1

Dark energy, gravitation and the Copernican principle

JEAN-PHILIPPE UZAN

1.1 Cosmological models and their hypotheses

1.1.1 Introduction

The progress of physical cosmology during the past ten years has led to a "standard" cosmological model in agreement with all available data. Its parameters are measured with increasing precision but it requires the introduction of a dark sector, including both dark matter and dark energy, attracting the attention of both observers and theoreticians.

Among all the observational conclusions, the existence of a recent acceleration phase of the cosmic expansion has become more and more robust. The quest for the understanding of its physical origin is however just starting (Peebles and Ratra, 2003; Peter and Uzan, 2005; Copeland *et al.*, 2006; Uzan, 2007). Models and speculations are flourishing and we may wonder to what extent the observations of our local universe may reveal the physical nature of the dark energy. In particular, there exist limitations to this quest intrinsic to cosmology, related to the fact that most observations are located on our past light-cone (Ellis, 1975), and to finite volume effects (Bernardeau and Uzan, 2004) that can make many physically acceptable possibilities indistinguishable in practice.

This text discusses the relations between the cosmic acceleration and the theory of gravitation and more generally with the hypotheses underlying the construction of our cosmological model, such as the validity of general relativity on astrophysical scales and the Copernican principle. We hope to illustrate that cosmological data now have the potential to test these hypotheses, which go beyond the measurements of the parameters.

Dark Energy: Observational and Theoretical Approaches, ed. Pilar Ruiz-Lapuente. Published by Cambridge University Press. © Cambridge University Press 2010.

1.1.2 Cosmology, physics and astronomy

Cosmology sits at the crossroads between theoretical physics and astronomy.

Theoretical physics, based on physical laws, tries to describe the fundamental components of nature and their interactions. These laws can be probed locally by experiments. These laws need to be extrapolated to construct cosmological models. Hence, any new idea or discovery concerning these laws can naturally call for an extension of our cosmological model (e.g. introducing massive neutrinos in cosmology is now mandatory).

Astronomy confronts us with phenomena that we have to understand and explain consistently. This often requires the introduction of hypotheses beyond those of the physical theories (Section 1.1.3) in order to “save the phenomena” (Duhem, 1908), as is actually the case with the dark sector of our cosmological model. Needless to say, even if a cosmological model is in agreement with all observations, whatever their accuracy, it does not prove that it is the “correct” model of the universe, in the sense that it is the correct cosmological extrapolation and solution of the local physical laws.

Dark energy confronts us with a compatibility problem since, in order to “save the phenomena” of the observations, we have to include new ingredients (cosmological constant, matter fields or interactions) beyond those of our established physical theories. However, the required value for the simplest dark energy model, i.e. the cosmological constant, is more than 60 orders of magnitude smaller than what is expected from theoretical grounds (Section 1.1.6). This tension between what is required by astronomy and what is expected from physics reminds us of the twenty-centuries long debate between Aristotelians and Ptolemaeans (Duhem, 1913), that was resolved not only by the Copernican model but more importantly by a better understanding of the physics, since Newton’s gravity was compatible only with one of these three models that, at the time, could not be distinguished observationally.

1.1.3 Hypotheses of our cosmological model

The construction of any cosmological model relies on four main hypotheses:

(H1) a theory of gravity,
(H2) a description of the matter contained in the universe and its non-gravitational interactions,
(H3) symmetry hypotheses, and
(H4) a hypothesis on the global structure, i.e. the topology, of the universe.

These hypotheses are not on the same footing, since H1 and H2 refer to the physical theories. These two hypotheses are, however, not sufficient to solve the field equations and we must make an assumption on the symmetries (H3) of the solutions describing our universe on large scales, while H4 is an assumption on some global properties of these cosmological solutions, with the same local geometry.

Our reference cosmological model is the ΛCDM model. It assumes that gravity is described by general relativity (H1), that the universe contains the fields of the standard model of particle physics plus some dark matter and a cosmological constant, the last two having no physical explanation at the moment. Note that in the cosmological context this involves an extra assumption, since what will be required by the Einstein equations is the effective stress–energy tensor averaged on large scales. It thus implicitly refers to a, usually not explicit, averaging procedure (Ellis and Buchert, 2005). It also deeply involves the Copernican principle as a symmetry hypothesis (H3), without which the Einstein equations usually cannot be solved, and usually assumes that the spatial sections are simply connected (H4). H2 and H3 imply that the description of standard matter reduces to a mixture of a pressureless fluid and a radiation perfect fluid.

1.1.4 Copernican principle

The *cosmological principle* supposes that the universe is spatially isotropic and homogeneous. In particular, this implies that there exists a privileged class of observers, called fundamental observers, who all see an isotropic universe around them. It implies the existence of a cosmic time and states that all the properties of the universe are the same everywhere at the same cosmic time. It is supposed to hold for the smoothed-out structure of the universe on large scales. Indeed, this principle has to be applied in a statistical sense since there exist structures in the universe.

We can distinguish it from the *Copernican principle* which merely states that we do not live in a special place (the center) in the universe. As long as isotropy around the observer holds, this principle actually leads to the same conclusion as the cosmological principle.

The cosmological principle makes definite predictions about all unobservable regions beyond the observable universe. It completely determines the entire structure of the universe, even for regions that cannot be observed. From this point of view, this hypothesis, which cannot be tested, is very strong. On the other hand, it leads to a complete model of the universe. The Copernican principle has more modest consequences and leads to the same conclusions but only for the observable universe where isotropy has been verified. It does not make any prediction on the structure of the universe for unobserved regions (in particular, space could be homogeneous and non-isotropic on scales larger than the observable universe). We refer to Bondi (1960), North (1965) and Ellis (1975) for further discussions on the definition of these two principles.

We emphasize that, as will be discussed in the next section, our reference cosmological model includes a primordial phase of inflation in order to explain the origin of the large-scale structures of the universe. Inflation gives a theoretical prejudice in favor of the Copernican principle since it predicts that all classical (i.e. non-quantum) inhomogeneities (curvature, shear, etc.) have been washed-out during this phase. If it is sufficiently long, we expect the principle to hold on scales much larger than those of the observable universe, hence backing-up the cosmological principle, since unobservable regions today arise from the same causal process that affected the conditions in our local universe. While the standard predictions of inflation are in agreement with all astronomical data, we should not forget that it is only a theoretical argument to which we shall come back if we find observable evidence against isotropy (Pereira *et al.*, 2007; Pitrou *et al.*, 2008), curvature (Uzan *et al.*, 2003), and homogeneity (e.g. a spatial topology of the universe).

These principles lead to a Robertson–Walker (RW) geometry with metric

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\gamma_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j, \tag{1.1}$$

where t is the cosmic time and γ_{ij} is the spatial metric on the constant time hypersurfaces, which are homogeneous and isotropic, and thus of constant curvature. It follows that the metric is reduced to a single function of time, the scale factor, a. This implies that there is a one-to-one mapping between the cosmic time and the redshift z:

$$1 + z = \frac{a_0}{a(t)}, \tag{1.2}$$

if the expansion is monotonous.

1.1.5 ΛCDM reference model

The dynamics of the scale factor can be determined from the Einstein equations which reduce for the metric (1.1) to the Friedmann equations:

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} + \frac{\Lambda}{3}, \tag{1.3}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3}. \tag{1.4}$$

$H \equiv \dot{a}/a$ is the Hubble function and $K = 0, \pm 1$ is the curvature of the spatial sections. G and Λ are the Newton and cosmological constants. ρ and P are respectively the energy density and pressure of the cosmic fluids and are related by

$$\dot{\rho} + 3H(\rho + P) = 0.$$

Defining the dimensionless density parameters as

$$\Omega = \frac{8\pi G\rho}{3H^2}, \quad \Omega_\Lambda = \frac{\Lambda}{3H^2}, \quad \Omega_K = -\frac{K}{H^2 a^2}, \tag{1.5}$$

respectively for the matter, the cosmological constant and the curvature, the first Friedmann equation can be rewritten as

$$\begin{aligned} E^2(z) &\equiv \left(\frac{H}{H_0}\right)^2 \\ &= \Omega_{\mathrm{rad}0}(1+z)^4 + \Omega_{\mathrm{mat}0}(1+z)^3 + \Omega_{K0}(1+z)^2 + \Omega_{\Lambda 0}, \end{aligned} \tag{1.6}$$

with $\Omega_{K0} = 1 - \Omega_{\mathrm{rad}0} - \Omega_{\mathrm{mat}0} - \Omega_{\Lambda 0}$. All background observables, such as the luminosity distance, the angular distance, etc., are functions of $E(z)$ and are thus not independent.

Besides this background description, the ΛCDM also accounts for an understanding of the large-scale structure of our universe (galaxy distribution, cosmic microwave background anisotropy) by using the theory of cosmological perturbations at linear order. In particular, in the sub-Hubble regime, the growth rate of the density perturbations is also a function of $E(z)$.

One must, however, extend this minimal description by a primordial phase in order to solve the standard cosmological problems (flatness, horizon, etc.). In our reference model, we assume that this phase is described by an inflationary period during which the expansion of the universe is almost exponentially accelerated. In such a case, the initial conditions for the gravitational dynamics that will lead to the large-scale structure are also determined so that our model is completely predictive. We refer to Chapter 8 of Peter and Uzan (2005) for a detailed description of these issues that are part of our cosmological model but not directly related to our actual discussion.

In this framework, the dark energy is well defined and reduces to a single number equivalent to a fluid with equation of state $w = P/\rho = -1$. This model is compatible with all astronomical data, which roughly indicates that

$$\Omega_{\Lambda 0} \simeq 0.73, \quad \Omega_{\mathrm{mat}0} \simeq 0.27, \quad \Omega_{K0} \simeq 0.$$

1.1.6 The cosmological constant problem

This model is theoretically well-defined, observationally acceptable, phenomenologically simple and economical. From the perspective of general relativity the value of Λ is completely free and there is no argument allowing us to fix it, or equivalently, the length scale $\ell_\Lambda = |\Lambda_0|^{-1/2}$, where Λ_0 is the astronomically deduced value of the cosmological constant. Cosmology roughly imposes that

$$|\Lambda_0| \le H_0^2 \iff \ell_\Lambda \le H_0^{-1} \sim 10^{26}\,\mathrm{m} \sim 10^{41}\,\mathrm{GeV}^{-1}.$$

In itself this value is no problem, as long as we only consider classical physics. Notice, however, that it is disproportionately large compared to the natural scale fixed by the Planck length:

$$\ell_\Lambda > 10^{60}\ell_\mathrm{P} \iff \frac{\Lambda_0}{M_\mathrm{Pl}^2} < 10^{-120} \iff \rho_{\Lambda_0} < 10^{-120} M_\mathrm{Pl}^4 \sim 10^{-47}\,\mathrm{GeV}^4, \tag{1.7}$$

when expressed in terms of energy density.

The main problem arises from the interpretation of the cosmological constant. The local Lorentz invariance of the vacuum implies that its energy–momentum tensor must take the form (Zel'dovich, 1988) $\langle T_{\mu\nu}^{\mathrm{vac}}\rangle = -\langle\rho\rangle g_{\mu\nu}$, which is equivalent to that of a cosmological constant. From the quantum point of view, the vacuum energy receives a contribution of the order of

$$\langle\rho\rangle_{\mathrm{vac}}^{\mathrm{EW}} \sim (200\,\mathrm{GeV})^4, \qquad \langle\rho\rangle_{\mathrm{vac}}^{\mathrm{Pl}} \sim (10^{18}\,\mathrm{GeV})^4, \tag{1.8}$$

arising from the zero point energy, respectively fixing the cutoff frequency of the theory to the electroweak scale or to the Planck scale. This contribution implies a disagreement of respectively 60 to 120 orders of magnitude with astronomical observations!

This is the cosmological constant problem (Weinberg, 1989). It amounts to understanding why

$$|\rho_{\Lambda_0}| = |\rho_\Lambda + \langle\rho\rangle_{\mathrm{vac}}| < 10^{-47}\,\mathrm{GeV}^4 \tag{1.9}$$

or equivalently,

$$|\Lambda_0| = |\Lambda + 8\pi G\langle\rho\rangle_{\mathrm{vac}}| < 10^{-120} M_{\mathrm{Pl}}^2, \tag{1.10}$$

i.e. why ρ_{Λ_0} is so small today, but non-zero.

Today, there is no known solution to this problem and two approaches have been designed. One the one hand, one sticks to this model and extends the cosmological model in order to explain why we observe such a small value of the cosmological constant (Garriga and Vilenkin, 2004; Carr and Ellis, 2008). We shall come back to this approach later. On the other hand, one hopes that there should exist a physical mechanism to exactly cancel the cosmological constant and looks for another mechanism to explain the observed acceleration of the universe.

1.1.7 The equation of state of dark energy

The equation of state of the dark energy is obtained from the expansion history, assuming the standard Friedmann equation. It is thus given by the general expression (Martin *et al.*, 2006)

$$3\Omega_{\mathrm{de}} w_{\mathrm{de}} = -1 + \Omega_K + 2q, \tag{1.11}$$

q being the deceleration parameter,

$$q \equiv -\frac{a\ddot{a}}{\dot{a}^2} = -1 + \frac{1}{2}(1+z)\frac{\mathrm{d}\ln H^2}{\mathrm{d}z}. \tag{1.12}$$

This expression (1.11) does not assume the validity of general relativity or any theory of gravity, but gives the relation between the dynamics of the expansion history and the property of the matter that would lead to this acceleration if general relativity described gravity. Thus, the equation of state, as defined in Eq. (1.11), reduces to the ratio of the pressure, P_{de}, to the energy density, ρ_{de}, of an effective dark energy fluid under this assumption only, that is if

$$H^2 = \frac{8\pi G}{3}(\rho + \rho_{\mathrm{de}}) - \frac{K}{a^2}, \tag{1.13}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + \rho_{\mathrm{de}} + 3P + 3P_{\mathrm{de}}). \tag{1.14}$$

All the background information about dark energy is thus encapsulated in the single function $w_{\mathrm{de}}(z)$. Most observational constraints on the dark energy equation of state refer to this definition.

1.2 Modifying the minimal ΛCDM

The *Copernican principle* implies that the spacetime metric reduces to a single function, the scale factor $a(t)$, which can be Taylor expanded as $a(t) = a_0 + H_0 (t - t_0) - \frac{1}{2}q_0 H_0^2 (t - t_0)^2 + \cdots$. It follows that the conclusion that the cosmic expansion is accelerating ($q_0 < 0$) does not involve any hypothesis about the theory of gravity (other than that the spacetime geometry can be described by a metric) or the matter content, as long as this principle holds.

The assumption that the Copernican principle holds, and the fact that it is so central in drawing our conclusion on the acceleration of the expansion, splits our investigation into two avenues. Either we assume that the Copernican principle holds and we have to modify the laws of fundamental physics or we abandon the

Copernican principle, hoping to explain dark energy without any new physics but at the expense of living in a particular place in the universe. While the first solution is more orthodox from a cosmological point of view, the second is indeed more conservative from a physical point of view. It will be addressed in Section 1.2.4. We are thus faced with a choice between "simple" cosmological solutions with new physics and more involved cosmological solutions of standard physics.

This section focuses on the first approach. If general relativity holds then Eq. (1.4) tells us that the dynamics has to be dominated by a dark energy fluid with $w_{\mathrm{de}} < -\frac{1}{3}$ for the expansion to be accelerated. The simplest solution is indeed the cosmological constant Λ for which $w_{\mathrm{de}} = -1$ and which is the only model not introducing new degrees of freedom.

1.2.1 General classification of physical models

1.2.1.1 General relativity

Einstein's theory of gravity relies on two independent hypotheses.

First, the theory rests on the Einstein equivalence principle, which includes the universality of free-fall, the local position and local Lorentz invariances in its weak form (as other metric theories) and is conjectured to satisfy it in its strong form. We refer to Will (1981) for a detailed explanation of these principles and their implications. The weak equivalence principle can be mathematically implemented by assuming that all matter fields are minimally coupled to a single metric tensor $g_{\mu\nu}$. This metric defines the length and times measured by laboratory clocks and rods so that it can be called the *physical metric*. This implies that the action for any matter field, ψ say, can be written as $S_{\mathrm{matter}}[\psi, g_{\mu\nu}]$. This so-called *metric coupling* ensures in particular the validity of the universality of free-fall.

The action for the gravitational sector is given by the Einstein–Hilbert action

$$S_{\mathrm{gravity}} = \frac{c^3}{16\pi G}\int \mathrm{d}^4x\sqrt{-g_*}R_*, \tag{1.15}$$

where $g^*_{\mu\nu}$ is a massless spin-2 field called the *Einstein metric*. The second hypothesis states that both metrics coincide

$$g_{\mu\nu} = g^*_{\mu\nu}.$$

The underlying physics of our reference cosmological model (i.e. hypotheses H1 and H2) is thus described by the action

$$S_{\mathrm{gravity}} = \frac{c^3}{16\pi G}\int \mathrm{d}^4x\sqrt{-g}(R - 2\boldsymbol{\Lambda}) + \sum_{\text{standard model}+\mathbf{CDM}} S_{\mathrm{matter}}[\psi_i, g_{\mu\nu}], \tag{1.16}$$

which includes all known matter fields plus two unknown components (in bold face).

1.2.1.2 Local experimental constraints

The assumption of a metric coupling is well tested in the solar system. First, it implies that all non-gravitational constants are spacetime independent, and have been tested to a very high accuracy in many physical systems and for various fundamental constants (Uzan, 2003, 2004; Uzan and Leclercq, 2008), e.g. at the 10^{-7} level for the fine structure constant on time scales ranging to 2–4 Gyrs. Second, the isotropy has been tested from the constraint on the possible quadrupolar shift of nuclear energy levels (Prestage *et al.*, 1985; Chupp *et al.*, 1989; Lamoreaux *et al.*, 1986) proving that matter couples to a unique metric tensor at the 10^{-27} level. Third, the universality of free-fall of test bodies in an external gravitational field at the 10^{-13} level has been tested in the laboratory (Baessler *et al.*, 1999; Adelberger *et al.*, 2001). The Lunar Laser ranging experiment (Williams *et al.*, 2004), which compares the relative acceleration of the Earth and Moon in the gravitational field of the Sun, also probes the strong equivalence principle at the 10^{-4} level. Fourth, the Einstein effect (or gravitational redshift) states that two identical clocks located at two different positions in a static Newton potential U and compared by means of electromagnetic signals will exhibit a difference in clock rates of $1 + [U_1 - U_2]/c^2$, where U is the gravitational potential. This effect has been measured at the 2×10^{-4} level (Vessot and Levine, 1978).

The parameterized post-Newtonian formalism (PPN) is a general formalism that introduces 10 phenomenological parameters to describe any possible deviation from general relativity at the first post-Newtonian order (Will, 1981). The formalism assumes that gravity is described by a metric and that it does not involve any characteristic scale. In its simplest form, it reduces to the two Eddington parameters entering the metric of the Schwartzschild metric in isotropic coordinates:

$$g_{00} = -1 + \frac{2Gm}{rc^2} - 2\beta^{\mathrm{PPN}}\left(\frac{2Gm}{rc^2}\right)^2, \qquad g_{ij} = \left(1 + 2\gamma^{\mathrm{PPN}}\frac{2Gm}{rc^2}\right)\delta_{ij}.$$

Indeed, general relativity predicts $\beta^{\mathrm{PPN}} = \gamma^{\mathrm{PPN}} = 1$. These two phenomenological parameters are constrained by: (1) the shift of the Mercury perihelion (Shapiro *et al.*, 1990), which implies that $|2\gamma^{\mathrm{PPN}} - \beta^{\mathrm{PPN}} - 1| < 3 \times 10^{-3}$; (2) the Lunar Laser ranging experiment (Williams *et al.*, 2004) which implies $|4\beta^{\mathrm{PPN}} - \gamma^{\mathrm{PPN}} - 3| = (4.4 \pm 4.5) \times 10^{-4}$ and (3) the deflection of electromagnetic signals, which are all controlled by γ^{PPN}. For instance, very long baseline interferometry (Shapiro *et al.*, 2004) implies that $|\gamma^{\mathrm{PPN}} - 1| = 4 \times 10^{-4}$, while measurement of the time delay variation to the Cassini spacecraft (Bertotti *et al.*, 2003) sets $\gamma^{\text{PPN}} - 1 = (2.1 \pm 2.3) \times 10^{-5}$.

The PPN formalism does not allow us to test finite range effects that could be caused e.g. by a massive degree of freedom. In that case one expects a Yukawa-type deviation from the Newton potential,

$$V = \frac{Gm}{r}\left(1 + \alpha e^{-r/\lambda}\right),$$

that can be probed by "fifth force" experimental searches. λ characterizes the range of the Yukawa deviation while its strength α may also include a composition dependence (Uzan, 2003). The constraints on (λ, α) are summarized in Hoyle *et al.* (2004), which typically shows that $\alpha < 10^{-2}$ on scales ranging from the millimeter to the solar system size.

In general relativity, the graviton is massless. One can, however, give it a mass, but this is very constrained. In particular, around a Minkowski background, the mass term must have the very specific form of the Pauli–Fierz type in order to avoid ghosts (see below for a more precise definition) being excited. This mass term is, however, inconsistent with solar system constraints because there exists a discontinuity (van Dam and Veltman, 1970; Zakharov, 1970) between the case of a strictly massless graviton and a very light one. In particular, such a term can be ruled out from the Mercury perihelion shift.

General relativity is also tested with pulsars (Damour and Esposito-Farèse, 1998; Esposito-Farèse, 2005) and in the strong field regime (Psaltis, 2008). For more details we refer to Will (1981), Damour and Lilley (2008) and Turyshev (2008). Needless to say, any extension of general relativity has to pass these constraints. However, deviations from general relativity can be larger in the past, as we shall see, which makes cosmology an interesting physical system to extend these constraints.

1.2.1.3 Universality classes

There are many possibilities to extend this minimal physical framework. Let us start by defining universality classes (Uzan, 2007) by restricting our discussion to field theories. This helps in identifying the new degrees of freedom and their couplings.

The first two classes assume that gravitation is well described by general relativity and introduce new degrees of freedom beyond those of the standard model of particle physics. This means that one adds a new term $S_{\text{de}}[\psi; g_{\mu\nu}]$ in the action (1.16) while keeping the Einstein–Hilbert action and the coupling of all the fields (standard matter and dark matter) unchanged. They are:

1. *Class A* consists of models in which the acceleration is driven by the gravitational effect of the new fields. They thus must have an equation of state smaller than $-\frac{1}{3}$. They are