1
The purpose of this book

For many years after its appearance, general relativity (GR) was regarded as an exotic extension of Newtonian gravity, that was only necessary for high-precision measurements in the Solar System and for describing the expansion of the Universe. However, the increasing precision of physical and astronomical measurement is transforming GR into an indispensable tool, and not merely a small correction to Newton's theory.

It is commonly stated that we have entered the era of precision cosmology, in which a number of important observations have reached a degree of precision, and a level of agreement with theory, that is comparable with many Earth-based physics experiments. One of the consequences of this advance is the need to examine at what point our usual, well-worn assumptions begin to compromise the accuracy of our models, and whether more general theoretical methods are needed to maintain calculational accuracy. Historically, each advance in astronomical measurement has produced many new discoveries, and revealed more of the structure of the cosmos, such as voids, walls, filaments, etc. As we map out the Universe around us – its mass distribution and flow patterns – in ever greater detail, the nonlinear behaviour of cosmic structures will become increasingly apparent, and the methods of inhomogeneous cosmology will come into their own. Inhomogeneous solutions of Einstein's field equations provide models of both small and large structures that are fully nonlinear.

It is widely assumed that the Universe, when viewed on a large enough scale, is homogeneous and can be described by an FLRW model. The successes of the Concordance model are built on using a spatially homogeneous and isotropic background metric combined with first-order perturbation theory. Although this assumption has been appropriate up to now, and underlies many important developments in cosmology, it is not the whole story. The modern successes have led to considerable hubris and overconfidence that we have pinned down the matter content and evolution of the key epochs of the post-grand unification Universe, even though all admit we don't know the underlying physics of dark matter and dark energy. Therefore, despite these successes, we must not lose sight of the fact that the present-day Universe is actually very inhomogeneous. If we keep this fact out of our minds, then we ignore knowledge and techniques that will be essential in understanding the real Universe and its multitudinous components in the era of precision cosmology. The relationship between a lumpy universe and an averaged homogeneous one is still not all that well understood, though there are some promising investigations (Buchert, 2000, 2001; Buchert and Carfora, 2002, 2003; Räsänen, 2006; Buchert, 2008; Ellis, 2008; Leith, Ng, and Wiltshire, 2008). The assumption of homogeneity – so essential in developing the basics of cosmology – must now be considered just a zeroth-order approximation; and similarly linear perturbation theory a first-order approximation, whose domain of validity is an early, nearly homogeneous Universe.

The use of perturbations relies on two myths that we wish to briefly discuss here:

1. Since for galaxy clusters the gravitational potential ϕ obeys $\phi/c^2 \ll 1$ (typically $\phi/c^2 \approx 10^{-5}$), the spacetime is nearly flat and linear perturbations work to sufficient accuracy.

2. Where linear perturbations become unsatisfactory, one can go to higher orders of perturbation.

This is how these myths disagree with reality:

1. The fact is that the quantity ϕ/c^2 is a measure of the so-called curvature contrast (perturbation of the 3-dimensional curvature of space divided by the local value of the curvature). But in order to remain safely within the 'linear regime', *both* the curvature contrast and the density contrast, $\Delta\rho/\rho$, must remain small. Moreover, the curvature contrast is a very imprecise indicator of the quality of approximation, since, for example in the Lemaître–Tolman model with a negative curvature background, it is decreasing with time, irrespective of the initial conditions (see Sec. 18.10 in Plebański and Krasiński, 2006), while the density contrast is increasing. Thus, results of such approximate calculations cannot be safely extrapolated into the future. At present, for various galaxy clusters $\Delta\rho/\rho$ is contained between about 5 and about 4000 (see Table 4.1 here), and still much more for smaller structures, so we are already outside the range where the series of approximations can be expected to converge to an exact result. If an exact calculation gives a result that is close to one obtained earlier by a perturbative method, then this is a confirmation that should be welcomed rather than criticised by saying 'it was obvious that the result had to be close to the earlier-calculated one, so the exact treatment was needless.'

2. Higher-order perturbations can improve the accuracy achieved in the first order only when we are still within the range of convergence of the series of approximations (or, in astronomical terminology, within the 'linear regime'). When we are outside that range, second-order corrections will turn out to be larger than the first-order result, and thus worthless. Indeed, second-order calculations are so complex, and involve so many terms, that investigations must focus on particular phenomena and set all other unrelated terms to zero.

3. Although various assertions are made about the requirements for perturbation theory to be valid, we are not aware of any proof of the domain of validity at linear or any other order, nor any proof of convergence. There do seem to be significant difficulties and uncertainties with the method, as discussed in the

following examples. Van Elst *et al.* (1997) show that most irrotational silent cosmological models have a linearisation instability. Bruna and Girbau (1999, 2005) analysed the linearisation stability of Robertson–Walker models, and showed the closed model is not linearisation stable. Notari (2006) suggests that, due to the self-gravitation of perturbations, their contribution at the present day has grown much larger than expected from standard perturbation theory. In a brief review of nonlinear methods for cosmological perturbations, Matarrese (1996) concludes that there are effects that violate the standard wisdom about when relativistic effects are important. Crocce and Scoccimarro (2006) investigate terms up to third order in a perturbation expansion, and they show that terms of successive orders can have different signs, and that second- and third-order terms can exceed first-order terms at larger wavelengths (see their Fig. 1). To address this problem, they introduce renormalised perturbation theory, and though in the examples given the new approach keeps all its terms positive, and higher-order terms contribute at ever smaller scales, the convergence of the renormalised series is also unproven. Losic and Unruh (2008) and Unruh and Losic (2008) find that second-order perturbations could become stronger than first order in models of slow-roll inflation. In the 'nonlinear regime' there is just no escape from exact methods.

On the other hand, the observational data do allow plenty of scope for explanations based on inhomogeneous models. This fact is often hidden under layers of vigorous advertising for traditional methods, but is slowly coming to light – see the broad and impartial review by Sarkar (2008). Similar comments, although strongly biased toward averaging techniques seen as the ultimate best method, can be found in the papers by Wiltshire (2007).

In fact, recent works have tried to combine the averaging method 'à la Buchert' with cosmological linear perturbation theory, claimed to be valid at scales where nonlinear evolution is supposed to enter into play, to study the 'backreaction' of inhomogeneities on a 'homogeneous background' (Li and Schwarz, 2007, 2008). This shows that the cosmological community is becoming aware that the effects of inhomogeneities must be taken into account even if the methods to do so are still in their infancy.

More reasons why approximate results are uncertain, misleading or not fully useful will be given later in this book.

During the 80 years after the Friedmann–Lemaître–Robertson–Walker papers, relativity has advanced much farther on many fronts and produced results that can be directly used in interpreting cosmological observations. In the present text we demonstrate several examples of such applications of exact methods of relativity, mostly (but not exclusively) taken from our own papers. This line of research has a tradition going back to the 1930s, and many powerful exact results had been derived long ago – they are presented systematically in the review by Krasiński (1997). Here we concentrate on the papers that relied on modern astronomical data.

In the first part of this book we present those elements of exact relativistic cosmology that are directly applicable to the interpretation of observations. We assume that the reader is familiar with relativity at the graduate level. We describe the basic properties of three classes of models that have already proved usable: (I) The Lemaître (1933)–Tolman (1934) (L–T) model, which is an exact solution of Einstein's equations with spherical symmetry and dust source – the simplest generalisation of the Friedmann models. (II) The model for which we propose the name 'Lemaître model' – for good historical reasons, see Lemaître (1933) and Sec. 2.2, but which is known in the literature as the Misner–Sharp (1964) approach. This is not an explicit solution, but a metric whose components obey a set of two equations. This scheme, well suited to numerical treatment, describes the evolution of a spherically symmetric perfect fluid with nonzero pressure gradient. (III) The Szekeres (1975) model, which is an exact dust solution with no symmetry, generalising L–T. This is currently the most sophisticated known exact solution of Einstein's equations of cosmological relevance. In Part I of the book we also introduce the relativistic description of light propagation, which includes a detailed presentation of apparent and event horizons in the L–T model. This is a necessary introduction to the discussion of formation and evolution of nonstationary black holes.

In Part II we then present and discuss the applications of exact relativistic methods to actual problems of observational cosmology. These include:

1. **Formation of a galaxy with a central black hole.** This is a problem for whose solution exact methods of relativity are essential, since it is impossible even to define a black hole in a meaningful way by Newtonian methods or perturbations of an FLRW background. Since this was the first investigation of this kind, we used the simplest model that was applicable to it – the L–T model. Its spherical symmetry is a drawback, since most galaxies are not spherically symmetric, and they rotate in addition. However, our investigation may be useful for preliminary qualitative understanding of the process. We considered two possible mechanisms of formation of such an entity: a gravitational collapse of an ordinary ball of dust, and a condensation forming around a pre-existing wormhole.

   We aimed to reproduce the density distribution and mass of one actually observed galaxy with a central black hole, M87. For this, we used the density profile in that galaxy, believed to be known from observations, and for the density profile within the growing black hole, about which nothing is and nothing can be known from observations, we chose a simple model, joined onto the galaxy profile.

   To set realistic initial conditions we assumed that by the time of decoupling of matter and radiation the condensation that would later grow to become the galaxy had a density contrast consistent with the implications of measurements of temperature anisotropies of the CMB radiation. There is a problem with this that keeps coming up in several other investigations: the currently best achieved angular resolution of measurements of temperature is $0.1°$, while a typical proto-galaxy would occupy a region of angular diameter of only $0.004°$ on the CMB sky (see Table 4.1). Thus, at present there are just no adequate observational data to constrain our model. Lacking any better choice, we took the limits known for the scale of $0.1°$ to apply to a single galaxy.

It turned out that the galaxy–black hole structure can be generated by both mechanisms, but for the condensation around a wormhole the black hole appears in a much shorter time, almost instantly after the Big Bang. In that section we stressed that the horizon whose position can be approximately determined from observations is the apparent horizon, not the event horizon.

2. **Formation and evolution of galaxy clusters and voids, using the L–T model.** Inhomogeneities are naturally generated in an inhomogeneous model by the form of the arbitrary functions defining the model. We did not therefore discuss their origin.

   Rather, we inferred the amplitude of perturbations at the time of decoupling from the constraints imposed by the measured temperature fluctuations of the CMB radiation. Then we showed that perturbations at decoupling that are consistent with observational constraints can generate a galaxy cluster whose calculated parameters corresponding to the present epoch (1.5×10^{10} y after the Big Bang) agree with the observed parameters of a galaxy cluster chosen at random from the Abell catalogue. Here we encountered again the problem mentioned above: the observational upper limit on the temperature anisotropy of the CMB radiation is determined at angular scales about 20 times larger than the angular diameter that an Abell proto-cluster would occupy on the CMB sky. With better resolutions sure to be achieved in the future, our constraints will have to be re-evaluated. For voids, by the same method, we achieved a qualitative agreement with observations, but for the density within a void we obtained values several times larger than those observed.

   The presence of non-baryonic dark matter (which does not interact with photons) would solve this discrepancy, since at the last scattering instant the amplitude of dark matter fluctuations could be larger than temperature fluctuations. However, this discrepancy can also be removed by more conventional methods, which do not require large fluctuations at last scattering. It may perhaps be removed by a more careful choice of the amplitudes and profiles of density and velocity in the proto-voids, but meanwhile one of us (K.B.) put forward another solution of the problem (see below).

   These investigations also demonstrated that velocity perturbations are distinctly more efficient in generating structures than are density perturbations. This calls for a revision of the long-standing structure formation paradigm that relies on the belief that density perturbations alone are responsible for the origin of structures. Matter at the edge of today's typical void should occupy a region of angular diameter 0.1° on the CMB sky, which is comparable to the current best resolution achieved in observations. Thus, for voids we are on the verge of being able to test our models against observations.

3. **A more precise description of the formation of voids.** In order to solve the problem of insufficiently low densities within voids obtained in the approach discussed above, we used the Lemaître model describing a mixture of inhomogeneous dust and inhomogeneous radiation. Since the Einstein equations with no further assumptions are indeterminate in this case (the number of unknown functions is larger by one than the number of equations), something had to be assumed about the radiation distribution. Lacking any data, we took the simplest assumption – that the comoving spatial extent of the perturbations of radiation density does not change with time (however the amplitude of radiation density does change with time in the same way as in the Friedmann models). Proceeding from this, we were able to reproduce the current density distribution in voids with arbitrary precision. The assumption about the profile of radiation density will most probably have to be modified in the future, but our result shows that this is a workable approach to the problem of void formation. This investigation highlights the fact that the decoupling of matter and radiation is not a single instant, as is usually assumed for simplicity, but a process extended in time.
4. **The evolution of double structures (cluster–void pairs).** We showed that the evolution of pairs where a void sits on the edge of a cluster, or vice versa, can be described using the quasi-spherical Szekeres (1975) model. Actually, the Szekeres model can be used to describe multiple structures as well, but because of its mathematical limitations the parameters of the additional objects are no longer free; in particular the peripheral structures come out too large. This model helps us to understand some actually observed facts, such as a faster evolution of large voids at larger distances from condensations compared to voids in the proximity of condensations. An additional bonus of this investigation is a physical interpretation of the arbitrary functions in the Szekeres model (they define the direction of the dipole component of the density distribution).
5. **Interpretation of the type Ia supernova dimming as an effect of inhomogeneities.** We showed that, in principle, the observed excessive dimming of type Ia supernovae with distance can be accounted for, on the basis of proper general relativity, without introducing 'dark energy', if one permits inhomogeneity in the matter distribution. This description is more natural than the standard assumption, which postulates a new and completely unknown form of energy. This might also be a way of dealing with the 'coincidence problem' attached to the time when the cosmological constant becomes

CAMBRIDGE

Cambridge University Press
978-0-521-76914-3 - Structures in the Universe by Exact Methods: Formation, Evolution, Interactions
Krzysztof Bolejko, Andrzej Krasinski, Charles Hellaby and Marie-Noelle Celerier
Excerpt
More information

dominant over the matter density. Even if the cosmological constant appears as a very natural geometrical tool in Einstein's equations, this implies that we live in a preferred epoch, when the densities of ordinary matter and 'dark energy' are comparable. (The cosmological term becomes negligible at very high mass densities and extremely dominant at very low mass densities.)

We described how this problem can be dealt with in the more general setting, then we gave some examples of exact inhomogeneous models with no cosmological constant which can be found in the recent literature, and which fit cosmological observations such as supernova observations, baryon acoustic oscillations and $H(z)$ measurements, and in addition recover the position of the CMB power spectrum peaks. It was thus proved that evolving inhomogeneities can mimic, partly or totally, the effects of the 'dark energy' component of the Concordance model.

Note, however, that the 'dark energy' problem can be approached in a still different way. The cosmological models of the Robertson–Walker class are supposed to apply at large scales, to an already-averaged geometry and matter distribution in the Universe. So far, there exists no universally accepted definition of averaging that would be both exact and covariant. However, there are several definitions of approximate or non-covariant averaging procedures (see, for example, the last chapter in Krasiński, 1997, and more recent references in Célérier, 2007b) which agree in one point: averaging the nonlinear Einstein equations produces an additional term that mimics negative pressure (i.e. repulsion) in the large-scale energy–momentum tensor. We did not discuss this approach here because of the lack of a general consensus concerning the right method, but this line of research also seems more physical than postulating a new form of energy.

6. **Solving the 'horizon problem' without the use of inflationary models.** We showed that this supposed problem can be solved using the L–T model with appropriately chosen arbitrary functions. This is preferable to the use of inflation, which makes assumptions about physical conditions in the Universe at such early times that any kind of direct verification is currently impossible. Moreover, the solution we proposed solves the horizon problem whatever the location of the observer in spacetime while inflation solves it only temporarily. Alternatives to inflation deserve at least to be considered, but, unfortunately, they tend to be suppressed in the noise of lobbying that has surrounded the inflationary paradigm from the very beginning.
7. **Influence of inhomogeneous structures in the path of a light ray on the observed temperature distribution of the CMB radiation.** We first examined whether (part of) the dipole moment of the CMB might be of a cosmological nature, by putting the observer off the centre of a particular class of L–T models. This result must be taken with caution, since the moments beyond the quadrupole were not calculated here. However, it was

the first time, to our knowledge, that the geodesic equations for non-radial photons and formulae for the dipole and the quadrupole were established for L–T models. We also gave an account of a study similar to ours, which was completed later on and in which the octopole was also calculated. For further investigation we used the quasi-spherical Szekeres model to describe localised condensations or voids, matched into the homogeneous Friedmann background. The light ray under investigation proceeded from a source that emitted it during the decoupling epoch, through several inhomogeneous structures, to an observer registering it at the present epoch. It turned out that the temperature anisotropies caused by the structures are smaller than those generated by the Sachs–Wolfe (1967) effect, unless the observer is situated within one of those structures.

The investigations reported here are preliminary, and most of them will have to be refined or revised in the future, when more precise data become available, and methods of self-consistent interpretation of observational data against inhomogeneous models are developed. The point we wish to make at present is that this branch of cosmology is already advanced to some degree and its results and potential contributions deserve to be more widely recognised.

PART I

Theoretical background

2

Exact solutions of Einstein's equations that are used in cosmology

2.1 The Lemaître–Tolman model

2.1.1 Basic properties

The Lemaître (1933)–Tolman (1934) (L–T) model is a spherically symmetric nonstatic solution of the Einstein equations with a dust source, i.e. the matter tensor is $T^{\alpha\beta} = \rho u^\alpha u^\beta$. The coordinates are assumed to be comoving, so that the 4-velocity of matter is $u^\alpha = \delta^\alpha_t$. See Krasiński (1997) and Plebański and Krasiński (2006) for an extensive list of properties and other work on this model. Its metric is (in units in which $c = 1$ and in the synchronous time gauge):

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{{R_{,r}}^2}{1+2E(r)}\mathrm{d}r^2 - R^2(t,r)(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2), \tag{2.1}$$

where $E(r)$ is an arbitrary function, $R_{,r} = \partial R/\partial r$, and $R(t,r)$ obeys

$${R_{,t}}^2 = 2E + \frac{2M}{R} + \frac{\Lambda}{3}R^2, \tag{2.2}$$

where $R_{,t} = \partial R/\partial t$ and Λ is the cosmological constant. Equation (2.2) is a first integral of the Einstein equations, and $M = M(r)$ is another arbitrary function of integration. The mass density in energy units is:

$$\kappa\rho = \frac{2M_{,r}}{R^2 R_{,r}}, \qquad \text{where } \kappa = \frac{8\pi G}{c^4}. \tag{2.3}$$

The metric of the space $t =$ const would be flat if $E(r)$ were set to zero, so E determines the curvature of space at each r value. Comparison of (2.2) with the Newtonian energy equation for a spherically symmetric dust distribution indicates that $M(r)$ is the gravitational mass contained within the comoving spherical shell at any given r, while $E(r)$ is the energy per unit mass of the particles in that shell.

Equation (2.2) can be solved by simple integration:

$$\int_0^R \frac{\mathrm{d}\widetilde{R}}{\sqrt{2E + 2M/\widetilde{R} + \frac{1}{3}\Lambda\widetilde{R}^2}} = t - t_B(r), \tag{2.4}$$

where t_B appears as an integration function (the bang time), and is an arbitrary function of r. This means that the Big Bang is not simultaneous as in the Friedmann models, but occurs at different times at different distances from the origin.