Modified gravity and the stability of de Sitter space

Valerio Faraoni

Physics Department, Bishop’s University
Lennoxville, Québec, Canada J1M 1Z7

(Dated: September 2, 2005)

Within the context of modified gravity and dark energy scenarios of the accelerated universe, we study the stability of de Sitter space with respect to inhomogeneous perturbations using a gauge-independent formalism. In modified gravity the stability condition is exactly the same that one obtains from a homogeneous perturbation analysis, while the stability condition in scalar–tensor gravity is more restrictive.

The recent discovery that the expansion of the universe is accelerated, obtained by studying type Ia supernovae [1], and the cosmic microwave background experiments showing that the universe has nearly critical density [2], call for a theoretical explanation. Two classes of models are predominant in the literature: the first class assumes that there is a form of dark energy (or quintessence) unclustered at all scales, that accounts for 70% of the energy density \( \rho \) of the universe. This dark energy, of unknown nature, is necessarily exotic: to generate acceleration in Einstein gravity it must have negative pressure \( P_{DE} < -\rho_{DE}/3 \). The best fit to the observational data favours an even more exotic dark energy (phantom energy or superquintessence) with equation of state parameter \( w \equiv P_{DE}/\rho_{DE} < -1 \), which is evolving in time [3]. Were a value \( w < -1 \) to be confirmed by observations, it can not be explained by general relativity with a canonical scalar field \( \phi \), which is the most common model of dark energy, because of the Einstein–Friedmann equation \( \ddot{H} = -\kappa (\rho + \rho_\phi)/6 = -\kappa \dot{\phi}^2/2 \leq 0 \), which is incompatible with \( P_\phi < -\rho_\phi \) (equivalent to superacceleration \( \ddot{H} > 0 \)). To model an equation of state parameter \( w < -1 \), a phantom field [4] or a field coupled nonminimally to the Ricci curvature \( R \) (see Refs. [5] [6] for reviews) have been considered. These theories can be seen as special cases of scalar–tensor gravity, described by the action

\[
S = \int d^4x \sqrt{-g} \left[ \psi(\phi)^R - \frac{1}{2} \nabla c \phi \nabla c \phi - V(\phi) \right].
\]

The second class of models does not require the presence of exotic dark energy but modifies gravity at large scales by introducing non–linear (in \( R \)) corrections to the Einstein–Hilbert Lagrangian which become dominant only at late times (low curvatures) [8]–[10]. These theories often suffer from problems with the post–Newtonian limit [11] [12] [13] or from instabilities [14], and are not yet accepted as completely viable theories, but they are nevertheless interesting as the cosmic acceleration that we are observing may be the first sign of a departure from Einstein’s gravity. Furthermore, these models are motivated by certain compactifications of M–theory [15].

In both dark energy models and modified gravity, depending on the model adopted, the universe may accelerate forever or end its existence at a finite time in the future in a Big Rip or sudden future singularity [17] [18] [19]. Such singularities have been classified in Ref. [16]; according to this classification, it is known that singularities of type I can occur in these models [18], but singularities of other types are not excluded.

The fate of the universe depends on the presence and size of the attraction basins of attractor solutions in the phase space. In many models of both dark energy and modified gravity, a de Sitter attractor solution is found [21]. In this paper we address the issue of the stability of de Sitter space in modified gravity and scalar–tensor theories. It is straightforward to assess stability with respect to homogeneous (time–dependent only) perturbations: however, it is more significant to establish whether de Sitter space is also stable with respect to more general inhomogeneous (space– and time–dependent) perturbations. This is a much more difficult task because of the gauge–dependence problems associated with this kind of perturbations [22]. In the following we show that for modified gravity the stability condition obtained with a gauge–independent inhomogeneous perturbation analysis reduces to that obtained in the much simpler homogeneous perturbation analysis, whereas this is not the case for scalar–tensor gravity.

We begin from the generalized gravity action

\[
S = \int d^4x \sqrt{-g} \left[ \frac{f(\phi, R)}{2} - \frac{\omega(\phi)}{2} \nabla c \phi \nabla c \phi - V(\phi) \right],
\]

which contains scalar–tensor gravity as the case \( f(\phi, R) = \psi(\phi) R \), and modified gravity \( f(R) \) when the scalar field \( \phi \) is absent and \( f_{RR} \neq 0 \). In the spatially flat Friedmann–Lemaître–Robertson–Walker (FLRW) metric

\[
ds^2 = -dt^2 + a^2(t) (dx^2 + dy^2 + dz^2),
\]

the field equations are

\[
H^2 = \frac{1}{3F} \left( \frac{\omega}{2} \dot{\phi}^2 + \frac{RF}{2} - \frac{f}{2} + V - 3H \dot{F} \right),
\]

\[
\dot{H} = -\frac{1}{2F} \left( \omega \dot{\phi}^2 + \dot{F} - H \dot{F} \right),
\]

\[
\ddot{\phi} + 3H \dot{\phi} + \frac{1}{2\omega} \frac{d\omega}{d\phi} \dot{\phi}^2 - \frac{\partial f}{\partial \phi} + 2 \frac{dV}{d\phi} = 0,
\]
where \( F \equiv \partial f / \partial R \), \( H \equiv a / \dot{a} \), and an overdot denotes differentiation with respect to the comoving time \( t \). By choosing \((H, \phi)\) as dynamical variables, the equilibrium points of the dynamical system (4)–(6) are de Sitter spaces with constant scalar field \((H_0, \phi_0)\). These solutions exist subject to the conditions

\[
6H_0^2 F_0 - f_0 + 2V_0 = 0 ,
\]

\[
f_0' - 2V_0' = 0 ,
\]

where \( F_0 \equiv F(\phi_0, R_0) \), \( f_0 \equiv f(\phi_0, R_0) \), \( V_0 \equiv V(\phi_0) \), \( V_0' \equiv \frac{dV}{dR}|_{\phi_0} \), a prime denotes differentiation with respect to \( \phi \), and \( R_0 = 12H_0^2 \).

Inhomogeneous perturbations of de Sitter space are investigated by using the covariant and gauge–invariant formalism of Bardeen–Ellis–Bruni–Hwang [22] in the version studied by Hwang [23] for generalized gravity. The metric perturbations are defined by

\[
\delta g_{00} = -a^2 (1 + 2AY) ,
\]

\[
\delta g_{0i} = -a^2 B Y_i ,
\]

\[
\delta g_{ij} = a^2 [h_{ij} (1 + 2H_L) + 2H_T Y_{ij}] ,
\]

where the scalar harmonics \( Y \) are the eigenfunctions of the eigenvalue problem \( \nabla_i \nabla^i Y = -k^2 Y \). Here \( h_{ij} \) is the three-dimensional metric of the FLRW background and the operator \( \nabla_i \) is the covariant derivative associated with \( h_{ij} \), while \( k \) is an eigenvalue. The vector and tensor harmonics \( Y_i \) and \( Y_{ij} \) obey

\[
Y_i = -\frac{1}{k} \nabla_i Y , \quad Y_{ij} = \frac{1}{k^2} \nabla_i \nabla_j Y + \frac{1}{3} Y h_{ij} .
\]

Bardeen’s gauge–invariant potentials

\[
\Phi_H = H_L + \frac{H_T}{3} + \frac{\dot{a}}{k} \left( B - \frac{a}{k} \dot{H}_T \right) ,
\]

\[
\Phi_A = A + \frac{\dot{a}}{k} \left( B - \frac{a}{k} \dot{H}_T \right) + \frac{a}{k} \left[ \dot{B} - \frac{1}{k} \left( a \dot{H}_T \right) \right] ,
\]

and the Ellis–Bruni variable

\[
\Delta \phi = \delta \phi + \frac{a}{k} \dot{\phi} \left( B - \frac{a}{k} \dot{H}_T \right) ,
\]

are used, with equations similar to eq. (15) defining the gauge–invariant variables \( \Delta f, \Delta F, \) and \( \Delta R \). The first order equations obeyed by the gauge–invariant perturbations are given in Ref. [22] and they simplify considerably in the de Sitter background \((H_0, \phi_0)\). To first order, they are:

\[
\Delta \ddot{\phi} + 3H_0 \Delta \dot{\phi} + \left[ \frac{k^2}{a^2} - \frac{1}{2\omega_0} (f_0'' - 2V_0'') \right] \Delta \phi = \frac{f_{\phi R}}{2\omega_0} \Delta R ,
\]

\[
\Delta \ddot{F} + 3H_0 \Delta \dot{F} + \left( \frac{k^2}{a^2} - 4H_0^2 \right) \Delta F + \frac{F_0}{3} \Delta R = 0 ,
\]

\[
\dot{H}_T + 3H_0 \dot{H}_T + \frac{k^2}{a^2} H_T = 0 ,
\]

\[
- \dot{\Phi}_H + H_0 \Phi_A = \frac{1}{2} \left( \frac{\Delta \dot{F}}{F_0} - H_0 \frac{\Delta F}{F_0} \right) ,
\]

\[
\Phi_H = -\frac{1}{2} \frac{\Delta F}{F_0} ,
\]

\[
\Phi_A + \Phi_H = -\frac{\Delta F}{F_0} ,
\]

\[
\dot{\Phi}_H + 3H_0 \Phi_H - H_0 \Phi_A - 3H_0^2 \Phi_A = -\frac{1}{2} \frac{\Delta \dot{F}}{F_0} - H_0 \frac{\Delta \dot{F}}{F_0} + \frac{3H_0^2}{2} \frac{\Delta F}{F_0} ,
\]
with

$$\Delta R = 6 \left[ \Phi_H + 4H_0\dot{\Phi}_H + \frac{2k^2}{3a^2} \dot{\Phi}_H - H_0\ddot{\Phi}_A + \left( \frac{k^2}{3a^2} - 4H_0^2 \right) \Phi_A \right], \quad (23)$$

Furthermore, vector perturbations do not have any effect to first order in the absence of ordinary matter and de Sitter space is always stable with respect to first order tensor perturbations, as can be seen from eq. 18, so we only need to worry about scalar perturbations (see Ref. 24 for details). We first consider modified gravity theories obtained by setting \(\Phi = 1\) and \(f = f(R)\) with \(f_{RR} \neq 0\) in the action \(\mathcal{L}\). The gauge–invariant perturbations are related by

$$\Phi_H = \Phi_A = -\frac{\Delta F}{2F_0}, \quad (24)$$

$$\Delta R = 6 \left[ \Phi_H + 3H_0\Phi_H + \left( \frac{k^2}{a^2} - 4H_0^2 \right) \Phi_H \right], \quad (25)$$

where \(a = a_{eq}H_{eq}\). By using the fact that \(\Delta F = \frac{F_0}{2} \Delta R\), one obtains \(\Delta R = -\frac{2F_0}{f_{RR}} \Phi_H\). The perturbations \(\Phi_H\) and \(\Phi_A\) evolve according to

$$\ddot{\Phi}_H + 3H_0\ddot{\Phi}_H + \left( \frac{k^2}{a^2} - 4H_0^2 \right) \Phi_H = 0; \quad (26)$$

at late times the term \(k^2/a^2\) can be safely neglected and stability is achieved if the coefficient of \(\Phi_H\) in the last term of the left hand side of eq. 26 is non–negative, i.e. (using eq. 15), if

$$\frac{F_0^2 - 2fo f_{RR}}{F_0 f_{RR}} \geq 0. \quad (27)$$

The spatial dependence of the inhomogeneous perturbations is encoded in the eigenvector \(k\) of the spherical harmonics; the fact that the only term containing \(k\) (or the physical wave vector \(k_{phys} = k/a\)) in eq. 26 becomes negligible on a de Sitter background implies that the spatial dependence effectively disappears from the analysis. Eq. 26 coincides with the stability condition that can be obtained by a straightforward homogeneous perturbation analysis of eqs. 4 and 5. Hence, in the stability analysis of de Sitter space in modified gravity theories one can safely neglect inhomogeneous perturbations and limit oneself to the much more approachable homogeneous perturbations, thus bypassing the gauge–dependence problems. However, this conclusion could not be drawn \textit{a priori} but it necessarily relies on the inhomogeneous perturbation analysis presented. Further, this result has been shown to be true only for the stability of de Sitter space and not for different attractor solutions that may be present in the phase space.

Naively, the physical reason for this considerable formal simplification could be looked for in the fact that, during the quasi–exponential expansion of the universe, inhomogeneities (and anisotropies \[25\]) are redshifted away; this is not the whole story though, because the simplification found for modified gravity does not occur in scalar–tensor theories. In fact, in this case, eqs. 11, 24, and 25, together with

$$\frac{\Delta F}{F_0} = \frac{\delta \varphi}{\varphi_0} \Delta R, \quad (28)$$

yield

$$\Delta \varphi + 3H_0 \Delta \varphi + \left[ k^2 \frac{f''}{a^2} - V'' + \frac{6f_{RR}}{F_0^2} \frac{H_0^2}{\omega_0 c^2} \right] \frac{\Delta \varphi}{\varphi_0} = 0. \quad (29)$$

if \(1 + \frac{3f_{RR}^2}{(2\omega_0 f_0)} \neq 0\). The stability condition of de Sitter space in scalar–tensor gravity then becomes \[24\]

$$\frac{(f'' - V'')}{\omega_0 c^2} \geq 0. \quad (30)$$

For general scalar–tensor theories, this condition is more restrictive than the corresponding stability condition obtained from a straightforward homogeneous perturbation analysis of eqs. 4–6, which is

$$\frac{f''}{\omega_0} - \frac{V''}{\omega_0} \leq 0. \quad (31)$$

However, for scalar–tensor theories of the form

$$S = \int d^4x \sqrt{-g} \left[ \phi R - \frac{\omega(\phi)}{\phi} \nabla^\alpha \phi \nabla_\alpha \phi - V(\phi) \right] \quad (32)$$

with a single coupling function, eqs. 40 and 51 coincide \[26\]. The reason for the failure of these equations to coincide in the general case can be traced to the right hand side of eq. 10, in which perturbations \(\Delta R\) act as a source for the perturbation \(\Delta \varphi\) (roughly speaking, perturbations \(\delta H\) in the Hubble parameter, and their derivatives, source scalar field perturbations \(\delta \varphi\)); such a term is absent in the homogeneous perturbation analysis of the Klein–Gordon equation \[10\]. Analogously, such a term is absent in eq. 26 obeyed by the gauge–independent perturbations \(\Phi_H\) (or \(\Phi_A\)) in modified gravity. This shows that although modified gravity is mathematically equivalent to a scalar–tensor theory \[11, 28\], the corresponding physics is not completely equivalent.
As an example of application of the stability condition, let us consider the theory described by 

\[ f(R) = R - \frac{\mu^4}{R}. \]  

(33)

The de Sitter space must satisfy the condition \( R_0 = 12H_0^2 = \sqrt{3}\mu^2 \) and the stability condition \( 27 \) can never be satisfied: this de Sitter space is always unstable. This situation can be ameliorated by a quadratic correction – in the theory with

\[ f(R) = R - \frac{\mu^4}{R} + aR^2, \]  

(34)

the condition for the existence of de Sitter space is again \( R_0 = \sqrt{3}\mu^2 \); by applying the stability condition \( 27 \) one obtains that de Sitter space is stable if

\[ a > \frac{1}{3\sqrt{3}\mu^2} \]  

(35)

and unstable otherwise (in particular for negative \( a \)). This result agrees with those of Refs. \( 27 \), which follow from an independent analysis of the effective potential in the Einstein conformal frame version of the scalar–tensor theory equivalent to the action \( 2 \) with \( f(R) \) specified by eq. \( 34 \). Furthermore, the theory described by corrections in both \( 1/R \) and \( R^2 \) has much better chances of passing the Solar System tests than the theory based on simple \( 1/R \) corrections \( 27 \).

A more complete discussion of the relation between stability of modified gravity and of scalar–tensor theories, and the application of eqs. \( 27 \) and \( 30 \) to other specific dark energy and modified gravity scenarios will be presented elsewhere.

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by a grant from the Senate Research Committee of Bishop’s University.

\[ \ast \text{vfarano@cs-linux.ubishops.ca} \]


[19] In certain scenarios, matter dominates again at late times and the universe eventually enters a decelerated era; however, it is very difficult to cross the barrier $w = -1$ from the region $w < -1$ [20].


