we are currently observing was, in the past, much bigger than the corresponding extension of the particle horizon: in other words, many parts of the currently observed Universe were not causally connected. If we rescale, for instance, the proper size of the present observable Universe, H_0^{-1}, down to the Planck epoch $t = t_{\mathrm{P}}$, when the horizon size was $H_{\mathrm{P}}^{-1} = M_{\mathrm{P}}^{-1}$, we obtain a proper radius much larger than the horizon:

$$\frac{a(t_{\mathrm{P}})}{a(t_0)}H_0^{-1} = \frac{r(t_0)}{r(t_{\mathrm{P}})}M_{\mathrm{P}}^{-1} \sim 10^{29}M_{\mathrm{P}}^{-1} \gg M_{\mathrm{P}}^{-1}. \tag{1.89}$$

Given such initial conditions, we are led to the questions: why is the current Universe so homogeneous and isotropic, or why is the average CMB temperature everywhere the same, as if all the portions of space we are now observing were in the past in causal contact, and had time to interact and thermalize?

An interesting solution of this problem arises from noticing that the ratio between the horizon size ($\sim H^{-1}$) and the proper size of a spatial region ($\sim a$) is governed by the same function $r(t) = (aH)^{-1}$ as controls the ratio between the spatial curvature and the space-time curvature. A sufficiently long inflationary phase, which makes $r(t)$ decreasing and which is able to solve the flatness problem, can thus simultaneously also solve the horizon problem. Indeed, if $r(t)$ decreases as time goes on, the causally connected regions expand faster than the Hubble horizon: at the end of inflation one then precisely obtains a configuration which corresponds to the "unnatural" initial conditions of the standard cosmological scenario (see Fig. 1.1).

How long does the inflationary phase have to be in order to solve the flatness and horizon problems? The answer depends on both the expansion rate and the

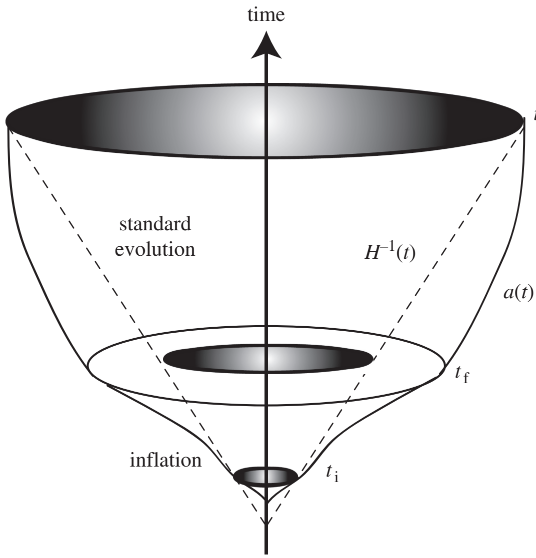

Figure 1.1 Qualitative evolution of the Hubble horizon (dashed line) and of the scale factor (solid curve). The time coordinate is on the vertical axis, while the horizontal axes are space coordinates spanning a two-dimensional spatial section of the cosmological manifold. The inflationary phase extends from t_{i} to t_{f}, the standard cosmological phase from t_{f} to the present time t_0. The shaded areas represent causally connected regions at different epochs. At the beginning of the standard evolution the size of the currently observed Universe (bounded by $a(t)$) is larger than the corresponding Hubble radius (bounded by H^{-1}); all its parts, however, emerge from a spatial region that is causally connected at the beginning of inflation.

beginning of the inflationary epoch. In any case, the decrease of the function $r(t)$, from the beginning t_{i} to the end t_{f} of inflation, has to be large enough to compensate for its subsequent increase from t_{f} to the present time t_0. This defines the following necessary condition to be satisfied by a successful inflationary epoch:

$$\left(\frac{r_{\mathrm{f}}}{r_{\mathrm{i}}}\right) \lesssim \left(\frac{r_{\mathrm{f}}}{r_0}\right), \tag{1.90}$$

where $r_{\mathrm{f}} \equiv r(t_{\mathrm{f}})$, and so on. Assuming for the inflationary phase the power-law evolution (1.86), one then obtains the condition

$$\left(\frac{t_{\mathrm{f}}}{t_{\mathrm{i}}}\right)^{1-\beta} = \left(\frac{H_{\mathrm{i}}}{H_{\mathrm{f}}}\right)^{1-\beta} \lesssim \left(\frac{r_{\mathrm{f}}}{r_{\mathrm{eq}}}\right)\left(\frac{r_{\mathrm{eq}}}{r_0}\right) = \left(\frac{H_{\mathrm{eq}}}{H_{\mathrm{f}}}\right)^{1/2}\left(\frac{H_0}{H_{\mathrm{eq}}}\right)^{1/3}, \tag{1.91}$$

which determines t_{f}, i.e. the scale H_{f} at which inflation ends, as a function of β and t_{i}. It may be useful to notice that, for scale factors following a power-law evolution in time, the function r is proportional to the conformal time coordinate: $r \sim \dot{a}^{-1} \sim \eta = \int a^{-1}\,\mathrm{d}t$. The condition (1.90) then directly provides the minimum duration of inflation in conformal time, i.e.

$$\left|\frac{\eta_{\mathrm{f}}}{\eta_{\mathrm{i}}}\right| \lesssim \left|\frac{\eta_{\mathrm{f}}}{\eta_0}\right|. \tag{1.92}$$

We have inserted the absolute value because, as we shall see later, inflationary (accelerated) scale factors are parametrized by a power-law evolution within a range of negative values of the conformal time coordinate.

The presence of a primordial inflationary phase, characterized by accelerated kinematics, is today universally accepted as the most natural complement of the subsequent decelerated expansion, driven by the standard radiation/matter fluids. The presence of such an inflationary phase allows us to explain the peculiar initial conditions of the standard cosmological model; in addition, it provides a dynamical mechanism for the origin of the large-scale structures and of the small CMB anisotropies (as will be discussed in Chapter 8). It is therefore natural to try to address in the inflationary context the crucial problem of the standard cosmological model, i.e. the presence of the initial singularity.

The premises are encouraging. One of the *necessary* (although not sufficient) conditions for avoiding the singularity is the violation of the so-called condition of "geodesic convergence" [24], for any time-like or null geodesic u^μ. This condition reads, in our notations,

$$R_{\mu\nu}u^\mu u^\nu \geq 0\,, \qquad u_\mu u^\mu \geq 0\,, \tag{1.93}$$

and is also equivalent, using the Einstein equations, to the so-called "strong energy condition" imposed on the gravitational sources,

$$T_{\mu\nu}u^\mu u^\nu \geq \frac{1}{2}T u_\mu u^\mu. \tag{1.94}$$

For a comoving geodesic $u^\mu = (1, \vec{0})$ of the Robertson–Walker metric, and for a perfect fluid source, the above conditions are violated when $p < -\rho/3$, which implies $R_0^0 = -3\ddot{a}/a < 0$, i.e. just when the expansion is accelerated ($\ddot{a} > 0$), and hence inflationary. It is worth noticing, at this point, that a typical example of an inflationary solution (which is also, historically, the first example [25, 28]) is the de Sitter solution, which describes a maximally symmetric, four-dimensional manifold with constant positive curvature, and which is indeed a regular solution of the Einstein equations (all curvature invariants are constant and finite everywhere).

1.2.2 de Sitter inflation

The matter source for the de Sitter solution corresponds to an effective energy-momentum tensor of the type (1.31), where the cosmological constant Λ may be interpreted as the "vacuum" energy density associated with a scalar field (the "inflaton"), frozen at the minimum $\phi = \phi_0$ of an appropriate potential. Consider the action for a self-interacting scalar field ϕ, minimally coupled to gravity,

$$S_m = \frac{1}{2} \int \mathrm{d}^4x \sqrt{-g} \left[g^{\mu\nu}\, \partial_\mu \phi\, \partial_\nu \phi - 2V(\phi) \right]. \tag{1.95}$$

The variation of S_m with respect to ϕ yields the equation of motion

$$\nabla_\mu \nabla^\mu \phi + \frac{\partial V}{\partial \phi} = 0, \tag{1.96}$$

which admits the constant solution $\phi = \phi_0$, $\partial_\mu \phi_0 = 0$, provided ϕ_0 extremizes the scalar potential $(\partial V/\partial\phi)_{\phi=\phi_0} = 0$. The energy-momentum tensor of the scalar field,

$$T_{\mu\nu} = \partial_\mu \phi\, \partial_\nu \phi - \frac{1}{2} g_{\mu\nu} \left[(\nabla\phi)^2 - 2V(\phi) \right], \tag{1.97}$$

for $\phi = \phi_0$ assumes the form (1.31), with $\Lambda = V(\phi_0) = \text{const}$, and the corresponding Einstein equations (1.26) and (1.27) are identically satisfied by the regular de Sitter solution with constant scalar curvature $R = -8\pi G T = -32\pi G V(\phi_0)$.

Using an appropriate, spatially flat ($K = 0$) chart, the solution can be represented in exponential form (see Eq. (1.32)) as

$$\begin{aligned} &\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)|\mathrm{d}\vec{x}|^2, \qquad a(t) = \mathrm{e}^{Ht}, \\ &H = \left[\frac{V(\phi_0)}{3\, M_\mathrm{P}^2} \right]^{1/2} = \text{const}, \qquad -\infty \le t \le \infty. \end{aligned} \tag{1.98}$$

This solution describes accelerated expansion at constant curvature, $\dot{a} > 0$, $\ddot{a} > 0$, $\dot{H} = 0$. Introducing the conformal time coordinate,

$$\eta = \int^t \frac{\mathrm{d}t'}{a(t')} = -\frac{\mathrm{e}^{-Ht}}{H} = -\frac{1}{aH}, \qquad -\infty \le \eta \le 0, \tag{1.99}$$

it can be written in a conformally flat form,

$$\mathrm{d}s^2 = a^2(\eta) \left(\mathrm{d}\eta^2 - |\mathrm{d}\vec{x}|^2 \right), \qquad a(\eta) = (-H\eta)^{-1}, \tag{1.100}$$

and the condition (1.90) of sufficient inflation becomes

$$\frac{r_\mathrm{i}}{r_\mathrm{f}} = \left| \frac{\eta_\mathrm{i}}{\eta_\mathrm{f}} \right| = \frac{a_\mathrm{f}}{a_\mathrm{i}} = \mathrm{e}^{H(t_\mathrm{f} - t_\mathrm{i})} \gtrsim \frac{r_0}{r_\mathrm{f}}. \tag{1.101}$$

It can be easily checked, in this form, that a very short duration (in units of H^{-1}) of the accelerated phase may be enough to compensate for large variations of the function r, even if inflation occurs at very primordial epochs.

Suppose, for instance, that the inflationary phase ends when $H_\mathrm{f} = 10^{-5} M_\mathrm{P}$ (in many models, higher values of the curvature scale are inconsistent with CMB observations, as we shall see in Chapter 7), and that the Universe immediately enters the radiation-dominated regime. The radiation temperature associated with t_f is then of the order of the grand unification theory (GUT) scale, $T_\mathrm{f} \sim 10^{15} - 10^{16}\,\mathrm{GeV}$, according to Eq. (1.74). For exponential (or quasi-exponential) inflation, on the other hand, the time duration of the accelerated phase $\Delta t = t_\mathrm{f} - t_\mathrm{i}$ can be conveniently expressed in terms of the "e-folding factor", $N = \ln(a_\mathrm{f}/a_\mathrm{i})$. In terms of N, the condition (1.101) then becomes

$$\begin{aligned} N = \ln\left(\frac{a_\mathrm{f}}{a_\mathrm{i}}\right) &= H\Delta t \\ &\gtrsim \ln\left(\frac{r_0}{r_\mathrm{f}}\right) = \frac{1}{2}\ln\left(\frac{H_\mathrm{f}}{H_\mathrm{eq}}\right) + \frac{1}{3}\ln\left(\frac{H_\mathrm{eq}}{H_0}\right) \end{aligned} \tag{1.102}$$

(we have used Eq. (1.91)). For $H_\mathrm{f} = 10^{-5} M_\mathrm{P}$ one obtains $N \gtrsim \ln 10^{27} \simeq 62$, i.e. $\Delta t \gtrsim 62\, H^{-1}$, where H is the curvature scale of the de Sitter manifold.

The de Sitter solution may give an appropriate description of the primordial inflationary phase; however, it cannot be extended forward in time towards "too late" epochs, since the Universe must enter into the standard decelerated phase that allows nucleosynthesis and the formation of large-scale structures, and that eventually converges to our present cosmological configuration. The transition (also called "graceful exit") between the inflationary and the standard regime is usually implemented, in conventional models of inflation, by assuming that the scalar field is not exactly constant, frozen at the minimum of its potential; instead, it is initially displaced from this minimum, and "slow-rolls" towards it.

1.2.3 Slow-roll inflation

To illustrate this possibility we start by considering the cosmological equations with the energy-momentum tensor of the scalar field as the only gravitational source. Also, we assume that we are given an initial spatial domain of size smaller than (or comparable to) the initial horizon radius H^{-1}, in which the spatial inhomogeneities of the scalar field are negligible, $|\partial_i \phi| \ll |\dot\phi|$. Restricting to this spatial domain we can then neglect the spatial dependence of our variables, and we can treat the scalar field as a perfect fluid *at rest* in the comoving frame, with the following effective energy density and pressure,

$$\rho = \frac{\dot{\phi}^2}{2} + V(\phi)\,, \qquad p = \frac{\dot{\phi}^2}{2} - V(\phi) \tag{1.103}$$

(see Eq. (1.97)), acting as the source of a homogeneous and isotropic Robertson–Walker geometry. Combining Eqs. (1.25) and (1.26), neglecting K, and using the identity $\ddot{a}/a = \dot{H} + H^2$, we are then led to the following independent Einstein equations:

$$3H^2 = 8\pi G\rho = 8\pi G\left(\frac{\dot{\phi}^2}{2} + V\right), \tag{1.104}$$

$$2\dot{H} = -8\pi G(\rho + p) = -8\pi G\dot{\phi}^2\,. \tag{1.105}$$

We may add to this system the scalar field equation (1.96), which reads

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\partial V}{\partial \phi} = 0. \tag{1.106}$$

For $\dot{\phi} \neq 0$ this equation is not independent, however, since it can be obtained by the differentiation of Eq. (1.104) and its combination with Eq. (1.105).

The dynamics of the slow-roll regime can be conveniently illustrated by using the scalar field as the independent variable (replacing cosmic time) of our differential equations. Denoting with a prime the differentiation with respect to ϕ, and dividing Eq. (1.105) by $\dot{\phi}$ (assuming a monotonic evolution with $\dot{\phi} \neq 0$), we obtain

$$2H' = -\lambda_{\mathrm{P}}^2\dot{\phi} \tag{1.107}$$

(recall that $\lambda_{\mathrm{P}}^2 = 8\pi G$). Inserting $\dot{\phi}$ from this equation into Eq. (1.104) we are led to the first-order equation

$$H'^2 - \frac{3}{2}\lambda_{\mathrm{P}}^2 H^2 = -\frac{1}{2}\lambda_{\mathrm{P}}^4 V, \tag{1.108}$$

which is equivalent to the Hamilton–Jacobi equation for the gravity-scalar field system [29]. Let us also define, for later applications, the following useful parameters:

$$\epsilon_H = -\frac{\dot{H}}{H^2} = -\frac{\mathrm{d}\ln H}{\mathrm{d}\ln a} = \frac{2}{\lambda_{\mathrm{P}}^2}\frac{H'^2}{H^2}, \tag{1.109}$$

$$\eta_H = -\frac{\ddot{\phi}}{H\dot{\phi}} = -\frac{\mathrm{d}\ln\dot{\phi}}{\mathrm{d}\ln a} = \frac{2}{\lambda_{\mathrm{P}}^2}\frac{H''}{H} \tag{1.110}$$

(we have used Eq. (1.107) to obtain the last equalities of both definitions). The subscript H makes explicit reference to the Hamilton–Jacobi formalism.

The so-called slow-roll regime corresponds to a sufficiently slow evolution of the scalar field, initially dominated by the "geometrical friction" term $H\dot{\phi}$, and characterized by a kinetic energy which is negligible with respect to the scalar potential. More precisely, the scalar field is slow-rolling if the following conditions are valid:

$$\ddot{\phi} \ll H\dot{\phi}, \qquad \dot{\phi}^2 \ll V, \qquad \dot{H} \ll H^2. \tag{1.111}$$

The slow-roll regime is thus implemented in the limit in which the parameters (1.109) and (1.110) are very small ($\epsilon_H \ll 1$, $\eta_H \ll 1$), and very slowly varying ($\dot{\epsilon}_H \simeq 0$, $\dot{\eta}_H \simeq 0$, to first order). In this limit we can immediately integrate Eq. (1.109) to obtain $H = (\epsilon_H t)^{-1}$. A second integration leads to the scale factor $a(t) \sim t^{1/\epsilon_H}$. Using conformal time,

$$a(\eta) \sim (-\eta)^{-(1+\epsilon_H)}. \tag{1.112}$$

One thus obtains an inflationary (i.e. accelerated) scale factor which approximates the de Sitter metric in the limit $\epsilon_H \to 0$ (see Eq. (1.100)).

The cosmological dynamics during the slow-roll regime is well described by the two independent equations

$$3H^2 = \lambda_{\mathrm{P}}^2 V, \qquad 3H\dot{\phi} = -V', \tag{1.113}$$

obtained from Eqs. (1.108) and (1.106), respectively, using the conditions $\epsilon_H \ll 1$, $\eta_H \ll 1$. Differentiating with respect to ϕ the first equation, we obtain

$$\frac{H'}{H} = \frac{V'}{2V}. \tag{1.114}$$

Inserting this condition into the exact definitions (1.109) and (1.110) we are led to approximate relations defining two new parameters, ϵ and η [30], satisfying

$$\epsilon_H \simeq \frac{1}{2\lambda_{\mathrm{P}}^2}\left(\frac{V'}{V}\right)^2 \equiv \epsilon, \qquad \eta_H \simeq -\epsilon + \eta, \qquad \eta \equiv \frac{V''}{\lambda_{\mathrm{P}}^2 V}, \tag{1.115}$$

and often used for the computation of the spectra of the metric perturbations amplified by inflation (see Section 8.2). The smallness of these two parameters guarantees the "flatness" of the potential $V(\phi)$, and the consequent slowness of the motion of ϕ towards the minimum.

The slow-roll equations (1.113) can be formally integrated, for any given $V(\phi)$, using the exact differential relations $\mathrm{d}a/a = H\mathrm{d}t$, $\mathrm{d}t = \mathrm{d}\phi/\dot{\phi}$, and writing the scale factor in the form

$$a(t) = a_{\rm i} \exp\left(\int_{t_{\rm i}}^{t} H\,{\rm d}t\right) = a_{\rm i} \exp\left(\int_{\phi_{\rm i}}^{\phi(t)} \frac{H\,{\rm d}\phi}{\dot{\phi}}\right)$$
$$= a_{\rm i} \exp\left(-\lambda_{\rm P}^2 \int_{\phi_{\rm i}}^{\phi(t)} \frac{V}{V'}{\rm d}\phi\right), \tag{1.116}$$

while $\phi(t)$ is obtained by integrating the equation

$$\dot{\phi} = -\frac{V'}{3H} = -\frac{1}{\sqrt{3}\lambda_{\rm P}} \frac{V'}{V^{1/2}}. \tag{1.117}$$

This solution is consistent provided the evolution of ϕ given by Eq. (1.117) is sufficiently slow ($\dot{\phi} \to 0$), and the scale factor (1.116) approximates the exponential de Sitter solution ($H \to$ const). A useful parameter, in such a context, is the number of e-folds $N(t)$ between a given time t and the end of inflation $t_{\rm f}$,

$$N(t) = \ln \frac{a_{\rm f}}{a(t)} = \int_{t}^{t_{\rm f}} H\,{\rm d}t. \tag{1.118}$$

Using Eq. (1.116) we can relate $N(t)$ to the corresponding value of the inflaton field at the same time t, namely,

$$N(t) = N(\phi(t)) = \lambda_{\rm P}^2 \int_{\phi_{\rm f}}^{\phi} \frac{V}{V'}{\rm d}\phi. \tag{1.119}$$

This relation will be applied in Chapter 8 to parametrize the primordial spectrum of metric perturbations obtained in the context of slow-roll inflation.

A simple example of an inflationary solution of the slow-roll type can be implemented, in practice, using an appropriate exponential potential [31],

$$V(\phi) = V_0\, {\rm e}^{-\lambda_{\rm P}\phi\sqrt{2/p}}, \tag{1.120}$$

where p and V_0 are positive parameters. In this case the Einstein equations (1.104) and (1.105) are solved by the particular exact solution

$$a = a_0 t^p,$$
$$\lambda_{\rm P}\phi = \sqrt{2p} \ln\left[\lambda_{\rm P} t \sqrt{\frac{V_0}{p(3p-1)}}\right], \tag{1.121}$$

which for $p > 1$ satisfies the kinematic conditions of power-law inflation (see Eq. (1.87)). The computation of H, $\dot{H}$, $\dot{\phi}$ and $\ddot{\phi}$ for this solution, together with the use of the exact definitions (1.109) and (1.110), leads to

$$\epsilon_H = \eta_H = p^{-1}. \tag{1.122}$$

For $p \gg 1$ the above solution (1.121) thus describes a phase of slow-roll inflation, which approaches de Sitter inflation in the limit $p \to \infty$.

Another efficient mechanism for generating slow-roll solutions is based on simple polynomial potentials of the type $V \sim \phi^n$, provided they are flat enough to satisfy the conditions $\epsilon \ll 1$ and $\eta \ll 1$ (a typical example is the so-called "chaotic" inflationary scenario [32], which includes the simplest case $n = 2$). For such potentials $V'/V = n/\phi$, and the total e-folding factor (computed from Eqs. (1.118) and (1.119)) takes the form

$$N = \ln\left(\frac{a_\mathrm{f}}{a_\mathrm{i}}\right) = \frac{\lambda_\mathrm{P}^2}{2n}(\phi_\mathrm{i}^2 - \phi_\mathrm{f}^2)\,. \tag{1.123}$$

Moreover, $V''/V \sim \phi^{-2}$, and the condition $\eta \ll 1$ requires very large values of the initial inflaton field in Planck units, $(\lambda_\mathrm{P}\phi_\mathrm{i})^2 \gg 1$, for the slow-roll regime to be valid. But this automatically guarantees an efficient inflationary expansion, $N \gg 1$, according to Eq. (1.123).

It must be noted that the slow-roll parameters associated with a polynomial potential are (slowly) evolving in time during inflation, in contrast with the case of the exponential potential where the parameters are constant (see Eq. (1.122)), and are in principle associated with an "eternal" duration of the phase of inflation. For models based on polynomial potentials the end of the inflationary phase may automatically occur as soon as the rolling velocity of the inflation increases, near to the minimum of the potential. In particular, when the effective mass V'' becomes of order H, the inflaton enters a regime of rapid oscillations characterized by the approximate equality of kinetic and potential energy, $\langle\dot{\phi}^2\rangle \simeq 2\langle V\rangle$. This regime preludes the inflaton decay and the consequent production of a cosmic background of relativistic particles, eventually becoming the dominant source of the standard, radiation-dominated era [5].

1.2.4 Initial singularity

A phase of slow-roll evolution, of the type illustrated by the above examples, seems to provide a more realistic (and probably even more natural) model of inflation than the one based upon the de Sitter solution, which requires instead a scalar field rigidly trapped at the minimum of its potential. Slow-roll solutions, however, do not describe a regular geometry like the de Sitter manifold, and therefore do not provide a solution to the singularity problem of the standard cosmological model. Indeed, the curvature decreases (even if slowly) during the slow-roll phase and this implies that, going backward in time, the Universe emerges from a singular state. The suppression of $\dot{H}$ during the slow-roll evolution moves the singularity backward in time towards much earlier epochs than in the standard scenario, but it does not remove it.

It should be noted, on the other hand, that even the exact inflationary solution (1.98), describing exponential expansion at constant curvature, does not completely remove the initial singularity. This solution, in fact, represents the de Sitter manifold in a chart that is *not* geodesically complete: a geodesic observer of such a coordinate system, starting from the origin, can reach a point at infinite spatial distance during a finite proper-time interval.

The geodesic incompleteness of the solution (1.98) is shown by recalling that the four-dimensional de Sitter manifold can be represented [33] as a pseudo-hypersphere (or hyperboloid) of radius $R_0 = \left(3M_P^2/\Lambda\right)^{1/2} = H^{-1} =$ const, embedded into a five-dimensional pseudo-Euclidean space with metric $\eta_{AB} = (+,-,-,-,-)$, spanned by the cartesian coordinates $z^A = (z^0, z^1, \dots z^4)$. The hyperboloid has equation

$$-\eta_{AB}z^A z^B = (z^i)^2 + (z^4)^2 - (z^0)^2 = H^{-2}, \tag{1.124}$$

where $A, B = 0, \dots 4$, and $i = 1, 2, 3$. The metric (1.98) can then be obtained by defining on the hyperboloid the intrinsic, four-dimensional cartesian chart $x^\mu = (t, x^i)$, and embedding the hypersurface into the higher-dimensional manifold through the following parametric equations,

$$\begin{aligned} z^i &= \mathrm{e}^{Ht} x^i, \\ z^0 &= \frac{1}{H}\sinh(Ht) + \frac{H}{2}\mathrm{e}^{Ht} x_i^2, \\ z^4 &= \frac{1}{H}\cosh(Ht) - \frac{H}{2}\mathrm{e}^{Ht} x_i^2, \end{aligned} \tag{1.125}$$

satisfying Eq. (1.124). Differentiating, and substituting into the five-dimensional form $\mathrm{d}s^2 = \eta_{AB}\,\mathrm{d}z^A\,\mathrm{d}z^B$, one obtains the line-element (1.98) with exponential scale factor. However, even for x_i and t ranging from $-\infty$ to $+\infty$, the given parametrization does not cover the full de Sitter manifold, but only a portion of it, defined by the condition $z^0 \geq -z^4$ (for $t \to -\infty$ one reaches the border of the parametrized region, marked by the null ray $z^0 = -z_4$).

A geodesically complete chart, covering the whole de Sitter hyperboloid, is obtained by considering a solution of Eq. (1.26) with $\rho = -p = \Lambda$ *and* with non-vanishing (constant, positive) spatial curvature $K = H^2 = (\Lambda/3M_P^2)$. The corresponding four-dimensional metric can then be written in the form

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \cosh^2(Ht)\left[\frac{\mathrm{d}r^2}{1 - H^2 r^2} + r^2\,\mathrm{d}\Omega^2\right], \tag{1.126}$$

and is related to the five-dimensional hyperboloid through the parametric equations

$$
\begin{aligned}
z^0 &= H^{-1}\sinh(Ht), \\
z^1 &= H^{-1}\cosh(Ht)\cos\chi, \\
z^2 &= H^{-1}\cosh(Ht)\sin\chi\cos\theta, \\
z^3 &= H^{-1}\cosh(Ht)\sin\chi\sin\theta\cos\phi, \\
z^4 &= H^{-1}\cosh(Ht)\sin\chi\sin\theta\sin\phi.
\end{aligned} \tag{1.127}
$$

Their differentiation, and substitution into the five-dimensional Minkowski form, leads to

$$
ds^2 = dt^2 - H^{-2}\cosh^2(Ht)\left[d\chi^2 + \sin^2\chi d\Omega^2\right], \tag{1.128}
$$

which reduces to Eq. (1.126) after setting $H^{-1}\sin\chi = r$. It is straightforward to check that the intrinsic chart $x^\mu = (t, \chi, \theta, \phi)$, with $-\infty \le t \le \infty$, $0 \le \chi \le \pi$, $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$, provides a full coverage of the hypersurface (1.124) (see for instance [33]).

By using the regular, complete de Sitter solution to eliminate the initial singularity we are led to a picture in which the primordial Universe enters a phase that can be extended (in a geodesically complete way) towards past infinity, according to the metric (1.126), keeping a constant, finite curvature controlled by Λ. However, the kinematic properties of such a phase are determined by the scale factor $a(t) = \cosh(Ht)$, describing a Universe which is initially contracting (at $t \to -\infty$), starting from an infinitely large spatial extension, and which becomes inflationary expanding only at large enough positive times, $t \to +\infty$.

Unfortunately, in models where the complete de Sitter solution is due to the potential energy of a scalar field satisfying standard causality and weak energy ($\rho \ge 0$) conditions, it seems impossible (using the Einstein equations) to include a smooth transition from the contracting to the expanding phase [34, 35]: starting from the exponentially contracting state, the Universe is doomed to collapse towards the singularity $a \to 0$, without “bouncing” to reach eventually the phase of accelerated expansion. In other words, known models of standard, potential-dominated inflation cannot be “past-eternal” [36].

Thus, for a successful model of de Sitter (or quasi de Sitter) potential-dominated inflation, the Universe has to enter the exponential regime already in the state of expansion. Such a state, as we have seen, cannot be arbitrarily extended backward in time without singularities, even in the case of the exact solution (1.98). We can say, therefore, that an inflationary phase driven by the potential energy of a scalar field mitigates the rapid growth of the curvature typical of the standard cosmological model, and shifts back in time the position of the initial singularity, without completely removing it, however (see Fig. 1.2)).

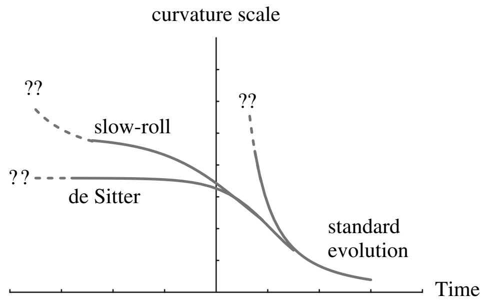

Figure 1.2 Qualitative evolution of the curvature scale in the standard cosmological model, and in models of de Sitter inflation and slow-roll inflation.

The problem of the initial configuration of the standard cosmological model, solved by inflation, then reappears (even if in a more relaxed form) for the inflationary phase, whose effectiveness still depends on the choice of an appropriate initial state. The question that arises is, in particular, the following: does there exist a dynamical mechanism able to "prepare" the appropriate initial inflationary state, producing (for instance) a homogeneous space-time domain that is already characterized by an exponential expansion (the "second half" of the de Sitter solution), and that can smoothly evolve towards the standard cosmological configuration?

One possible approach to this issue is provided by the methods of quantum cosmology (see Chapter 6). Using the Wheeler–De Witt equation [37, 38] it is possible to compute, for instance, the probability that our Universe emerges in the appropriate inflationary state directly from the vacuum (through a process conventionally called "tunneling from nothing" [39, 40, 41]). Such a probability, unfortunately, is strongly dependent on the initial quantum state representing the Universe before the transition, and this state is unknown, as it should be determined in correspondence with the initial singularity. There are various possible prescriptions for choosing the appropriate boundary conditions [39–43]: they are however "ad hoc", and lead to different (and strongly contrasting) results, leaving the debate still open.

Another possible, semiclassical approach is the one based on the "chaotic" inflationary scenario [32, 44]. In this approach the initial values of the scalar field are randomly distributed over different space-time regions, and those regions are characterized by different degrees of homogeneity with respect to the horizon scale. If, in some region, the scalar field happens to be sufficiently homogeneous, sufficiently large (in Planck units) and displaced from the minimum, then a phase of slow-roll inflation is triggered, and that initial region can evolve towards a configuration similar to the Universe in which we are living. In other space-time regions, where such conditions are not satisfied, inflation does not occur, and the subsequent evolution diverges from the path leading to the present cosmological state.

We should mention, finally, that even after a satisfactory explanation of the initial conditions, the scalar potential-dominated inflationary scenario suffers from other conceptual difficulties (see [45] for a recent discussion), such as the cosmological constant problem, the so-called "trans-Planckian" problem (see Section 5.3). String theory, as we shall see in the following chapters, may support inflationary mechanisms different from those based on the potential energy of a scalar field. As a consequence, different primordial scenarios are also possible, based on initial configurations other than the highly curved, hot and dense state approaching the initial singularity, typical of the standard model and of the inflationary models considered in this section.

References

[1] S. Weinberg, *Gravitation and Cosmology* (New York: John Wiley & Sons, 1972).
[2] J. B. Zeldovic and I. D. Novikov, *The Structure and Evolution of the Universe* (Chicago: University of Chicago Press, 1983).
[3] P. Coles and F. Lucchin, *Cosmology* (Chichester, UK: John Wiley & Sons, 2000).
[4] A. D. Linde, *Particle Physics and Inflationary Cosmology* (New York: Harwood, 1990).
[5] E. W. Kolb and M. S. Turner, *The Early Universe* (Redwood City, CA: Addison-Wesley, 1990).
[6] A. R. Liddle and D. H. Lyth, *Cosmological Inflation and Large-Scale Structure* (Cambridge, UK: Cambridge University Press, 2000).
[7] E. P. Hubble, *Proc. Nat. Acad. Sci.* **15** (1927) 168.
[8] A. A. Penzias and R. W. Wilson, *Ap. J.* **142** (1965) 419.
[9] G. W. Gibbons and S. Hawking, *Phys. Rev.* **D15** (1977) 2752.
[10] F. Sylos Labini, M. Montuori and L. Pietronero, *Phys. Rep.* **293** (1998) 61.
[11] W. Rindler, *Mon. Not. R. Astron. Soc.* **116** (1956) 6.
[12] Particle Data Group web pages at pdg.lbl.gov/.
[13] S. Pelmutter *et al.*, *Nature* **391** (1998) 51.
[14] A. G. Riess *et al.*, *Astronom. J.* **116** (1998) 1009.
[15] P. de Bernardis *et al.*, *Nature* **404** (2000) 955.
[16] S. Hanay *et al.*, *Ap. J. Lett.* **545** (2000) L5.
[17] N. W. Alverson *et al.*, *Ap. J.* **568** (2002) 38.
[18] D. N. Spergel *et al.*, *Ap. J. Suppl.* **148** (2003) 175.
[19] A. G. Riess *et al.*, *Ap. J.* **607** (2004) 665.
[20] D. N. Spergel *et al.*, astro-ph/0603449.
[21] P. B. Pal, *Pramana* **54** (2000) 79.
[22] R. R. Caldwell, M. Kamionkowski and N. N. Weinberg, *Phys. Rev. Lett.* **91** (2003) 07130.
[23] J. D. Barrow, *Class. Quantum Grav.* **21** (2004) 5619.
[24] S. W. Hawking and G. R. F. Ellis, *The Large Scale Structure of Space-Time* (Cambridge: Cambridge University Press, 1973).
[25] A. Guth, *Phys. Rev.* **D23** (1981) 347.
[26] A. D. Linde, *Phys. Lett.* **B108** (1982) 389.
[27] A. Albrecht and P. J. Steinhardt, *Phys. Rev. Lett.* **48** (1982) 1220.
[28] A. A. Starobinski, *Phys. Lett.* **B91** (1980) 99.
[29] D. S. Salopek and J. R. Bond, *Phys. Rev.* **D42** (1990) 3936.
[30] A. D. Liddle and D. Lyth, *Phys. Lett.* **B291** (1992) 391.

[31] F. Lucchin and S. Matarrese, *Phys. Rev.* **D32** (1985) 1316.
[32] A. D. Linde, *Phys. Lett.* **B129** (1983) 177.
[33] W. Rindler, *Essential Relativity* (New York: Van Nostrand Company, 1969).
[34] A. Vilenkin, *Phys. Rev.* **D46** (1992) 2355.
[35] A. Borde and A. Vilenkin, *Phys. Rev. Lett.* **72** (1994) 3305.
[36] A. Borde, A. Guth and A. Vilenkin, *Phys. Rev. Lett.* **90** (2003) 151301.
[37] J. A. Wheeler, in *Battelle Rencontres*, eds. C. De Witt and J. A. Wheeler (New York: Benjamin, 1968).
[38] B. S. De Witt, *Phys. Rev.* **D160** (1967) 1113.
[39] A. Vilenkin, *Phys. Rev.* **D30** (1984) 509.
[40] A. D. Linde, *Sov. Phys. JETP* **60** (1984) 211.
[41] V. A. Rubakov, *Phys. Lett.* **B148** (1984) 280.
[42] J. B. Hartle and S. W. Hawking, *Phys. Rev.* **D28** (1983) 2960.
[43] S. W. Hawking, *Nucl. Phys.* **B239** (1984) 257.
[44] A. D. Linde, *Phys. Lett.* **B351** (1995) 99.
[45] R. Brandenberger, in *Proc. YKIS 2005* (Yukawa Institute for Theoretical Physics, Kyoto, Japan, July 2005), hep-th/0509159.

2

The basic string cosmology equations

The aim of this chapter is to present the effective string theory equations governing the low-energy dynamics of the gravitational field and of its sources. Such equations are not postulated ad hoc but, as we shall see in the next chapter, they are required for the consistency of a quantum theory of strings propagating in a curved manifold, and interacting with other fields possibly present in the background. For a more systematic approach to these equations, the analysis of this chapter should probably follow the discussion of string quantization and the computation of the spectrum of the bosonic string states, which will be presented in Chapter 3. However, in the context of a cosmologically oriented book, we have preferred to postpone the string theory motivations in favor of a more immediate presentation of the basic string gravity equations, lying at the foundations of string cosmology just like the Einstein equations are at the foundations of standard cosmology.

For our purposes we only need to recall that the exact string theory equations, for all fields (including gravity) present in the string spectrum, can be approximated by a perturbative expansion, in general in two ways [1]: (*i*) as a higher-derivatives expansion (namely, as an expansion in powers of the "curvatures", or field strengths), and (*ii*) as an expansion in powers of the coupling parameter g_s^2, controlling the intensity of the string interactions. This second expansion is similar to the "loop" expansion of conventional quantum field theory, while the first one is peculiar to strings, since it is controlled by the fundamental length λ_s appearing in the (two-dimensional) string action integral (see Chapter 3); such an expansion disappears in the point-particle limit $\lambda_s \to 0$.

The discussion of this chapter will concentrate on the tree-level equations for the fundamental massless (boson) fields present in all models of strings and superstrings [1, 2] (here "tree-level" means that the equations are truncated to lowest order in both the curvature expansion, controlled by the parameter $2\pi\alpha' = \lambda_s^2$, and the loop expansion, controlled by g_s^2). Such equations can be derived from an effective action which is valid for manifolds with low enough curvature, $\alpha' R \ll 1$, and for fields with weak enough interactions, $g_s^2 \ll 1$ (the action is valid, therefore, in the low-energy, perturbative regime). Nevertheless, they can be basic equations even in a primordial cosmological context where there are scenarios – possibly suggested by string-duality symmetries – with perturbative initial configurations, well described by the low-energy equations [3] (see Chapter 4). Also, such equations are used in the context of the so-called "string gas cosmology" [4] that will be discussed in Chapter 6.

These low-energy equations will be explicitly derived from the action in the string frame (Section 2.1) and in the Einstein frame (Section 2.2). In the last section (and in the appendix) we will discuss the possible corrections induced by the addition of quadratic curvature terms to the effective gravitational equations, to first order in the α' expansion.

2.1 Tree-level equations

The gravitational (massless, bosonic) sector of the string effective action contains not only the metric but also (and even to lowest order) at least one more fundamental field: a scalar field ϕ, called the "dilaton". The corresponding tree-level action can be written as follows:

$$S = -\frac{1}{2\lambda_s^{d-1}} \int d^{d+1}x \sqrt{|g|}\, e^{-\phi} \left[R + (\nabla\phi)^2 \right] + S_\Sigma + S_m. \tag{2.1}$$

Here S_Σ is the boundary term required to reproduce the standard Einstein equations in the general-relativistic limit, and S_m is the action of all other fields, possibly coupled to ϕ and to $g_{\mu\nu}$ as prescribed by the conformal invariance of fundamental string interactions (see the discussion of the next chapter). Note that we have used (and we shall often use) the compact notation $(\nabla\phi)^2 = \nabla_\mu\phi\nabla^\mu\phi$.

The above equation is written adopting the so-called "string frame" (S-frame) parametrization of the action, where ϕ is dimensionless, and where the metric $g_{\mu\nu}$ is the same metric to which a fundamental string is minimally coupled, and with respect to which a free "test" string evolves geodesically. Otherwise stated, the action (2.1) is parametrized by the same metric field present in the two-dimensional action integral governing the motion of a fundamental string in a curved background (as illustrated in Chapter 3).

It should be noted, also, that we have generically considered the action for a $D = (d+1)$-dimensional space-time manifold. As we shall see later, the quantum theory of an extended object like a string can be consistently formulated only if the number of dimensions is fixed at a critical value $D = D_{\rm crit}$ (for instance, $D_{\rm crit} = 26$ for the bosonic string, $D_{\rm crit} = 10$ for a superstring [1, 2]). We will often also consider a number of dimensions less than critical – in particular,

$D=4$ – assuming, in that case, that the background fields have a factorizable structure, that the integral over the remaining $D_{\rm crit}-D$ spatial dimensions gives only a trivial (finite) volume factor, and that such an extra factor has been absorbed by an appropriate rescaling of the dilaton.

The constant length $\lambda_{\rm s}$ appearing in the action (2.1) represents the characteristic proper extension of a quantized one-dimensional object like a fundamental string, and provides the natural units of length ($\lambda_{\rm s}$) and energy ($\lambda_{\rm s}^{-1}=M_{\rm s}$) for a physical model based on the S-frame action [5]. The comparison with the $(d+1)$-dimensional Einstein–Hilbert action,

$$S=-\frac{1}{2\lambda_{\rm P}^{d-1}}\int {\rm d}^{d+1}x\sqrt{|g|}\;R, \tag{2.2}$$

immediately provides the (tree-level) relation between the string length and the Planck length,

$$\frac{\lambda_{\rm P}}{\lambda_{\rm s}}=\frac{M_{\rm s}}{M_{\rm P}}={\rm e}^{\phi/(d-1)}, \tag{2.3}$$

which clearly shows how the effective gravitational coupling, $8\pi G_D\equiv\lambda_{\rm P}^{d-1}$, is controlled by the dilaton, in string units, as $8\pi G_D=\lambda_{\rm s}^{d-1}\exp\phi$.

For a ϕ-independent matter action S_m, the action (2.1) would seem to describe a scalar-tensor model of gravity of the Brans–Dicke (BD) type, with BD parameter $\omega=-1$. In fact, if we set

$$\frac{{\rm e}^{-\phi}}{\lambda_{\rm s}^{d-1}}=\frac{\Phi}{8\pi G_D}, \tag{2.4}$$

the gravi-dilaton part of the action can be rewritten in the "canonical" BD form (see for instance [6]),

$$S_{\rm BD}=\frac{1}{8\pi G_D}\int {\rm d}^{d+1}x\sqrt{|g|}\left[-\Phi R+\omega\Phi^{-1}\left(\nabla\Phi\right)^2\right], \tag{2.5}$$

provided ω is fixed to the value -1.

Even for the gravi-dilaton sector, however, the analogy with a "pure" BD model is possibly valid only at the tree-level: in fact, after including the higher-loop corrections required by string theory in the strong coupling limit, the effective action may be rewritten in the form (2.5), only at the cost of defining a BD parameter which is dilaton dependent, $\omega=\omega(\Phi)$ (see for instance [7, 8]). In addition, the tree-level analogy with BD models only holds for a particular class of fields, whose S-frame action S_m is decoupled from the dilaton (for example, for the bosonic forms present in the Ramond–Ramond sector of type IIA and type IIB superstrings, see e.g. [1, 2] and Appendix 3B). String theory, in general, predicts a *non-minimal* and *non-universal* coupling of the various fields to the dilaton (see Chapter 9): it is thus impossible, in principle (even at the tree-level), to