1

Overview

Although cosmology can trace its beginnings back to Einstein's formulation of his general theory of relativity in 1915, which enabled the first mathematically consistent models of the Universe to be constructed, for most of the following century there was much uncertainty and debate about how to describe our Universe. Over those years the various necessary ingredients were introduced, such as the existence of dark matter, of the hot early phase of the Universe, of cosmological inflation, and eventually dark energy. In the latter part of the last century, cosmologists and their funding agencies came to realize the opportunity to deploy more ambitious observational programmes, both on the ground and on satellites, which began to bear fruit from 1990 onwards. The result is a golden age of cosmology, with the creation and observational verification of the first detailed models of our Universe, and an optimism that that description may survive far into the future. The objective, often described as *precision cosmology*, is to pin down the Universe's properties as best as possible, in many cases at the percent or few percent level. In particular, the landmark publication in 2003 of measurements of the cosmic microwave background by the Wilkinson Microwave Anisotropy probe (WMAP), seems certain to be identified as the moment when the Standard Cosmological Model became firmly established.

The key tool in understanding our Universe is the formation and evolution of structure in the Universe, from its early generation as the primordial density perturbation to its gravitational collapse to form galaxies. As we already argued in the introduction to our book *Cosmological Inflation and Large-Scale Structure*, the complete theory of structure formation, starting with the quantum fluctuations of a free field, continuing with general-relativistic gas dynamics, and ending with the free fall of photons and matter, is perhaps one of the most beautiful and complete in the entire field of physics. It has also demonstrated powerful predictive power, for instance anticipating the oscillatory structure of the cosmic microwave anisotropy spectrum more than twenty years before the anisotropies were measured in any form, and in detail making percent-level predictions that continue to be in accord with what are now percent-level observations.

The purpose of this book is to give a detailed account of the physics of density perturbations in the Universe, focussed around the form and implications of the primordial perturbation. We aim to describe the main astrophysical processes which transform the initial density perturbation into observables, such as the cosmic microwave anisotropies, and to show how these observable consequences can be tracked back to an origin which sheds light on fundamental physical processes in the early Universe.

The book is divided into parts, as follows.

Part I: Relativity gives the basics of general relativity along with the applications needed for cosmology, starting from a basic knowledge of special relativity.

Part II: The Universe after the first second concerns itself with the evolution of perturbations, starting with a primordial density perturbation, whose existence is at this point taken for granted. After a brief overview of the theory of the background (homogeneous) cosmology, density perturbations are defined and characterized, and their evolution studied in both Newtonian and relativistic frameworks. This evolution ultimately leads to the observable consequences of the theory.

Part III: Field theory sets the context for explaining the origin of the primordial density perturbation in terms of fundamental physics. It gives those aspects of field theory that are needed for Part IV, starting from a basic knowledge of quantum mechanics. Among the key ideas developed are scalar field dynamics, internal symmetry, supersymmetry, and the quantization of free fields.

Part IV: Inflation and the early Universe exploits these ideas to explain the leading theory for the origin of perturbations, cosmological inflation. We describe a number of variants on the basic inflationary theme. We conclude by developing the observational consequences of a wide range of inflationary scenarios, setting the challenge to distinguish amongst them using future observations.

The reader will notice that many references are given for the chapters of Part IV while very few are given for earlier chapters. This reflects a profound difference between the material in Part IV and that in Parts I–III. The theories covered in the first three parts have been around for at least several years, and in many cases for far longer. It is true that Nature may have chosen not use some of them. There may be no significant tensor or isocurvature perturbation, no supersymmetry, no axion, and no seesaw mechanism for neutrino masses. But the theories themselves are well established. As a result, most of the additional material consulted by the reader will consist of texts and reviews as opposed to topical research articles. The most appropriate sources of that kind will depend on the reader's background and future intentions, and we mention only a few possibilities.

In contrast, the study of the very early Universe covered in Part IV is at the cutting edge of current research. It is not covered in any text at present, and the coverage of reviews is quite patchy. The situation is also quite complicated, with a large menu of possibilities confronting many different kinds of observation. What we have done in Part IV is to get the reader started on a study of the main possibilities, pointing along the way to reviews and research papers that can be the basis of further study.

Notes on exercises

Most chapters end with a few exercises to allow the reader to practice applying the information given within the chapter. Several of these examples require some simple numerical calculations for their solution; in cosmology these days, it is practically impossible to avoid carrying out some numerical work at some stage. A typical task is the numerical computation of an integral that cannot be done analytically, or the evaluation of some special functions. These can be done via specially written programs, using library packages (e.g., *Numerical Recipes* [1], which is also an invaluable source of general information on scientific computation), or a computer algebra package such as Mathematica or Maple.

Units

In keeping with conventional notation in cosmology, we set the speed of light c equal to one, so that all velocities are measured as fractions of c. Where relevant, we also set the Planck constant $\hbar$ to one, so that there is only one independent mechanical unit. In particular, the phrases 'mass density' and 'energy density' become interchangeable. Often it is convenient to take this unit as energy, and we usually set the Boltzmann constant $k_{\rm B}$ equal to 1 so that temperature too is measured in energy units. (In normal units $k_{\rm B} = 8.618 \times 10^{-5}\,{\rm eV\,K^{-1}}$.)

Newton's gravitational constant G can be used to define the **reduced Planck mass** $M_{\rm Pl} = (8\pi G)^{-1/2}$. Thought of as a mass, $M_{\rm Pl} = 4.342 \times 10^{-6}$ g, which converts into an energy of 2.436×10^{18} GeV. We use the reduced Planck mass throughout, normally omitting the word 'reduced'. It is a factor $\sqrt{8\pi}$ less than the alternative definition of the Planck mass, never used in this book, which gives $m_{\rm Pl} = 1.22 \times 10^{19}$ GeV. We use $M_{\rm Pl}$ and G interchangeably, depending on the context. Inserting appropriate combinations of $\hbar$ and c, we also can obtain the reduced Planck time $T_{\rm Pl} \equiv \hbar/c^2 M_{\rm Pl} = 2.70 \times 10^{-43}$ s and reduced Planck length $L_{\rm Pl} \equiv \hbar/cM_{\rm Pl} = 8.10 \times 10^{-33}$ cm.

Reference

[1] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery. *Numerical Recipes: The Art of Scientific Computation*, 3rd edition (Cambridge: Cambridge University Press, 2007).

Part I

Relativity

2

Special relativity

In this chapter we review the basic special relativity formalism, expressed in a way that is readily extended to general relativity in the following chapter. We then study the energy–momentum tensor, both for a generic fluid and for a gas.

2.1 Minkowski coordinates and the relativity principle

A starting point for relativity is provided by the interval ds^2 between neighbouring points of spacetime, known as **events**. The interval may be regarded as a given concept, like the distance between two points in space.

Special relativity assumes the existence of coordinates in which

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \,. \tag{2.1}$$

(In a different convention, the sign of ds^2 is reversed.) Such coordinates are called **Minkowski coordinates**. They are also referred to as an inertial frame. The spatial coordinates are Cartesian, and the distance $d\ell$ between two points in space with same time coordinates is given by

$$d\ell^2 = ds^2 = dx^2 + dy^2 + dz^2. \tag{2.2}$$

It is convenient to use the index notation $(x^0, x^1, x^2, x^3) \equiv (t, x, y, z)$, and to denote a generic coordinate by x^μ.[1] We also use the notation $x^\mu = (t, \mathbf{x})$. Then we can define a **metric tensor** $\eta_{\mu\nu}$ as the diagonal matrix with elements $(-1, 1, 1, 1)$. Using it, Eq. (2.1) can be written

$$ds^2 = \eta_{\mu\nu}\, dx^\mu dx^\nu . \tag{2.3}$$

We adopt the summation convention: that there is a sum over every pair of identical

[1] We take Greek letters $(\mu, \nu, \dots)$ to run over the values $0, 1, 2, 3$, and italic letters $(i, j, \dots)$ to run over the values $1, 2, 3$.

spacetime indices. A summation of the above form, over the metric tensor, is called a contraction. In this case, μ and ν are contracted.

As in this example, sums over spacetime indices involve one upper index and one lower index, and spacetime coordinates always carry an upper index. In contrast, if an expression involves purely spatial Cartesian coordinates, one can take all indices as lower while still adopting the summation convention. In particular, the distance $d\ell$ between nearby points at a given time is given by

$$d\ell^2 = ds^2 = \delta_{ij} dx_i dx_j , \tag{2.4}$$

where δ_{ij} is the Kronecker delta, equal to 1 for equal indices and to 0 for unequal ones. (We will also denote it by δ^{ij} or δ^i_j according to the context.)

A transformation taking us from one inertial frame to another preserves the form of Eq. (2.3). The time-reversal transformation $t' = -t$ does this, but we fix the sign of t so that it increases going from past to future. A parity transformation, changing right-handedness to left-handedness, also does it, but we choose a right-handed coordinate system. (The parity transformation can be taken as a reversal of all three coordinates, or of just one of them.) With these restrictions, the transformations preserving the form of Eq. (2.1) are a translation of the spacetime origin and/or a Lorentz transformation. This is the Poincaré group of transformations.

A spacetime translation corresponds to new coordinates

$$x'^\mu = x^\mu + X^\mu , \tag{2.5}$$

with X^μ a constant. A Lorentz transformation corresponds to new coordinates

$$x'^\mu = \Lambda^\mu_{\ \nu} x^\nu , \tag{2.6}$$

with Λ^μ_ν a constant matrix satisfying

$$\Lambda^\alpha_{\ \mu} \Lambda^\beta_{\ \nu} \eta_{\alpha\beta} = \eta_{\mu\nu} . \tag{2.7}$$

Note that

$$\Lambda^\mu_{\ \nu} = \frac{\partial x'^\mu}{\partial x^\nu} . \tag{2.8}$$

It is often convenient to consider an infinitesimal Lorentz transformation,

$$\Lambda^\mu_{\ \nu} = \delta^\mu_\nu + \omega^\mu_{\ \nu} . \tag{2.9}$$

Requiring that the form of Eq. (2.3) is preserved, one sees that this is a Lorentz transformation if and only if the infinitesimal quantity $\omega_{\mu\nu} \equiv \eta_{\mu\alpha} \omega^\alpha_{\ \nu}$ is antisymmetric.

The most general Lorentz transformation is a rotation and/or a Lorentz boost. A rotation changes just the space coordinates and preserves the form of Eq. (2.4). It has the form

$$x'_i = R_{ij} x_j, \tag{2.10}$$

where the rotation matrix is orthogonal ($R^{-1} = R^{\mathrm{T}}$). For a rotation through angle ϕ about the z-axis,

$$R_{ij} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.11}$$

A Lorentz boost along (say) the x-axis mixes t and x components according to

$$x' = \gamma(x - vt), \qquad t' = \gamma(t - vx), \tag{2.12}$$

where $\gamma = (1 - v^2)^{-1/2}$ (recall that we set $c = 1$). The parameter v gives the relative velocity of the old and new frames.

Allowing only boosts with $v \ll 1$ gives the Galilean transformation $t' = t$ and $x' = x - vt$. This is the Newtonian description of spacetime, in which there is a universal time coordinate.

Special relativity ignores gravity and assumes the existence of Minkowski coordinates, which define an inertial frame. According to the **relativity principle** as originally formulated there is no preferred inertial frame, which means that the form of the equations remains the same under every transformation from one set of Minkowski coordinates to another. This was supposed to include time reversal (T) $t \to -t$, and the parity transformation (P) $x^i \to -x^i$ which reverses the handedness of the coordinates. According to quantum field theory, these transformations are related to charge conjugation (C), which interchanges particles with antiparticles in such a way that the combined CPT invariance is guaranteed, and this is verified by observation.

It was found in 1956 that the weak interaction is not invariant under the parity transformation P, though it seemed to be invariant under CP or equivalently T. In 1964 it was found that even these invariances are not exact. Therefore, at a fundamental level, the relativity principle should be applied only after t has been chosen to increase into the future, and (say) a right-handed coordinate system has been specified. This turns out not to be an issue in the usual scenarios of the early Universe perturbations, because in those scenarios there is no mechanism by which the violation of T and P leads to an observable effect. Scenarios have been proposed where that is not the case, so that for example the cosmic microwave background has net left-handed circular polarization, but no such effect has been observed.

2.2 Vectors and tensors with Minkowski coordinates

According to the relativity principle, the laws of physics should take on the same form in every inertial frame. To achieve this, the equations are written in terms of 4-scalars, 4-vectors and 4-tensors. These objects are invariant under spacetime translations, and they transform linearly under the Lorentz transformation.

2.2.1 4-scalars and 4-vectors

A 4-scalar is specified by a single number, and is invariant under the Lorentz transformation.[2] A 4-vector A^μ is specified by four components, transforming like dx^μ:

$$A'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} A^\nu. \tag{2.13}$$

As with ordinary vectors one can use a symbol to denote the vector itself as opposed to its components, such as $\vec{A}$. Also, one can define basis vectors $\vec{e}_\mu$ such that

$$\vec{A} = \sum_\mu A^\mu \vec{e}_\mu. \tag{2.14}$$

In a given inertial frame, each 4-vector is of the form $A^\mu = (A^0, \mathbf{A}) = (A^0, A^i)$ where as usual $\mathbf{A}$ denotes a 3-vector. The inner product of two 4-vectors is defined as $\eta_{\mu\nu} A^\mu B^\nu$. It is a 4-scalar and the 4-vectors are said to be orthogonal if it vanishes. From now on we generally refer to a 4-vector simply as a vector.

For any vector, it is useful to define the lower-component object

$$A_\mu = \eta_{\mu\nu} A^\nu, \tag{2.15}$$

or more explicitly, $A_0 = -A^0$, $A_i = A^i$. It is called a covariant vector while A^μ is called a contravariant vector (more properly one should talk about the covariant or contravariant components of the same vector). The scalar product of two vectors can be written $A_\mu B^\mu$. Going to a new inertial frame,

$$A'_\mu B'^\mu = A_\mu B^\mu. \tag{2.16}$$

The transformation law for A_μ follows from this, remembering that B'^μ is an arbitrary vector. Indeed, let us set every component of that object equal to zero except for one component $B'^\mu = 1$ (with μ either 0, 1, 2 or 3). Inserting this on the left-hand side and putting the transformed quantity $B^\nu = (\partial x^\nu / \partial x'^\mu) B'^\mu$ into the right-hand side, we get

$$A'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} A_\nu. \tag{2.17}$$

[2] If we consider in addition the effect of a parity transformation, we need to distinguish between true scalars (simply called scalars), which are invariant under the parity transformation, and pseudo-scalars which reverse their sign. A similar distinction needs to be made for vectors and tensors. We shall have occasional need of this distinction.

Cambridge University Press
978-0-521-82849-9 - The Primordial Density Perturbation: Cosmology, Inflation and the Origin of Structure
David H. Lyth and Andrew R. Liddle
Excerpt
More information

2.2.2 4-tensors

A second-rank 4-tensor is a sixteen-component object $C_{\mu\nu}$, which transforms like a product $A_\mu B_\nu$ of two vectors;

$$C'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}C_{\alpha\beta}. \tag{2.18}$$

Third- and higher-rank 4-tensors are defined in the same way. We can lower any component of a tensor with $\eta_{\mu\nu}$, for instance, $C^\mu{}_\nu = \eta_{\nu\lambda}C^{\mu\lambda}$. A tensor with all lower indices transforms like a product of covariant vectors and a tensor with mixed upper and lower indices transforms like a product of the appropriate mixture of vectors.

It is often useful to regard 4-vectors and 4-scalars as, respectively, first- and zeroth-rank tensors. Depending on the context, the term 'tensor' will either include those, or will mean an object of rank ≥ 2.

As with vectors, one can denote (say) a second-rank tensor by $\vec{C}$. One can also define basis tensors $\vec{e}_\mu \otimes \vec{e}_\nu$, such that

$$\vec{C} = \sum_{\mu,\nu} C^{\mu\nu}\vec{e}_\mu \otimes \vec{e}_\nu. \tag{2.19}$$

If we multiply two tensors and contract any number of indices we get another tensor; for instance $A_{\mu\nu}B^{\mu\alpha}$ is a tensor. An inverse of this statement is true; if the multiplication and contraction of an object with an *arbitrary* tensor gives another tensor, then that object is itself a tensor. We encountered a special case of this 'quotient theorem' in Eq. (2.16). As in that case, it can be proved simply by setting in turn each component of the arbitrary tensor to 1 with the others zero.

From Eq. (2.7), the metric tensor $\eta_{\mu\nu}$ is indeed a tensor, but a very special one. Its components are the same in every inertial frame, and $\eta^\mu{}_\nu = \eta_\nu{}^\mu = \delta^\mu_\nu$, where δ^μ_ν is again the Kronecker delta. Also, $\eta^{\mu\nu}$ has the same components as $\eta_{\mu\nu}$. The Levi–Civita symbol $\epsilon_{\mu\nu\alpha\beta}$ (totally antisymmetric with $\epsilon_{0123} = 1$) also has the same components in every frame. Any other tensor with the same components in every frame has to be constructed by multiplying and/or contracting these two.

If the components of a 4-vector or 4-tensor vanish in one coordinate system, they vanish in all coordinate systems. This means that a 4-vector or 4-tensor is defined uniquely by giving its components in any coordinate system.

In order to satisfy the principle of relativity, laws of physics are written in the form 'tensor = tensor', or in other words in the form 'tensor $= 0$'. Such an equation is said to be **covariant**; both sides transform in the same way. [3]

[3] At the quantum level one can also allow 'spinor' $= 0$, where the transformation of spinor components is given by Eq. (15.48).