PHYSICAL REVIEW D **96**, 044013 (2017)

Tensor $f(R)$ theory of gravity

Tomasz Stachowiak*

Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, 606-8501 Kyoto, Japan

(Received 28 February 2017; published 11 August 2017)

I propose an alternative $f(R)$ theory of gravity constructed by applying the function f directly to the Ricci tensor instead of the Ricci scalar. The main goal of this study is to derive the resulting modified Einstein equations for the metric case with the Levi-Cività connection, as well as for the general nonmetric connection with torsion. The modification is then applied to the Robertson-Walker metric so that the cosmological evolution corresponding to the standard model can be studied. An appealing feature is that even in the vacuum case, scenarios without initial singularity and exponential expansion can be recovered. Finally, formulas for possible observational tests are given.

DOI: 10.1103/PhysRevD.96.044013

I. INTRODUCTION

The foundation of the present work is to consider a modified Lagrangian (density), which depends functionally on the full Ricci tensor R_{ab}, not just on its trace $\mathcal{R}$ as is the case in the so-called $f(\mathcal{R})$ theories of gravity. The principles of relativity require that this modification be obtained covariantly, and not component-wise, so writing $f(R_{ab})$ could be misleading. Since f will be a tensor-valued function, for the sake of distinction from the usual $f(\mathcal{R})$ theory, the extension will be referred to as tensor $f(R)$.

The motivation in both cases is the same—the inclusion of higher-order-of-curvature effects which can classically be ignored, but which lead to important modifications in other regimes. Most notably, the Starobinsky inflation model [1] induced by quadratic terms is a particularly important result in this spirit. Although initially introduced on quantum gravity grounds, with corrections built from various contractions of the Ricci tensor, it is now often considered in the language of quadratic $f(\mathcal{R})$ theories [2].

Despite the initial similarity, the tensor $f(R)$ gravity presented here differs considerably from the usual one, and the goal of this article is to focus first on the development of this new theory, with a comparative study left for future work. Accordingly, the notation and mathematical setting will be given as well as the modified Einstein equations. Not to stop at the abstract level, I will also consider possible applications to cosmology, with a view to nonsingular evolution, and provide basic formulas to be used in observational cosmology.

Notable differences and similarities with the ordinary $f(\mathcal{R})$ theory will be pointed out throughout the derivations in Secs. II, IV, and V, but for a more complete, general review of the standard approach, the reader might want to consult Refs. [3,4], or [5] and references therein.

II. CONSTRUCTION OF THE MODIFIED ACTION

In the usual $f(\mathcal{R})$ theories, one postulates the Lagrangian

$$\mathcal{L}_0 = f(\mathrm{tr}[R]) = f(R_{ac}g^{ca}), \tag{1}$$

with the summation convention used, and the covariant metric tensor denoted by g_{ac}. On purely abstract grounds, the order in which f and trace appear is not fixed, so instead of the above I will consider the Lagrangian to be

$$\mathcal{L}_g = \mathrm{tr}[f(R)] = [f(R)]_{ac}g^{ca}, \tag{2}$$

where the square brackets are used to indicate elements of a matrix, and the bare symbol R has to refer to the tensor, not the scalar, as explained below.

A similar idea has been studied before by Borowiec [6,7], but it differed from the present work in two ways. First, it used a torsionless metric, and second, the Lagrangian depended on polynomial invariants of the Ricci tensor $\mathrm{tr}[R^k]$. Such scalars formed with powers k higher than the space-time dimension can be reduced to the lower ones by using the characteristic polynomial. However, this cannot, in general, be done explicitly for transcendental functions—i.e., when one needs to use an infinite series of powers of R. What is more, the coefficients of the characteristic polynomial themselves depend on the components of R, leading to an unwieldy expression of an original function of R in terms of a function of the invariants $\mathrm{tr}[R^k]$. The present work aims at overcoming this problem, and also at including connections with the most general torsion and nonmetricity.

To proceed with the general treatment, the first thing to settle is what tensors and operators to use, and in particular

*tomasz@amp.i.kyoto-u.ac.jp

how to interpret $f(R)$. Power series immediately come to mind, so what is needed is a representation of R such that it can be composed with itself by matrix multiplication consistent with relativistic index contraction. In other words, R should be an endomorphism, for instance on the tangent bundle over the space-time.

To treat R as such an endomorphism, mixed indices have to be used so that the result composition $R^a{}_b R^b{}_c$ is again a mixed-indices tensor of the same valence. $R_a{}^b$ would do as well, but with the former choice, the eigenvalue problem can be written as

$$R^a{}_b v^b = \lambda v^a, \tag{3}$$

i.e., for eigenvectors rather than eigenforms, which seems more natural. The two are still equivalent through the musical isomorphism, and such an R is a self-adjoint operator with regard to the metric

$$\begin{aligned} \langle u, R(v)\rangle &= u^a g_{ab} R^b{}_c v^c = u^a R_{ca} v^c \\ &= R^b{}_a u^a g_{bc} v^c = \langle R(u), v\rangle, \end{aligned} \tag{4}$$

provided that R_{ab} is symmetric, which is the case for the Levi-Cività connection. When one allows for the torsion to be nonzero, the above requires a generalization given in Sec. IV.

In the bracket-component notation, f should act on R considered as a linear operator with matrix elements $[R]^a{}_b$, and should also give as the result an operator, whose elements are denoted by $[f(R)]^a{}_b$. For example, for the composition with itself it is convenient to write $[R\cdot R]^a{}_b = [R^2]^a{}_b$, so the superscript 2 refers to the operator power, not a component. Accordingly, R will signify the (1, 1) valence tensor, and for the Ricci scalar the contraction $R^a{}_a$ or $\mathcal{R}$ will be used. After "bracketing," the index notation is recovered, which allows for raising and lowering; for brevity, the brackets will be omitted in the simplest cases such as $[R]_{ab} = R_{ab}$.

For any analytic $f:\mathbb{R}\to\mathbb{R}$, the following definition of the matrix function $f^*:\mathbb{R}^{n^2}\to\mathbb{R}^{n^2}$ can be used[1]:

$$\begin{aligned} [f^*(R)]^a{}_b &:= \sum_{n=0}^{\infty} f_n [R^n]^a{}_b \\ &= f_0 \mathbb{1}^a{}_b + f_1 R^a{}_b + f_2 R^a{}_s R^s{}_b \\ &\quad + f_3 R^a{}_s R^s{}_t R^t{}_b + \cdots, \end{aligned} \tag{5}$$

where

$$f(\xi) = \sum_{n=0}^{\infty} f_n \xi^n, \tag{6}$$

and the sums are written explicitly, as they are not tensor contractions (R^n is an operator power as explained above). The above requires that the spectral radius of $\rho(R) = \lim_{n\to\infty}\|R^n\|^{1/n}$ be less than the radius of convergence of the series $f(\xi)$.

For example, when $f = \exp$, the above two Lagrangians are

$$\begin{aligned} \mathcal{L}_0 &= \exp(\mathcal{R}) = 1 + \mathcal{R} + \frac{1}{2!}\mathcal{R}^2 + \frac{1}{3!}\mathcal{R}^3 + \cdots \\ &= 1 + R^a{}_a + \frac{1}{2!}(R^a{}_a)^2 + \frac{1}{3!}(R^a{}_a)^3 + \cdots, \\ \mathcal{L}_g &= \mathrm{tr}\left[\mathbb{1} + R + \frac{1}{2!}R^2 + \frac{1}{3!}R^3 + \cdots\right] \\ &= d + R^a{}_a + \frac{1}{2!}R^a{}_b R^b{}_a + \frac{1}{3!}R^a{}_b R^b{}_c R^c{}_a + \cdots, \end{aligned} \tag{7}$$

where d is the dimension of the space-time. Thus, the first essential deviation appears at the quadratic level and is proportional to $f_2(R^a{}_b R^b{}_a - (R^a{}_a)^2)$ if the same f is used in both approaches. The difference is also evident when the Lagrangians are written in terms of the eigenvalues of R:

$$\mathcal{L}_0 = f\left(\sum_i \lambda_i\right) \quad \text{vs} \quad \mathcal{L}_g = \sum_i f(\lambda_i). \tag{8}$$

A degeneracy in λ_i might then lead to the same theories, e.g., when the traceless Ricci tensor vanishes: $\hat{R}^a{}_b := R^a{}_b - \frac{1}{d}\mathcal{R}\mathbb{1}^a{}_b = 0$. The Ricci tensor is then proportional to the identity matrix and $[f(R)]^a{}_a = df(\mathcal{R}/d)$, which, up to a simple rescaling of f, is the same as the Lagrangian $\mathcal{L}_0$. However, one has to be careful when making such substitutions directly in the action, because R is determined only after having solved the Einstein equations. If the assumption $\hat{R} = 0$ is justified from the beginning, the two theories coincide. We shall see in the examples below that even in an empty universe this condition might not hold generally but just for isolated solutions.

Note also that if f is determined, there is no freedom of choice for its constant term f_0, which naturally corresponds to the cosmological constant. In other words, in such a nonperturbative interpretation, its value is tied to the whole expansion and cannot be adjusted independently. The expansion around $R = 0$ also shows that when f is almost linear, then higher-order terms can be ignored in the weak field limit, leading to the Einstein-Hilbert action and a small perturbation of general relativity.

Although intuitive, the above definition is not very convenient when a function is real analytic but has complex singularities like $\tanh(\xi)$. A definition better suited for the

[1] As with R, it will be convenient to sometimes write the identity operator without indices, so in order to avoid confusion with the variation δ, I will use $\mathbb{1}^a{}_c$ instead of the Kronecker symbol $\delta^a{}_c$.

situation at hand is an elegant generalization of Cauchy's formula[2]

$$f(R) \coloneqq \frac{1}{2\pi\mathrm{i}}\int_C (\xi\mathbb{1} - R)^{-1} f(\xi)\mathrm{d}\xi \tag{9}$$

for a contour C which encloses the spectrum of R but not the singularities of $f(\xi)$. The two definitions agree for fairly general assumptions, and for a function that is real on the real axis, the matrix $f(R)$ will also be real [8].

The dimension of $f(R)$ affects how the function is given, because R has the units of curvature, and so should the Lagrangian. At first, it seems two constants are necessary to give $f(R) = C_0\tilde{f}(R/C_1)$ in terms of a function $\tilde{f}$ which only contains dimensionless parameters, but this can be rewritten as

$$f(R) = C_0\tilde{f}\left(\frac{C_0}{C_1}\frac{R}{C_0}\right) \to C_0\tilde{f}(R/C_0), \tag{10}$$

with a redefined dimensionless $\tilde{f}$. The remaining constant C_0 can then be further rescaled using the cosmological or the Hubble constant depending on context—this is done in Sec. V.

Having defined $\mathrm{tr}[f(R)]$, the total action, including the matter Lagrangian $\mathcal{L}_M$, is taken to be

$$\mathcal{S} = \int \left(\frac{1}{16\pi\mathcal{G}}\mathcal{L}_g + \mathcal{L}_M\right)\sqrt{-g}\mathrm{d}^4x, \tag{11}$$

where $\mathcal{G}$ is the gravitational constant, and the modified Einstein equations can then be obtained in one of the two standard ways. One is to assume the Levi-Cività connection and take the metric as the dynamical variable; the other is to consider both the metric and the connection as dynamical. The former is called the metric and the latter the Palatini formulation (or, more generally, metric-affine).

In both cases, the variation of the $f(R)$ term is needed, and the second definition of a tensor function allows us to easily calculate it as

$$\begin{aligned}\delta\mathrm{tr}[f(R)] &= \mathrm{tr}\left[\frac{1}{2\pi\mathrm{i}}\int_C (\xi\mathbb{1} - R)^{-1}\delta R(\xi\mathbb{1} - R)^{-1} f(\xi)\mathrm{d}\xi\right] \\ &= \mathrm{tr}\left[\frac{1}{2\pi\mathrm{i}}\int_C (\xi\mathbb{1} - R)^{-2} f(\xi)\delta R\mathrm{d}\xi\right] \\ &= \mathrm{tr}\left[\frac{1}{2\pi\mathrm{i}}\int_C (\xi\mathbb{1} - R)^{-1} f'(\xi)\mathrm{d}\xi\delta R\right] = \mathrm{tr}[f'(R)\delta R],\end{aligned} \tag{12}$$

where the cyclic property of trace

$$\mathrm{tr}[X_1X_2\ldots X_k] = \mathrm{tr}[X_kX_1X_2\ldots X_{k-1}] \tag{13}$$

was used in the first line, and integration by parts was used in the second. Reexpressing δR with δg and $\delta\Gamma$ to arrive at the modified Einstein equation is the subject of the next two sections.

A. Definitions and notation

To shortly review the conventions used, the covariant derivative and the connection coefficients in a basis $\{e_a\}$ are related through

$$\nabla_{e_a} e_b = \Gamma^c{}_{ba} e_c, \tag{14}$$

so that for a coordinate basis $e_a = \partial_a$, one has

$$\nabla_a X^b = \partial_a X^b + \Gamma^b{}_{ca} X^c. \tag{15}$$

As Γ will not in general be symmetric in the lower indices, care needs to be taken regarding their order. The antisymmetric part of the connection defines the torsion as

$$T(X,Y) \coloneqq \nabla_X Y - \nabla_Y X - [X,Y] = e_a T^a{}_{bc} X^b Y^c, \tag{16}$$

and in a coordinate basis, where $[\partial_a, \partial_b] = 0$, it follows that

$$T^a{}_{bc} = \Gamma^a{}_{cb} - \Gamma^a{}_{bc}. \tag{17}$$

The Riemann tensor is given by[3]

$$\mathsf{R}(X,Y)Z \coloneqq \nabla_{[X}\nabla_{Y]}Z - \nabla_{[X,Y]}Z = e_d \mathsf{R}^d{}_{abc} Z^a X^b Y^c, \tag{18}$$

or, in term of components in a coordinate basis,

$$\mathsf{R}^d{}_{abc} = \partial_b\Gamma^d{}_{ac} - \partial_c\Gamma^d{}_{ab} + \Gamma^d{}_{sb}\Gamma^s{}_{ac} - \Gamma^d{}_{sc}\Gamma^s{}_{ab}, \tag{19}$$

and the Ricci tensor is the contraction

$$R_{ab} \coloneqq \mathsf{R}^c{}_{acb}. \tag{20}$$

Note, then, that although R_{ab} is constructed solely with the connection (curvature), for the operator $R^a{}_b = g^{ac}R_{cb}$ the metric is necessary. Finally, the signature will be taken to be $(-,+,+,+)$, and the speed of light equal to unity, so that coordinates have the dimension of length, and the metric itself is dimensionless.

III. THE METRIC APPROACH

The natural connection solely determined by the metric through $\nabla_a g_{bc} = 0$ and $T^a{}_{bc} = 0$ is the Levi-Cività connection. Its variation, as expressed by δg, is

$$\delta\Gamma^c{}_{ba} = \frac{1}{2}g^{cd}(\nabla_b\delta g_{ad} + \nabla_a\delta g_{db} - \nabla_d\delta g_{ba}), \tag{21}$$

[2] In what follows, f and f^* can safely be treated as the same object, so the star will be dropped.

[3] The brackets involving vectors denote commutation, not antisymmetrization—i.e., there is no prefactor of $\frac{1}{2}$.

and in turn for the covariant Ricci tensor, one has

$$\delta R_{ab} = \nabla_c(\delta\Gamma^c{}_{ab}) - \nabla_b(\delta\Gamma^c{}_{ac}), \tag{22}$$

which accordingly gives

$$\begin{aligned}\delta R_{ab} =& \frac{1}{2}(\nabla^d\nabla_b\delta g_{ad} + \nabla^d\nabla_a\delta g_{db} - \nabla^d\nabla_d\delta g_{ba}) \\ &+ -\frac{1}{2}(g^{cd}\nabla_b\nabla_a\delta g_{dc} + \nabla_b\nabla^d\delta g_{ad} - \nabla_b\nabla_d g^{cd}\delta g_{ca}) \\ =& \frac{1}{2}(\nabla^d\nabla_b\delta g_{ad} + \nabla^d\nabla_a\delta g_{bd} - \Box\delta g_{ab} - g^{cd}\nabla_b\nabla_a\delta g_{cd}).\end{aligned} \tag{23}$$

Next, by observing that

$$0 = \delta(\mathbb{1}^a{}_c) = g_{bc}\delta g^{ab} + g^{ab}\delta g_{bc}, \tag{24}$$

the variation of the operator R becomes

$$\delta R^a{}_b = g^{ac}(\delta R_{cb} - R^s{}_b\delta g_{cs}), \tag{25}$$

leading to

$$\begin{aligned}&\delta(\mathrm{tr}[f(R)]\sqrt{-g}) \\ &= \left([f'(R)]^a{}_b\delta R^b{}_a + \frac{1}{2}\mathrm{tr}[f(R)]g^{bd}\delta g_{bd}\right)\sqrt{-g} \\ &= \left([f'(R)]^{ac}\delta R_{ca} - [Rf'(R)]^{cd}\delta g_{dc} \right. \\ &\quad \left. + \frac{1}{2}\mathrm{tr}[f(R)]g^{bd}\delta g_{bd}\right)\sqrt{-g}.\end{aligned} \tag{26}$$

The variation δR_{ab} of (23) can be substituted into the above, and due to $\sqrt{-g}\nabla_a X^a = \partial_a(\sqrt{-g}X^a)$, each term containing the covariant derivative can be integrated by parts, provided that the variations vanish at the boundary or that the boundary is empty. The result is

$$\begin{aligned}&\delta(\mathrm{tr}[f(R)]\sqrt{-g}) \\ &= (\nabla_c\nabla^d[f'(R)]^{cb} - \frac{1}{2}\Box[f'(R)]^{bd} - \frac{1}{2}\nabla_a\nabla_c[f'(R)]^{ac}g^{bd} \\ &\quad - [Rf'(R)]^{bd} + \frac{1}{2}[f(R)]^a{}_a g^{bd})\sqrt{-g}\delta g_{bd}.\end{aligned} \tag{27}$$

Finally, defining the stress-energy tensor $\mathcal{T}$ by

$$\frac{\delta(\sqrt{-g}\mathcal{L}_M)}{\delta g_{bd}} =: \frac{1}{2}\mathcal{T}^{bd}\sqrt{-g}, \tag{28}$$

the condition $\delta\mathcal{S} = 0$ gives the following modified Einstein equations:

$$\boxed{\frac{1}{2}\Box[f'(R)]_{bd} - \nabla_c\nabla_b[f'(R)]^c{}_d + \frac{1}{2}\nabla^a\nabla^c[f'(R)]_{ac}g_{bd} + [Rf'(R)]_{bd} - \frac{1}{2}\mathrm{tr}[f(R)]g_{bd} = 8\pi\mathcal{G}\mathcal{T}_{bd}.} \tag{29}$$

As can be seen, the last two terms on the left-hand side reduce to the standard Einstein tensor for $f = \mathrm{Id}$, whereas the other terms are zero, since $f' = 1$.

IV. THE PALATINI APPROACH

In the more general case, the connection is independent of the metric, and there are two assumptions that can be relaxed here: vanishing torsion and metric compatibility. In general, the connection can be decomposed into the sum

$$\begin{aligned}\Gamma^a{}_{bc} &= \tilde{\Gamma}^a{}_{bc} + K^a{}_{bc} - C^a{}_{bc}, \\ K_{abc} &:= -\frac{1}{2}(T_{abc} + T_{bca} - T_{cab}),\end{aligned} \tag{30}$$

where $\tilde{\Gamma}$ is the Levi-Cività connection for g, K is called the contorsion tensor, and C describes the nonmetricity

$$C_{abc} := \frac{1}{2}(\nabla_c g_{ab} + \nabla_b g_{ca} - \nabla_a g_{bc}). \tag{31}$$

Accordingly, the variation of the Ricci tensor is now

$$\delta R_{ab} = \nabla_c(\delta\Gamma^c{}_{ab}) - \nabla_b(\delta\Gamma^c{}_{ac}) - T^d{}_{bc}\delta\Gamma^c{}_{ad}, \tag{32}$$

and neither the connection coefficients nor the Ricci tensor are symmetric in the lower indices. The eigenvalues of R might not be real anymore, in which case they appear in conjugate pairs. This means that the trace of $f(R)$ will still be real, for real analytic f.

There is, however, a possible natural generalization, because of the following identity[4]:

$$R_{ab} = R_{ba} + \nabla_{\underline{a}}T^c{}_{\underline{cb}} + T^c{}_{cd}T^d{}_{ab}, \tag{33}$$

[4]The underline denotes the sum over cyclic permutations.

which leads to the introduction of a new tensor, which is the symmetric part of R,

$$S_{ab} := R_{ab} - \frac{1}{2}(\nabla_{\underline{a}} T^c{}_{\underline{cb}} + T^c{}_{cd} T^d{}_{ab}). \tag{34}$$

These tensors have the same trace, so there is no need for S_{ab} in the standard $f(\mathcal{R})$ theories—the trace cancels the imaginary parts of the conjugate pairs of the eigenvalues. Here, the situation is different, because the function f is applied to the eigenvalues of R before the trace is taken, so although the final result is real, it also depends on the imaginary parts. The other reasons and equations for the $f(S)$ variant are given following the $f(R)$ derivation below.

In contrast to the preceding section, only first derivatives are present in the action, and the integration by parts requires an additional term, because the torsion affects the expression for covariant divergence:

$$\sqrt{-g}\nabla_a(X^a) = \partial_a(\sqrt{-g}X^a) + \sqrt{-g}(T^b{}_{ba} - C^b{}_{ba})X^a. \tag{35}$$

The total variation of the Lagrangian then becomes

$$\begin{aligned}\delta(\mathrm{tr}[f(R)]\sqrt{-g}) &= \left(P^{ba}\delta R_{ab} - [Rf'(R)]^{db}\delta g_{bd} + \frac{1}{2}[f(R)]^a{}_a g^{bd}\delta g_{bd}\right)\sqrt{-g}\\ &= \left(\frac{1}{2}\mathrm{tr}[f(R)]g^{bd} - [Rf'(R)]^{db}\right)\delta g_{bd}\sqrt{-g} + (\nabla_b P^{ba}\mathbb{1}^d{}_c - \nabla_c P^{da} - T^d{}_{bc}P^{ba}\\ &\quad + (C^s{}_{sb} - T^s{}_{sb})P^{ba}\mathbb{1}^d{}_c - (C^s{}_{sc} - T^s{}_{sc})P^{da})\delta\Gamma^c{}_{ad}\sqrt{-g},\end{aligned} \tag{36}$$

where the derivative tensor is denoted by $P_{ab} := [f'(R)]_{ab}$ for brevity.

In addition to the stress-energy tensor $\mathcal{T}$, a new quantity is necessary to reflect the fact that matter fields can, in general, depend on the connection—if only through the covariant derivative. The hyper-momentum tensor is defined thus:

$$\sqrt{-g}\mathcal{Q}_a{}^{bc} := \frac{\delta(\sqrt{-g}\mathcal{L}_M)}{\delta\Gamma^a{}_{bc}}, \tag{37}$$

and the modified Einstein equations can now be written as

$$\boxed{8\pi\mathcal{G}\mathcal{T}_{bd} = [Rf'(R)]_{(db)} - \frac{1}{2}\mathrm{tr}[f(R)]g_{bd}, \quad 8\pi\mathcal{G}\mathcal{Q}_c{}^{ad} = \nabla_b(\mathbb{1}^{[b}{}_c P^{d]a}) - (T-C)^s{}_{sb}\mathbb{1}^{[b}{}_c P^{d]a} - \frac{1}{2}T^d{}_{cb}P^{ba},} \tag{38}$$

where the symmetrization is necessary, because the variation δg_{bd} is symmetric, even though R_{bd} is not.

The second set of equations can be simplified if an auxiliary connection is defined to be

$$\hat{\Gamma}^a_{bc} := \Gamma^a{}_{bc} - \frac{1}{2}\mathbb{1}^a{}_b(T-C)^s{}_{sc}, \tag{39}$$

and using the associated covariant derivative $\hat{\nabla}$, the second set of Einstein equations reads

$$8\pi\mathcal{G}\mathcal{Q}_c{}^{ad} = \hat{\nabla}_b(\mathbb{1}^{[b}{}_c P^{d]a}) - \frac{1}{2}T^d{}_{cb}P^{ba}. \tag{40}$$

Additionally, contraction over the pair of indices $\{cd\}$ leads to

$$3\hat{\nabla}_b P^{ba} + T^s{}_{sb}P^{ba} + 16\pi\mathcal{G}\mathcal{Q}_s{}^{as} = 0, \tag{41}$$

which allows us to rewrite the main equations as

$$\begin{aligned}&[Rf'(R)]_{(db)} - \frac{1}{2}\mathrm{tr}[f(R)]g_{bd}\\ &\quad = 8\pi\mathcal{G}\mathcal{T}_{bd},\\ &\hat{\nabla}_c P^{da} - \left(T^d{}_{cb} - \frac{1}{3}T^s{}_{sb}\mathbb{1}^d{}_c\right)P^{ba}\\ &\quad = 16\pi\mathcal{G}\left(\mathcal{Q}_c{}^{ad} - \frac{1}{3}\mathcal{Q}_s{}^{as}\mathbb{1}^d{}_c\right).\end{aligned} \tag{42}$$

As in the ordinary $f(R)$ formulation, the torsion equations become algebraic for the Einstein-Hilbert case $f(R) = R$ because $P^{ab} = [f'(R)]^{ab} = g^{ab}$, so that derivatives of Γ only appear in R. Further, if the matter fields are such that $\mathcal{Q}_{abc} \equiv 0$, contractions of the torsion equations give

$$\begin{aligned}3C^s{}_{sa} &= -6C_{as}{}^s = 2T^s{}_{sa},\\ 2C_{(ad)c} &= T_{dac} + \frac{2}{3}T^s{}_{s(a}g_{c)d}.\end{aligned} \tag{43}$$

This means that if $T^s{}_{sa} = 0$, then $2C_{(ad)c} = T_{dac}$, and it follows immediately from (30) that $K_{abc} = C_{abc}$. But that, by definition, means the connection must be the Levi-Cività one.

In other words, for $f(R) = R$, zero hyper-momentum and totally antisymmetric torsion, the theory becomes standard general relativity. Note that for this to happen, it is not necessary to assume zero torsion from the beginning, just that all its traces vanish.

Since the Ricci tensor is, in general, no longer symmetric, the tensor P cannot be used directly to define a new metric for which Eq. (42) would define a metric connection. In the standard $f(\mathcal{R})$ theories, the tensor that enters is R itself, and it can be decomposed into (anti)symmetric parts at the level of the Einstein equations, as the function f is applied only to its trace, and all $f(R^a{}_a)$ terms are just scalars.

Here, the situation is different, in that even in the first set of equations the symmetrization is applied to $Rf'(R)$, not to R, and the second set of equations contains $f(R)$, not $f'(R)$. Because even for the second power one has $g^{bc}X_{c(d}X_{a)b} \neq X_{(ab)}g^{bc}X_{(cd)}$, symmetrizing the equations would not lead to a single distinguished tensor to be used as the new metric. Moreover, even though the components of R are real, it seems natural to consider a self-adjoint matrix, for which the action is directly related to the eigenvalues as in (8).

These problems could be overcome by constructing the action with the symmetric tensor S, introduced before, whose variation is simply $\delta S_{ab} = \frac{1}{2}(\delta R_{ab} + \delta R_{ba})$. The derivation is essentially the same as in (36), and the difference is that the tensor contracted with δg is already symmetric, so the Einstein equations are

$$8\pi\mathcal{G}\mathcal{T}_{bd} = [Sf'(S)]_{db} - \frac{1}{2}\mathrm{tr}[f(S)]g_{bd},$$
$$8\pi\mathcal{G}\mathcal{Q}_c{}^{ad} = \hat{\nabla}_b(\mathbb{1}^{[b}{}_c P^{d]a}) - \frac{1}{2}T^d{}_{cb}P^{ba}, \tag{44}$$

where now, by a slight abuse of notation, $P_{ab} = [f'(S)]_{ab}$, and the auxiliary covariant derivative is the one given by Eq. (39).

As before, the trace can be used to rewrite the second equation as (42), and following the same reasoning as for the standard $f(R)$ derivation, the torsionless connection with no hyper-momentum yields

$$\hat{\nabla}_c P^{da} = 0. \tag{45}$$

This would indicate that $\hat{\Gamma}$ is the Levi-Cività connection for the metric P^{da}, but the situation is complicated by the fact that the tensor $P = f'(R)$ is not conformally related to the original metric g, so the signature might not be the same, and the determinant of g is not directly proportional to that of P; also, the raising of indices in P does not amount to matrix inversion. It should also be kept in mind that with the standard extension of covariant derivative to tensor densities, which uses $\sqrt{|g|}$ to cancel the weight, the above equation can be rewritten as

$$\frac{1}{\sqrt{|g|}}\hat{\nabla}_c(\sqrt{|g|}P^{da}) = 0, \tag{46}$$

but this is not equivalent to

$$\frac{1}{\sqrt{|\det P|}}\hat{\nabla}_c(\sqrt{|\det P|}P^{da}) = 0, \tag{47}$$

unless $\sqrt{\det P}$ is used to extend $\hat{\nabla}$ to densities. Without specifying which extension is used, the condition $\nabla_c(\sqrt{|g|}g_{ab}) = 0$ does not necessarily indicate metricity, contrary to what can sometimes be found in the literature. Because of this freedom, and given the problems with inverting P, the more fundamental Eq. (45) is better as an indication of a metric connection in the present case.

With some effort, the Christoffel formula can be used to express $\hat{\Gamma}$ as a function of derivatives of P, but the derivatives of the connection coefficients are still involved in the nonlinear term $f'(S)$. The question is then whether they can be eliminated with the help of the remaining equations.

In the standard approach, the first set of the Einstein equations (44) can, in principle, be used to solve for the Ricci scalar and accordingly simplify the second set by using the Ricci tensor associated with the new metric and its Levi-Cività connection [3]. Here, one would have to solve nonlinear equations for the whole tensor S in order to eliminate the connection in the same manner. At present, it appears that this path of investigation is not applicable, because the equations involve full tensors R or S, not just their traces.

V. FRW DYNAMICS

The standard cosmological model is the basic example that needs to be considered in order to gain insight into the applicability of the proposed modification. The model assumes spatial homogeneity and isotropy, requiring the Robertson-Walker geometry, which in spherical coordinates $\{t, r, \theta, \varphi\}$ has the metric

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right) = g_{ab}\mathrm{d}x^a\mathrm{d}x^b, \tag{48}$$

where $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi$ is the standard metric on the unit sphere. The final assumption in this first attempt at modified cosmology will be that the RW metric provides the only dynamical variable—the scale factor $a(t)$—the connection is that of Levi-Cività, and the metric formalism can be used.

Accordingly, the matter source will be taken to be a homogeneous perfect fluid with density ρ and pressure p, so that the stress energy tensor is

$$\mathcal{T}_{ab} = pg_{ab} + (p+\rho)u_a u_b, \tag{49}$$

where the four-velocity in these coordinates is just $u = \partial_t$.

There are then effectively only two modified Einstein equations, one of third order and one of fourth corresponding to the $\mathcal{T}_{00}$ and $\mathcal{T}_{11}$ components of (29), respectively. However, the latter follows from the derivative of the former, which is the generalization of the Friedmann equation

$$H(3f''(\lambda_0) + f''(\lambda_1))\frac{\mathrm{d}\lambda_0}{\mathrm{d}t} = 16\pi\mathcal{G}\rho + 6H^2((\lambda_1 - \lambda_0)f''(\lambda_1) + f'(\lambda_1) - f'(\lambda_0)) + \lambda_0(f'(\lambda_0) + f'(\lambda_1)) - f(\lambda_0) - 3f(\lambda_1), \tag{50}$$

where λ are the eigenvalues of R

$$\lambda_0 = 3\frac{\ddot{a}}{a}, \qquad \lambda_1 = 2H^2 + \frac{2k}{a^2} + \frac{\ddot{a}}{a}, \tag{51}$$

H is the Hubble "constant" $H = \dot{a}/a$, and an overdot denotes the time derivative.

The present value of the constant, $H_0 \coloneqq H(0)$, is customarily used to obtain dimensionless quantities and, as discussed in Sec. II, there is still an unspecified constant in the function f. Although H_0^2 has the suitable dimension, it will not do as C_0, because the function f should be a fundamental quantity valid for all gravitational actions, not just the FRW cosmology, and thus cannot be defined with such specific constants. Instead, C_0 will become a physical parameter of the new theory, and the Hubble constant H_0 will serve to provide the dimensionless counterpart $c_0 \coloneqq C_0 H_0^{-2}$.

Of course, the roles could be reversed, with C_0 used instead of H_0, but for initial clarity it is better to keep with the convention of rescaling densities, time, etc., with H_0. The dimensionless eigenvalues are then

$$\alpha \coloneqq \lambda_0 H_0^{-2}, \qquad \beta \coloneqq \lambda_1 H_0^{-2}, \tag{52}$$

which gives e.g. $f(\lambda_0) = f(\alpha H_0^{-2})$ and leads to further simplification,

$$H_0^{-2} f(\lambda_0) = c_0 \tilde{f}(\alpha/c_0) \eqqcolon F(\alpha), \tag{53}$$

and similarly for β. The main equation can then be rewritten as

$$\begin{aligned} h(3F''(\alpha) + F''(\beta))\frac{\mathrm{d}\alpha}{\mathrm{d}\tau} \\ = 6\Omega + 6h^2((\beta - \alpha)F''(\beta) + F'(\beta) - F'(\alpha)) \\ + \alpha(F'(\alpha) + F'(\beta)) - F(\alpha) - 3F(\beta), \end{aligned} \tag{54}$$

where h, the density parameter, and dimensionless time are defined by

$$h \coloneqq \frac{H}{H_0}, \qquad \Omega \coloneqq \frac{8\pi\mathcal{G}\rho}{3H_0^2}, \quad \text{and} \quad \tau \coloneqq H_0 t. \tag{55}$$

The function F can then be specified with any suitable number of dimensionless parameters including c_0. It could be considered to be given *a priori* by some elementary function like $A\sin(B\xi)$, or defined by infinitely many expansion coefficients as the series (6). Yet to consider such coefficients as independent parameters would be to multiply entities beyond necessity, so I will adopt the former approach here.

A quantitative reason can also be given for this, in anticipation of the observational analysis. Finding the coefficients from the data would undoubtedly lead to better and better fits as the number of coefficients increases, but such a fit would come with a huge cost as measured by the Akaike or Bayesian information criteria, which are now standard tools of observational cosmology [9,10].

As for the nature of parameters in the present case, some more information can be gleaned from the zeroth- and first-order expansions of F, as they reproduce the standard model with the cosmological constant. The general form[5] is $F(\xi) = F_0 + F_1\xi$, but the overall rescaling of the Lagrangian is not important, and taking $F_1 = 1$ gives the ordinary Friedmann equation

$$H^2 = \frac{8\pi\mathcal{G}\rho}{3} - \frac{k}{a^2} - \frac{2}{3}H_0^2 F_0 \tag{56}$$

upon identifying the cosmological constant $\Lambda = -2H_0^2 F_0$. In terms of the original function f, this means that $f_0 = -\Lambda/2$, and it suggests that the cosmological constant itself could be used as a fundamental dimensional quantity by

$$f(\xi) = \frac{\Lambda}{2}\tilde{f}\left(\frac{2\xi}{\Lambda}\right), \tag{57}$$

[5] As before, ξ is just an auxiliary independent variable used to define functions and their rescalings.

with $\tilde{f}$ carrying no other free parameters. Using the respective density parameter $\Omega_\Lambda \coloneqq \Lambda/3$, this means that

$$F(\alpha) = \frac{1}{H_0^2} f(\lambda_0) = \frac{3\Omega_\Lambda}{2}\tilde{f}\left(\frac{2\alpha}{3\Omega_\Lambda}\right), \tag{58}$$

where the expansion of $\tilde{f}$ is then necessarily restricted to

$$\tilde{f}(\xi) = -1 + \xi + \mathcal{O}(\xi^2). \tag{59}$$

Turning now to the dynamics of this model, a minimal set of variables yielding a closed system can be built from the derivatives of $a(t)$, or rather their rescaled versions h and α, which are identically related by $\alpha = 3(\dot{h} + h^2)$. Also, the other of the eigenvalues can be eliminated through

$$\beta = 2h^2 + \frac{2\Omega_k}{a^2} + \frac{\alpha}{3}, \quad \text{with} \quad \Omega_k \coloneqq \frac{k}{H_0^2}, \tag{60}$$

although for shorter notation it will be better to keep the symbol β and understand it as a function of a, h and α, which will be the replacements for $\dot{a}$, $\ddot{a}$ and $\dddot{a}$.

Because the conservation law $\nabla^a \mathcal{T}_{ab} = 0$ still holds, the matter-energy density ρ is expressible in terms of a if one assumes an equation of state

$$p = (\gamma - 1)\rho \Rightarrow \rho(t) = \rho_0 a(t)^{-3\gamma} \Rightarrow \Omega = \sum_j \Omega_j a^{-3\gamma_j}. \tag{61}$$

Finally, introducing

$$W \coloneqq 6h^2((\beta - \alpha)F''(\beta) + F'(\beta) - F'(\alpha)) + \alpha(F'(\alpha) + F'(\beta)) - F(\alpha) - 3F(\beta), \tag{62}$$

for the sake of brevity, a dynamical system with three degrees of freedom described by the variables $\{\alpha, h, a\}$ is obtained:

$$\begin{cases} \dot{\alpha} = \frac{6\Omega + W(\alpha,\beta,h,a)}{(3F''(\alpha)+F''(\beta))h} & \eqqcolon v_1(\alpha,h,a), \\ \dot{h} = \frac{1}{3}\alpha - h^2 & \eqqcolon v_2(\alpha,h,a), \\ \dot{a} = ah & \eqqcolon v_3(\alpha,h,a), \end{cases} \tag{63}$$

where the dot now refers to the new time τ. Note that the denominator of v_1 would only be identically zero for the purely linear F, which is the standard general relativity. The form of v_2 and v_3 is dictated by the definition of h, and the essential dynamics lies with v_1. This is also where we find the difference in complexity between the new theory and $f(\mathcal{R})$, for which v_2 and v_3 are the same, but the first equation would read

$$\dot{\alpha} = 3(\beta - \alpha)h + \frac{6\Omega - F(3\beta+\alpha) + 2\alpha F'(3\beta+\alpha)}{12hF''(3\beta+\alpha)}. \tag{64}$$

The difference between the two equations in the simplest quadratic case $F(\xi) = -\frac{3}{2}\Omega_\Lambda + \xi + F_2\xi^2$ is just $(\Omega_\Lambda + \Omega - h^2)/(2F_2 h)$, which is nonzero exactly when the evolution deviates from the Friedmann equation. As was mentioned in Sec. II, if $\hat{R}$ vanishes, then a simple rescaling of F also leads to the same equations, but in this particular geometry the condition is very restrictive. For flat universes (as in the examples below), the only solutions with this property are the de Sitter ones, $h = \text{const}$, which do not exhaust all possible solutions, even when $\Omega = 0$. On the other hand, the difference disappears completely if we take different functions: $F(\xi) = F_1 + \xi + F_2\xi^2$ for $f(R)$ and $\tilde{F}(\xi) = 4F_1 + \xi + 3F_2\xi^2$ for $f(\mathcal{R})$; the theories are equivalent for the Robertson-Walker geometry at the quadratic level, even when $\hat{R} \neq 0$. However, no such simple relation could be found for cubic terms.

A general feature of the main system (63) is that if the geometry is flat, i.e., $k = 0$ and the density does not depend on the scale factor, like for the cosmological constant, then the first two equations decouple and give a planar system. In fact, one could simply assume that no ordinary matter enters the equations as Ω, but instead consider the higher-order terms of F as some sort of field imitating matter. For example, if $F(\xi) = -\frac{3}{2}\Omega_f + \xi + \frac{1}{2}F_2\xi^2$, the main equation (54) becomes

$$h^2 = \Omega_f + F_2\left(\dot{h}^2 - 2h\ddot{h} - \frac{14}{3}h^2\dot{h} + \frac{16}{9}h^4\right), \tag{65}$$

so that Ω_f acts as dark energy and the F_2 term acts as effective material content.

Another general, and problematic, feature of the $\dot{\alpha}$ equation is the singularity at $h = 0$, i.e., when expansion changes to contraction and vice versa. This is not a singularity of Eq. (54) and can lead to a valid solution provided that the numerator of v_1 vanishes as well. Thus, care has to be taken when using the dynamical system form, because the singularities might simply signify that the left-hand side of the original equation is zero, and vice versa: a zero of v_1 might in fact be a singularity of the original equation (63).

A. Examples of cosmological models

A very basic example illustrating these features is to take a flat, empty universe and assume the exponential function

$$F(\xi) = \Omega_f e^{\frac{\xi}{\Omega_f}} - 1 = \xi + \frac{\xi^2}{2\Omega_f} + \mathcal{O}(\xi^3), \tag{66}$$

which includes the linear action, but no cosmological constant in the usual sense. The specific form of v_1 is then

$$v_1\left(3\Omega_f\alpha,\sqrt{\frac{1}{2}\Omega_f}h,a\right)=\frac{4e^{-\alpha}+e^{2\alpha}(3\alpha-3h^2-1)+3e^{h^2}(h^4+(1-2\alpha)h^2+\alpha-1)}{\sqrt{2\Omega_f^3}(3e^{2\alpha}+e^{h^2})h},\tag{67}$$

where the additional factors in the arguments are only introduced to shorten the formula. It is still essentially transcendental, so one has to resort to qualitative analysis first to locate the points and regions of interest. This can be done with the help of Fig. 1, which shows the planar vector field (v_1, v_2) together with the locations of singular lines and zeros of the right-hand side v (left panel), and the phase portrait constructed from typical trajectories (right panel); the particular value of $\Omega_f = 3/2$ was chosen.

The left and right saddle points A_1 and A_2 correspond to time-reversed de Sitter and standard de Sitter solutions, respectively, and their positions $(h_0, 3h_0^2)$ are given by $h_0 = \pm\sqrt{2\Omega_f w/3}$, where w is the positive solution of $e^{-2w} + w = 1$.

The singular critical point B_1 could be considered as a static solution, because it lies on the singular line $h = 0$, but also on the $W = 0$ line, so in fact Eq. (54) is satisfied. For the vector field, on the other hand, the limit at B_1 is not well defined, as it depends on the path.

Importantly, there are no periodic orbits on either side of B_1, as the line $h = 0$ separates the neighborhood of B_1 into two elliptic sectors of opening π. The "closed" trajectories have B_1 as their limit point, so they are asymptotically static both in the past and in the future.

More physically realistic evolutions here seem to consist of trajectories that are attracted by A_2 and subsequently scattered along the unstable direction towards infinity. These are expanding universes with ever increasing acceleration, and also with initial singularity, which can be read from the phase portrait: going back in time, the trajectory has increasingly negative α, and discarding the exponentially small terms for large h and α, the right-hand side is approximately

$$\dot{\alpha} = 12h^3, \qquad \dot{h} = -h^2,\tag{68}$$

making α and h diverge in finite (negative) time.

There are also two mixed cases—i.e., trajectories going from a big bang becoming asymptotically static as they tend to B_1 and vice versa: asymptotically static in the past, but then getting scattered by A_2 into accelerated expansion. These exemplary behaviors of the scale factor and the Hubble constant are plotted in Fig. 2. Note that the time integration constant τ_0 such that $a(\tau_0) = 1$ cannot always be chosen to make $h(\tau_0) = 1$, so it is adjusted for each trajectory for better visibility in this and subsequent graphs.

It is probably more instructive to consider a more intricate model, which is furnished by taking a rational function

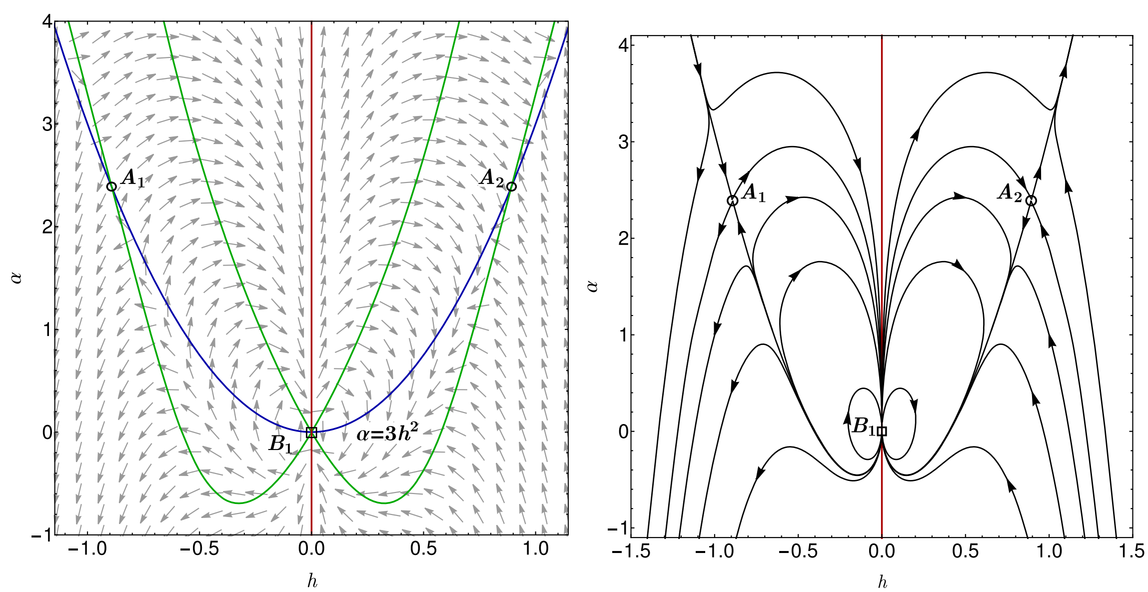

FIG. 1. The vector field (63) and its phase space diagram for a flat, empty universe with exponential Lagrangian function (66) and $\Omega_f = 3/2$. The green and red lines represent zero sets of the numerator and denominator of v_1, respectively. The blue parabola corresponds to $v_2 = 0$. Because of huge variation, the vector lengths are not drawn to scale to better show the discontinuity of direction at the singular line.

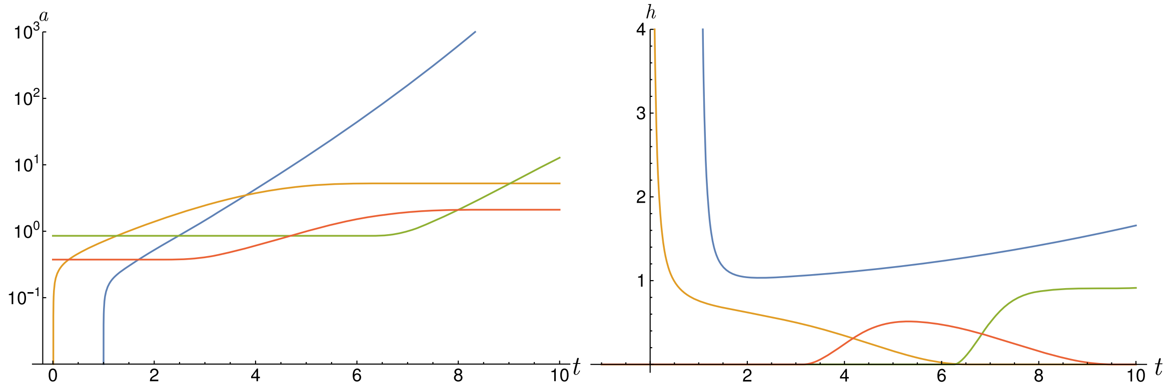

FIG. 2. Behavior of a and h for typical flat, empty universes corresponding to Eq. (66). Blue and orange curves are trajectories which have infinite h in the past, but the former escapes to infinite h while the latter is trapped by B_1. Green and red curves both start at B_1 in the infinite past, but the former escapes while the latter is recaptured.

$$\tilde{f} = -1 + \frac{\xi}{1-\xi^2} \Rightarrow F(\xi) = -\frac{3\Omega_\Lambda}{2} + \xi + \frac{4\xi^3}{9\Omega_\Lambda^2} + \mathcal{O}(\xi^5), \tag{69}$$

which includes the constant term, so it can be identified with the cosmological constant as in Eq. (58). Note that if the series were to be used, different expansions in different regions would be required. The reduction of the resulting powers of R with the characteristic polynomial would have to be carried out separately, which would lead to cumbersome expressions—if it were possible to obtain closed ones at all.

Direct substitution of this F into (63) produces a v_1 which is several lines long, so it is perhaps best to skip its specific form and, similarly to before, view the vector field and the various singular lines of the phase space; they are shown in the left panel of Fig. 3. The picture is now considerably more complex, with many more singular points of type B, for which both the numerator and denominator in $\dot{\alpha}$ vanish. These points signify a possible crossings through the otherwise impassable barriers indicated by the red lines.

There are still only two critical points A_1, A_2 located at $\left(\mp\sqrt{\frac{1}{2}\Omega_\Lambda}, \frac{3}{2}\Omega_\Lambda\right)$, which are asymptotic equilibria, and as before, they correspond to time-reversed de Sitter and standard de Sitter solutions, respectively. However, as the phase diagram of Fig. 3 shows, there are now two heteroclinic trajectories connecting them, one through B_1 at $(0, \frac{9}{2}\Omega_\Lambda)$ and the other through B_2 at $(0, \frac{3}{2}\Omega_\Lambda)$.

There is a complication here, not present in the previous example, though. The horizontal green lines at $\pm\frac{3}{2}\Omega_\Lambda$ are singularities of F, and so also of the Friedmann equation, but they cancel out in v_1, resulting in the straight-line trajectories. These are not singularities of curvature either, because α and h remain finite, so if one considers the action principle as purely formal to obtain the dynamical equations, these solutions could have some physical meaning.

A similar situation is found for the pair B_3 and B_4 located at $(\pm\sqrt{\Omega_\Lambda}, -\frac{3}{2}\Omega_\Lambda)$, except that the whole line can be thought of as just one trajectory for which h goes from ∞ to $-\infty$ in finite time. On both lines, the second equation $\dot{h} = v_2$ can be integrated to give

$$h = \sqrt{\pm\frac{1}{2}\Omega_\Lambda}\tanh\left(\sqrt{\pm\frac{1}{2}\Omega_\Lambda}(\tau-\tau_0)\right)$$
$$\Rightarrow a = \cosh\left(\sqrt{\pm\frac{1}{2}\Omega_\Lambda}(\tau-\tau_0)\right), \tag{70}$$

where the integration constant τ_0 can be complex, giving in effect three types of functions: tangent for the trajectory on the lower line, hyperbolic tangent for the A_1A_2 segment, and hyperbolic cotangent for the trajectories on the upper line that escape to $\pm\infty$. The dependence of the scale factor and h on time for these cases is shown in Figure 4. Additionally, the trajectories coming from infinity qualitatively reflect the behavior of the generic trajectories in the respective region in Fig. 3; in particular, the past singularity is reached in finite time.

Outside the singular lines, there are the two special heteroclinic orbits: from A_1 through B_1 to A_2 and from B_4 to A_2. The first is possible, because the equation can be regularized by considering α as a function of h so that $\alpha'(h) = v_1/v_2$, which leads to a local expansion at B_1,

$$\alpha = \frac{9}{2}\Omega_\Lambda - 6h^2 + \mathcal{O}(h^4). \tag{71}$$

This trajectory is similar to the one through B_2 but avoids the problem of singular action. The second case, upon closer inspection, also admits continuation through B_4, as is revealed by switching again to $h_1 \coloneqq h - \sqrt{\Omega_\Lambda}$ as the independent variable. The series for α can then be found:

$$\alpha = -\frac{3}{2}\Omega_\Lambda - 12\sqrt{3}h_1 + \mathcal{O}(h_1^2). \tag{72}$$

Both of these solutions are shown in Fig. 5: the first is probably the best candidate for a "bounce" universe, and the second has a big-bang singularity.

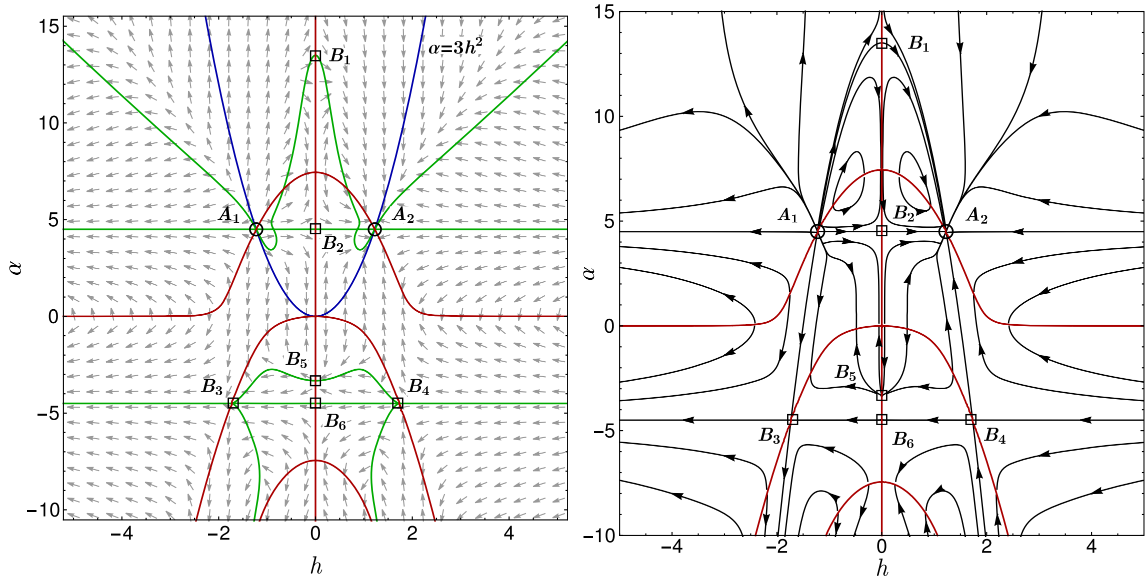

FIG. 3. The vector field (63) and its phase portrait for a flat, empty universe with rational F given by Eq. (69) and with $\Omega_\Lambda = 3$. The field is singular on the red curves. The green curves represent a vanishing numerator of v_1, and the blue parabola corresponds to $\dot{h} = 0$. Because of huge variation, the vector lengths are not drawn to scale, to better show the discontinuity of direction at the singular lines.

Looking more closely at the behavior at infinity also reveals an asymptotic relation of the form $\alpha \sim -6h^2$, which, together with the two previous expansions, suggests looking for the equation of the extended separatrix involving $\alpha + 6h^2$. Indeed, it turns out that there is a parabola through B_3, A_1, B_1, A_2, B_4 given by

$$U \coloneqq \alpha + 6h^2 - \frac{9}{2}\Omega_\Lambda = 0, \tag{73}$$

which is an invariant set, i.e.

$$\left.\frac{\mathrm{d}U}{\mathrm{d}\tau}\right|_{U=0} = 0, \tag{74}$$

as can be checked by direct substitution.

Eliminating α from $U = 0$ leaves a simple Riccati equation $2\dot{h} = 3\Omega_\Lambda - 6h^2$, which again gives trigonometric solutions for h and a akin to Eq. (70)—in particular, for the big-bang type

$$a = \sinh\left(\sqrt{\frac{9}{2}\Omega_\Lambda}(\tau - \tau_0)\right)^{1/3} \sim (\tau - \tau_0)^{1/3}, \tag{75}$$

which is the behavior of the standard Friedmann cosmology with the so-called stiff matter characterized by $p = \rho$. The same equation of state holds also for a minimally coupled massless scalar field ϕ, for which the energy density is just the kinetic term $\rho = \frac{1}{2}\dot{\phi}^2$, or approximately when the potential term can be neglected: $\frac{1}{2}\dot{\phi}^2 \gg V(\phi)$. This suggests a correspondence analogous to that of

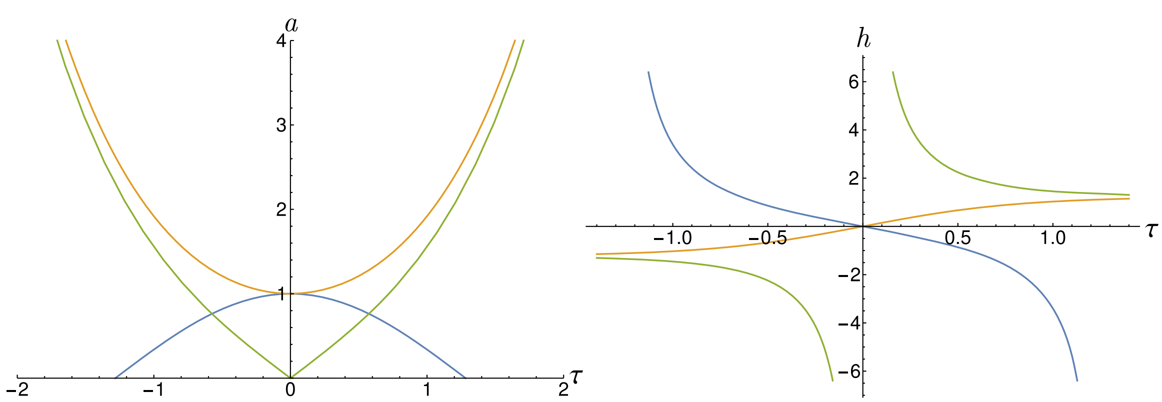

FIG. 4. Behavior of a and h for the singular lines $\alpha = \pm\frac{3}{2}\Omega_\Lambda$ of Fig. 3. The blue trajectory goes through $B_4B_6B_3$, the orange one connects A_1 to A_2, and the green one has A_1 and A_2 as limit points.

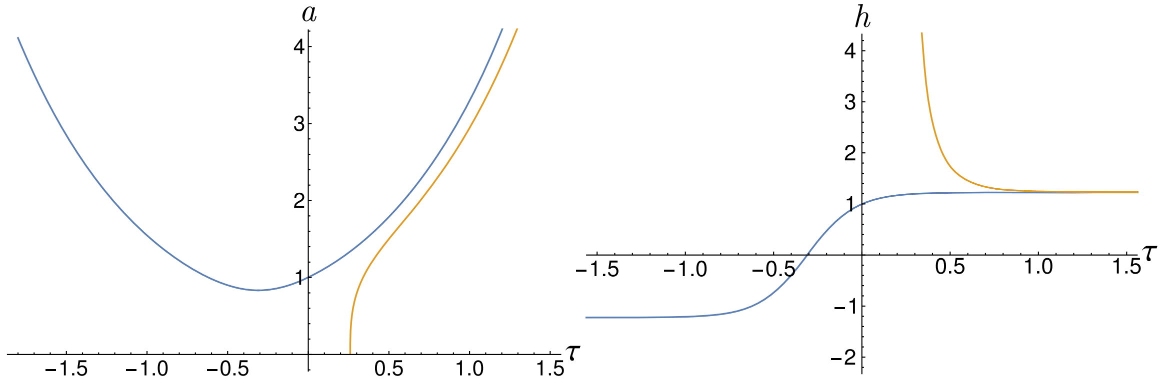

FIG. 5. The nontrivial heteroclinic trajectories of Fig. 3: the blue line corresponds to the one connecting A_1 and A_2 through B_1, and the orange line to the one coming from infinity to A_2 through B_4.

standard R^2 theories, which are conformally equivalent to scalar field cosmologies [11].

The introduction of matter through a nonzero Ω term means that the system (63) can no longer be simply visualized on a plane, but particular solutions can still easily be obtained numerically. The most important ingredient would be dust matter ($\gamma = 1$), and following that, radiation ($\gamma = \frac{4}{3}$), but since the latter constitutes a tiny fraction of Ω in the standard ΛCDM model, $\Omega = \Omega_m a^{-3}$ is assumed in the numerical integration. Thus, this particular model will depart from reality close to the big bang by ignoring the radiation-dominated GUT era and the inflationary phase, when the value of Λ is much larger than the ΛCDM one used below.

A surprising property to notice is that the parabola (73) is still an invariant set, and accordingly Eq. (75) gives a big-bang solution also with dust. This is due to the singular nature of the denominator in the Ω/F'' term in v_1. Although this means that the stiff matter component dominates in the earliest epochs, the "effective equation of state" p/ρ changes with time as the de Sitter state is reached. By analogy with the standard Friedmann cosmology, one can eliminate p from the second Einstein equation to obtain the time-dependent adiabatic index as

$$\gamma_\tau = \frac{2}{3}\left(1 - \frac{a\ddot{a}}{\dot{a}^2}\right) = \frac{2}{3}\left(1 - \frac{\alpha}{3h^2}\right). \tag{76}$$

This function can be used to compare the behavior of the density for the present model and the corresponding Friedmann equation including the stiff matter term, i.e.,

$$h^2 = \Omega_\Lambda + \Omega_m a^{-3} + \Omega_s a^{-6}, \qquad \Omega_\Lambda + \Omega_m + \Omega_s = 1. \tag{77}$$

The comparison is shown in Fig. 6.

In the present case, there is no constraint on the sum of all the Ω terms, and Ω_Λ and Ω_m need not be the same as in the ΛCDM model, because the Einstein equations are different. The parameter values are both subject to estimation from observations, but for the present qualitative comparison, one can use the asymptotic behavior of (75): $a \sim \exp(\sqrt{\Omega_\Lambda/2}\tau)$, which should correspond to the relevant asymptotics of ΛCDM, i.e., $a \sim \exp(\sqrt{0.7}\tau)$, so that $\Omega_\Lambda = 1.4$ is chosen for the $f(R)$ equations.

At any rate, the comparison shows that the universe whose trajectory lies in the first quadrant (Fig. 3) and tends to the de Sitter attractor A_2 has $\gamma = 2/3$ during the big bang (red in Fig. 6), so it corresponds to cosmic strings [12,13]. This is peculiar, because it means that the matter term (a^{-3}) must be canceled close to the initial singularity, so that only the a^{-2} term matters instead. It happens due to the trajectory approaching the horizontal singular line of $\alpha = 3\Omega_\Lambda/2$, so asymptotically the solution (70) holds and $a \sim (\tau - \tau_0)$. The transition from cosmic strings directly to exponential expansion makes this class of trajectories unlikely as physical models.

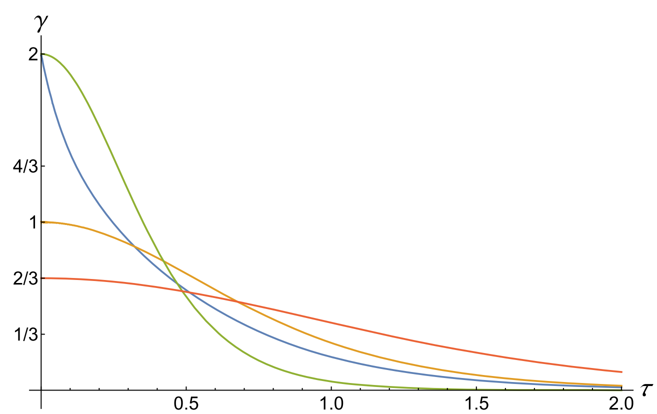

FIG. 6. The time-dependent index of the equation of state $p = (\gamma - 1)\rho$. The orange line corresponds to the ΛCDM model (77) with $\Omega_\Lambda = 0.69$ and $\Omega_m = 0.31$; the blue line is almost the same scenario, but with $\Omega_m = 0.3$ and the addition of $\Omega_s = 0.01$. The green and red lines are big-bang solutions for a flat universe with dust and rational F given by Eq. (69), $\Omega_\Lambda = 1.4$ and $\Omega_m = 0.6$. They both tend to A_2: the former through B_4 and the latter from the right of the first quadrant (see Fig. 3).

The heteroclinic trajectory (green in Fig. 6) is unchanged by dust with $\gamma \approx 2$ stationary at first, then decaying to the "dark energy" level. This decrease is faster than for the corresponding ΛCDM with stiff matter (blue), but the agreement is much better than in the previous scenario. The shape resembles more that of the standard ΛCDM (orange) in that there is no cusp, although different types of matter dominate initially. This in itself is not an obstacle, as it is unlikely that classical GR and dustlike matter determine the initial singularity anyway, and in the bouncing scenarios $a''(0) = 0$, so that γ could even tend to infinity.

An interesting analogy here is that the heteroclinic trajectory is unchanged by the addition of dust, so that it can be thought of as defined purely by the geometry and the function $f(R)$—quite as the cosmological constant can be thought of as a geometric term rather than an actual material component. In both cases, such content-independent gravity only makes sense as a model for the late homogeneous universe, not at smaller scales like black holes. Note also that this particular example (69) was deliberately chosen with a singularity so that it cannot be treated perturbatively. By itself, it may not be a replacement for ΛCDM, but its most prominent feature, the invariant manifold $U = 0$, appears as a guidepost in further generalizations, partly because it effectively reduces the fourth-order Einstein equations to an analogue of the Friedmann equation, which is easily solvable. One goal of future investigations will thus be to find models where such invariant curves exist and are nontrivially perturbed by matter.

Coming back to the general dynamics, an undesirable global feature of dynamics with a singular $F(\xi)$ is that the phase space is cut into several regions by the red lines, and the trajectories cannot be continued through them even with local analysis, because the vector field's directions are opposite on each side. Nevertheless, A_2 is a steady state attractor for almost the whole first quadrant, and there are two heteroclinic scenarios without singularities.

This behavior is more pronounced when one considers more peculiar setups—for example, with the periodic Lagrangian

$$F(\xi) = \Omega_\Lambda\left(-\frac{3}{2} + \tan\left(\frac{\xi}{\Omega_\Lambda}\right)\right). \tag{78}$$

Because F enters the equations with the rescaled eigenvalues α and β as its arguments, it is more convenient to eliminate h and use the eigenvalues as the dependent variables. In order to do that, a rescaled time $\mathrm{d}\sigma \coloneqq \mathrm{d}\tau/h$ can be used, giving for the flat case

$$\begin{cases} \frac{\mathrm{d}\alpha}{\mathrm{d}\sigma} = \frac{6\Omega+W}{3F''(\alpha)+F''(\beta)}, \\ \frac{\mathrm{d}\beta}{\mathrm{d}\sigma} = \frac{6\Omega+W}{9F''(\alpha)+3F''(\beta)} - (\alpha-\beta)(\frac{1}{3}\alpha-\beta). \end{cases} \tag{79}$$

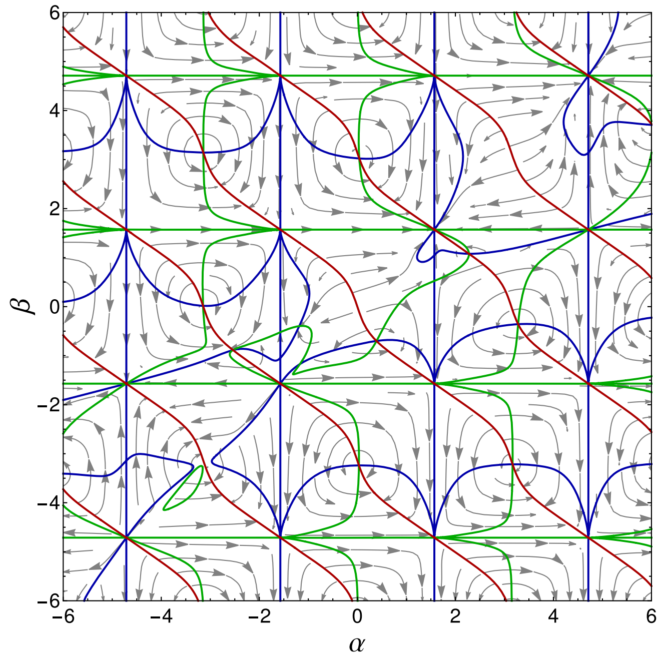

FIG. 7. The vector field and the singular lines for the system of Eq. (79) with trigonometric $F(\xi)$ of (78) and $\Omega_\Lambda = 1$.

This setup gives rise to a period cell structure of the phase space, as seen in Fig. 7, and there are infinitely many critical points and heteroclinic orbits to choose from.

At present, this cannot be considered to be more than a toy model, but it hints at the possibility of constructing a phase space with compartments for different epochs of evolution separated by the singular lines and transitions taking place through the critical points. The behaviors of h and a would need to be recovered from that of α and β in order to give physical interpretation, and at first glance, it is hard to judge whether the complexity comes from the choice of dependent variables, or is an intrinsic feature of the tensor $f(R)$ theory.

The determination of the actual (real, if one can call it that) $F(\xi)$, or f, is a question in itself, and at present it is hard to imagine what other fundamental theory could provide it. At the very least, it should be constrained by observations, but some new approach will be required not to merely fit subsequent polynomial approximations of a series if one wants to recover the complete function.

VI. OBSERVATIONAL FORMULAS

In order to assess the applicability of the proposed construction, one must turn to observational cosmology. The detailed numerical analysis is outside the scope of this article and will be deferred to future work. Nevertheless, some preparatory analysis is straightforward and can be given here.

The standard cosmological test relies on the supernova Ia data and the relationship between the redshift and luminosity. In the Friedmann case, there is a direct relation between H^2 and the redshift, so the integration of time and distance is straightforward. Here, the equations involve up to the third derivative of the scale factor, so another route needs to be taken: for small redshifts, a series formula binding various expansion coefficients can be given, while in the general case, the dynamical system has to be integrated.

Recall first that the redshift is linked to the scale factor by $z+1=a^{-1}$, for $a(0)=1$ at present, and that the luminosity distance to an object at comoving distance r is $d_L=r(1+z)$. Provided, then, that r can be expressed by z, this will allow us to calculate the apparent luminosity and relate to observations [14].

The required expression follows from the condition of the null geodesic: $\mathrm{d}s^2=0$, which for the metric (48) gives directly

$$r=\frac{1}{\sqrt{k}}\sin\left(\sqrt{k}\int\frac{\mathrm{d}t}{a}\right), \tag{80}$$

where a limit is understood for $k=0$. Assuming that a or z are monotonic functions of t that can be used for parametrization of the light path, the above can be rewritten as

$$d_L=\frac{1+z}{H_0\sqrt{\Omega_k}}\sin\left(\sqrt{\Omega_k}\int_0^z\frac{\mathrm{d}z}{h}\right). \tag{81}$$

In the standard model, H is simply given as a function of z by the Friedmann equation, and the integral can even be explicitly calculated by means of elliptic functions [12]. As mentioned above, this cannot be done here, but following Ref. [12], the main equation can be used to give constraints of the higher characteristics—the deceleration parameter q and the jerk j:

$$q\coloneqq-\frac{\ddot{a}a}{\dot{a}^2}=-\frac{\alpha}{3h^2},\qquad j\coloneqq\frac{\ddot{a}a^2}{\dot{a}^3}=\frac{\dot{\alpha}}{3h^3}-q. \tag{82}$$

A change of the independent variable from t (or τ) to z immediately gives

$$\frac{\mathrm{d}h}{\mathrm{d}z}=\frac{1+q}{1+z}h,\qquad\frac{\mathrm{d}q}{\mathrm{d}z}=\frac{j-2q^2-q}{1+z}, \tag{83}$$

which then allows us to expand h in the integral (81) in powers of z, so that the whole expression can be expanded as

$$d_L=\frac{z}{H_0}\left(1+\frac{1-q_0}{2}z-\frac{1+j_0-3q_0^2-q_0+\Omega_k}{6}z^2+\mathcal{O}(z^3)\right). \tag{84}$$

For small z, this provides a means to finding H_0, q_0 and j_0 from the luminosity data, but one also has to take into account that these parameters are not independent. In the standard model, q can be eliminated because $h'(z)$ is an explicit function of z and the density parameters Ω. Similarly here, the jerk is constrained by the main equation, which for this purpose becomes

$$j+q=\frac{6\Omega+W(\alpha,\beta,h,a)}{3(3F''(\alpha)+F''(\beta))h^4}, \tag{85}$$

with

$$\alpha=-3h^2q,\qquad\beta=h^2(2-q)+\Omega_k(1+z)^2. \tag{86}$$

So, given the function F, the constraint on j_0 is

$$j_0+q_0=\frac{6\Omega_0+W(-3q_0,2-q_0+\Omega_k,1,1)}{3(3F''(-3q_0)+F''(2-q_0+\Omega_k))}. \tag{87}$$

Finally, to obtain the luminosity distance for larger redshifts, where a series expansion is not practicable, an augmented dynamical system is a straightforward solution. Assuming again that z can be used as the independent variable, as is the case in exponential expansion, a dynamical equation for d_L is necessary instead of the integral (81).

The null geodesic condition gives

$$\frac{\mathrm{d}r}{\mathrm{d}z}=\frac{\sqrt{1-kr^2}}{H}, \tag{88}$$

and denoting the dimensionless distance by $l=H_0d_L$ leads to

$$\frac{\mathrm{d}l}{\mathrm{d}z}=\frac{l}{1+z}+\frac{\sqrt{(1+z)^2-\Omega_kl^2}}{h}, \tag{89}$$

while the basic system now reads

$$\begin{cases}\frac{\mathrm{d}\alpha}{\mathrm{d}z}=-\frac{6\Omega+W(\alpha,\beta,h,(1+z)^{-1})}{(1+z)(3F''(\alpha)+F''(\beta))h^2},\\ \frac{\mathrm{d}h}{\mathrm{d}z}=\frac{3h^2-\alpha}{3(1+z)h}.\end{cases} \tag{90}$$

Because z has become the independent variable, this system is non-autonomous and only two-dimensional (regardless of k and Ω). Even in the Friedmann case, for more complex $H(z)$, the integral (81) has to be obtained numerically. The only complication here is that three ordinary differential equations need to be integrated; their initial conditions follow from the definitions

$$l(0) = 0, \qquad \alpha(0) = -3q_0, \qquad h(0) = 1. \quad (91)$$

VII. CONCLUSIONS

The main modification of the gravitational action proposed here is to include terms nonlinear in curvature, but going further than polynomials, so that rational functions with a finite radius of convergence or even transcendental functions can be used. Additionally, instead of considering just a function of the Ricci scalar $f(\mathrm{tr}[R])$, the whole tensor can be treated as an argument, and the trace taken at the very end to produce a scalar Lagrangian density $\mathrm{tr}[f(R)]$. In the case of transcendental functions, this considerably changes the results, when compared to the ordinary $f(\mathcal{R})$ theories.

With a view to fully general treatment, such as including spin, the presented derivation is valid for affine connections with nonvanishing torsion and without the assumption of metricity. An important consequence is that for nonsymmetric Ricci tensors, one can no longer introduce an obvious metric conformal to the original g_{ab}. This stems from the nonlinear functions of the Ricci tensor entering the equations, instead of just functions of the Ricci scalar multiplying R_{ab} or g_{ab}.

Despite the difficulties, workable equations can be derived and applied to the Robertson-Walker geometry so that the analogue of the standard cosmological model may be studied. As is generally the case, the modified Einstein equations are of higher order, and instead of one Friedmann equation, one has a three-dimensional dynamical system.

An obvious complication is that the dynamical variables enter the equations both inside and outside the transcendental functions, which leaves little hope for explicit solutions. Nevertheless, these models are within reach, and if the function f is determined from other fundamental principles, the dynamics and observational consequences can still be effectively analyzed, as shown here.

The analysis of phase portraits for both rational and transcendental f reveals critical points which are attractors and which correspond to de Sitter solutions. More importantly, there also exist non-singular "big bounce" evolutions, which are heteroclinic trajectories, and explicit solutions for them can be given. For the dynamical systems to be two-dimensional, it was assumed that the curvature was zero and no ordinary matter was present. On the one hand, this allows for a complete visualization of the phase diagram, but on the other, it limits the physical applicability. Still, the late or present Universe with accelerated expansion can be modeled as the de Sitter attractor, while for the big bounce solutions the scale factor does not approach zero, so that matter density never dominates, and neglecting it is justifiable.

If dustlike matter is included, the separatrix of the above simplified rational model survives and the same explicit solutions hold. One still has both big bounce and big bang solutions, not unlike those of ΛCDM with stiff matter. In general, matter changes the early evolution around the separatrix but not on it. Thus, the next possible step in constructing a viable model seems to be identifying $f(R)$ such that it also has an invariant submanifold, but which depends on Ω_m, not just on the geometry and Λ.

In any case, the elegant feature here is that the cosmological constant can appear naturally because of how the theory is constructed—it is identified with the constant term of $f(R)$. Yet, even when this term was zero ($f = \exp -1$), the same sort of accelerated expansion appeared.

A more detailed study of the initial singularity in the presence of matter and curvature index k could lead to more interesting results still. For example, seeing how one of the scenarios imitates stiff matter, it will be interesting to ask if such cosmologies can be equivalent to standard general relativity with a scalar field, similarly to the ordinary R^2 case. It is also the quadratic $f(R)$ case for the Robertson-Walker geometry, when there is an equivalence with the ordinary $f(\mathcal{R})$, although it does not seem to extend to higher orders. Another convergence is found when the traceless Ricci tensor vanishes, so that R_{ab} is proportional to g_{ab} and the Einstein equations for both theories coincide. However, as the examples show, even for an empty universe this might correspond only to fixed points, not to general solutions of the full theory.

With a view to future work, some observational formulas are also given, so that the basic cosmological tests can be applied. A comparison to the standard model is in order to help guide the subsequent theoretical developments. Specifically, some constraints on the function f should be obtained. The crudest way would be to fit the first coefficients of its expansion, but of course there is no hope in recovering the whole series this way.

Rather, one might want to approach the problem by trying to fit a differential equation satisfied by f. Already for linear differential equations with rational coefficients this would reduce the number of parameters to finite, while at the same time allowing for a the vast family of (confluent) hypergeometric functions and their generalizations.

Future investigations could also address the question of reduction of the order of the dynamical system (63). For the Einstein-Hilbert action, the third derivative of the scale factor does not enter, and only the Friedmann equation, which is a relation between H and a, is left. Here, the equation involving the third derivative of the scale factor, or $\dot{\alpha}$, would be reduced if $3F''(\alpha) + F''(\beta) = 0$. For independent α and β this happens only if F is linear, so that GR is recovered.

If, on the other hand, there is a relation $\beta = \psi(\alpha)$, then a nontrivial solution to the functional equation $3J(\alpha) + J(\psi(\alpha)) = 0$ could potentially be found. Such a relation is in itself a second-order differential equation for the scale factor, so the dynamics is simplified, but it then also means that the function F is determined by $F''(\xi) = J(\xi)$.

Ideally, however, the function f should be mainly constrained by experiment, not just the simplicity of the resulting equations. If this theory passes the basic cosmological tests, analyzing it in a wider context of gravitational physics will help address this issue. Questions of instabilities will have to be answered, although as suggested by Ref. [3], the Palatini approach, applicable here, provides a setting to avoid at least the Ostrogradski instability. In general, issues such as ghost fields, semiclassical stability and post-Newtonian (Solar System) tests will be required, and hopefully undertaken, to ascertain the overall viability of the presented extension.

I would like to thank professor Marek Szydłowski for many valuable discussions of the manuscript.

[1] A. A. Starobinsky, A new type of isotropic cosmological models without singularity, Phys. Lett. B **91B**, 99 (1980).

[2] G. Allemandi, A. Borowiec, and M. Francaviglia, Accelerated cosmological models in Ricci squared gravity, Phys. Rev. D **70**, 103503 (2004).

[3] G. J. Olmo, Introduction to palatini theories of gravity and nonsingular cosmologies, in *Open Questions in Cosmology*, edited by G. J. Olmo (InTech Publishing, 2012), Chap. 7.

[4] S. Nojiri and S. D. Odintsov, Unified cosmic history in modified gravity: From $F(R)$ theory to Lorentz non-invariant models, Phys. Rep. **505**, 59 (2011).

[5] T. P. Satiriou and V. Faraoni, $f(R)$ theories of gravity, Rev. Mod. Phys. **82**, 451 (2010).

[6] A. Borowiec, Nonlinear Lagrangians of the Ricci type, arXiv:gr-qc/9906043.

[7] A. Borowiec, Metric-polynomial structures and gravitational Lagrangians, Inst. Phys. Conf. Ser. **173**, 241 (2002).

[8] N. J. Higham, *Functions of Matrices* (Society for Industrial and Applied Mathematics, Philadelphia, 2008).

[9] A. R. Liddle, How many cosmological parameters?, Mon. Not. R. Astron. Soc. **351**, L49 (2004).

[10] A. Kurek and M. Szydłowski, The ΛCDM model in the lead: A Bayesian cosmological model comparison, Astrophys. J. **675**, 1 (2008).

[11] B. Whitt, Fourth-order gravity as general relativity plus matter, Phys. Lett. **145B**, 176 (1984).

[12] M. P. Dabrowski and T. Stachowiak, Phantom Friedmann cosmologies and higher-order characteristics of expansion, Ann. Phys. (Amsterdam) **321**, 771 (2006).

[13] M. P. Dabrowski, Oscillating Friedman cosmology, Ann. Phys. (N.Y.) **248**, 199 (1996).

[14] S. Weinberg, *Cosmology* (Oxford University Press, New York, 2008).