

**Problems with propagation and time evolution in  $f(T)$  gravity**

Yen Chin Ong\*

*Graduate Institute of Astrophysics, National Taiwan University, Taipei 10617, Taiwan  
and Leung Center for Cosmology and Particle Astrophysics, National Taiwan University, Taipei 10617, Taiwan*

Keisuke Izumi†

*Leung Center for Cosmology and Particle Astrophysics, National Taiwan University, Taipei 10617, Taiwan*

James M. Nester‡

*Department of Physics, National Central University, Chungli 320, Taiwan; Graduate Institute of Astronomy,  
National Central University, Chungli 320, Taiwan; Center for Mathematics and Theoretical Physics,  
National Central University, Chungli 320, Taiwan; and Institute of Physics, Academia Sinica, Taipei 11529, Taiwan*

Pisin Chen§

*Graduate Institute of Astrophysics, National Taiwan University, Taipei 10617, Taiwan; Leung Center for Cosmology  
and Particle Astrophysics, National Taiwan University, Taipei 10617, Taiwan; Department of Physics, National Taiwan University,  
Taipei 10617, Taiwan; and Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory,  
Stanford University, Stanford, California 94305, USA*

(Received 19 April 2013; published 10 July 2013)

Teleparallel theories of gravity have a long history. They include a special case referred to as the teleparallel equivalent of general relativity (TEGR, also known as  $GR_{\parallel}$ ). Recently this theory has been generalized to  $f(T)$  gravity. Tight constraints from observations suggest that  $f(T)$  gravity is not as robust as initially hoped. This might hint at hitherto undiscovered problems at the theoretical level. In this work, we point out that a generic  $f(T)$  theory can be expected to have certain problems including superluminal propagating modes, the presence of which can be revealed by using the characteristic equations that govern the dynamics in  $f(T)$  gravity and/or the Hamiltonian structure of the theory via Dirac constraint analysis. We use several examples from simpler gauge field theories to explain how such superluminal modes could arise. We also point out problems with the Cauchy development of a constant time hypersurface in Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime in  $f(T)$  gravity. The time evolution from a FLRW (and as a special case, Minkowski spacetime) initial condition is not unique.

DOI: [10.1103/PhysRevD.88.024019](https://doi.org/10.1103/PhysRevD.88.024019)

PACS numbers: 04.50.Kd, 95.30.Sf

**I. INTRODUCTION**

Einstein proposed the idea of teleparallelism, or *Fernparallelismus* (also known as absolute parallelism, distant parallelism), in order to unify gravity and electromagnetism [1]. Unlike general relativity, spacetime in teleparallelism is equipped with a connection with zero curvature, but nonvanishing torsion. Since the spacetime is flat, the parallel transport of a vector is independent of path, and hence the name *teleparallel*, meaning “parallel at a distance.” Einstein’s quest for unification via teleparallel theory can be found in the interesting account of [2]. Despite Einstein’s failure to formulate a unified theory, teleparallelism was later revived and studied as a pure theory of gravity. The main motivations were (1) that the teleparallel formulation was considered to have advantages with regard to the identification of the energy-momentum of gravitating systems, and (2) teleparallel theory can be

regarded as a gauge theory of local translations (see, e.g., [3–11]).<sup>1</sup> In fact Einstein’s general theory of relativity (GR) can be rewritten in teleparallel language (see, e.g., [13–16]). The result is a theory equivalent to GR known variously as the teleparallel equivalent of General Relativity (TEGR) or  $GR_{\parallel}$ . Recently TEGR has been generalized to  $f(T)$  theory, a theory of modified gravity formed in the same spirit as generalizing general relativity to  $f(R)$  gravity [17,18].

It seems, however, that observational data puts rather tight constraints on  $f(T)$  gravity, effectively reducing the theory (at least for some choices of  $f(T)$ ) back to TEGR (However, see also [19,20]). This includes constraints from considering a varying fine structure constant [21], the formation of large-scale structures [22] and a varying gravitational constant [23], as well as solar system constraints [24]. This could hint at problems with  $f(T)$  gravity at the theoretical level.

In this paper we will briefly review TEGR and more general theories of teleparallel gravity, including  $f(T)$  gravity, in Sec. II. In Sec. III, we discuss the degrees of freedom

\* ongyenchin@member.ams.org

† izumi@phys.ntu.edu.tw

‡ nester@phy.ncu.edu.tw

§ chen@slac.stanford.edu

<sup>1</sup>For recent study of teleparallelism as a *higher* gauge theory, see [12].

for  $f(T)$  theory. We know that there are extra degrees of freedom in the theory beyond those of GR; that the number of extra degrees of freedom is generically 3 was clearly revealed by the Dirac constraint analysis carried out by Li *et al.* [25]. Despite this, in a careful linear perturbative analysis up to second order on flat Friedmann-Lemaître-Robertson-Walker (FLRW) background, no extra mode appears [26]. In the case of Minkowski background, this had already been pointed out in [25]. Disappearance of extra degrees of freedom is known to happen in some gravity theories (e.g., in massive gravity, see [27]); this can be caused by the accidental disappearance of kinetic terms at the linear level. If so, as we shall argue, it is likely to be a signal of superluminal propagation. We will also discuss the related danger of the nonlinear constraint effect, which plagued some other theories of gravity in the past. Our conclusion is that  $f(T)$  theory is very likely to give rise to unwanted problems due to the effects of nonlinear constraints, in addition to propagation outside of the null cone. As a consequence, the number and type of degrees of freedom could vary as a physical system evolves. This is further explored via the method of characteristics in Sec. IV.

The method of characteristics is properly introduced in Sec. IV, where we also provide an illuminating example in an application to the nonlinear Proca field. The same example is given a Hamiltonian analysis that clearly shows the close connection of superluminal propagating modes and nonlinear constraints. Although the characteristic equations allow one to obtain superluminal propagating degrees of freedom, one must be careful to distinguish between *physical* degrees of freedom and gauge degrees of freedom, as the example of the Maxwell field illustrates.

In Sec. V, we investigate the dynamical equations in  $f(T)$  theory and derive a system of partial differential equations that describe the characteristics of the theory. This calculation reveals nonlinear features that are very prone to allow superluminal propagation outside of the metric null cone.<sup>2</sup> Furthermore, we also show that the constant-time hypersurface in FLRW spacetime, and as a special case Minkowski spacetime, can be a Cauchy horizon, which means that even the infinitesimal future time evolution from it is not unique. We conclude with some discussion in Sec. VI.

## II. TELEPARALLEL THEORIES OF GRAVITATION

Let  $M$  be a semi-Riemannian manifold of signature  $(-, +, +, +)$  with metric  $g$ . The teleparallel theories we wish to discuss<sup>3</sup> use the vierbein field or tetrad  $e_a(x^\mu)$  as

<sup>2</sup>We should remark here that superluminal propagation may or may not be achronal, so it is not certain that such propagation really does lead to a violation of causality. We make further comments on this in Sec. IV.

<sup>3</sup>There are more general teleparallel theories which also have nonmetricity [7]; in fact, there are symmetric teleparallel theories—with vanishing torsion but nonvanishing nonmetricity—including one that is equivalent to GR [28,29].

the fundamental dynamical variables. We use Greek letters for spacetime coordinate indices and Latin letters to label the tetrad vectors.

Suppose that the manifold is parallelizable, i.e., there exist  $n$  vector fields  $\{v_1, \dots, v_n\}$  such that at *any* point  $p \in M$  the tangent vectors  $v_i|_p$ 's provide a basis of the tangent space at  $p$ . Note that this particular set of vector fields must be able to span the tangent spaces at all points of the manifold, i.e., it is a set of everywhere linearly independent globally nonvanishing vector fields. Then we can view the mapping between the bases of the tangent space in a coordinate frame  $\{\partial_\mu\}$  to that of a noncoordinate frame  $\{e_a\}$  as an isomorphism  $TM \rightarrow M \times \mathbb{R}^4$ . Orthonormality of the tetrad is thus imposed by the Minkowski metric on the tangent space  $T_p M \cong \mathbb{R}^{3,1} = (\mathbb{R}^4, \eta)$ , related to the metric on the manifold  $M$  by  $g_{\mu\nu} = \eta_{ab} e^a_\mu e^b_\nu$ .

One then equips the manifold  $M$  with the metric-compatible Weitzenböck connection [30,31]  $\overset{w}{\nabla}$  defined by

$$\overset{w}{\nabla}_X Y := (XY^a) e_a, \quad (1)$$

where  $Y = Y^a e_a$ . Equivalently, the connection coefficients are

$$\overset{w}{\Gamma}{}^\lambda_{\mu\nu} = e_a^\lambda \partial_\nu e^a_\mu = -e^a_\mu \partial_\nu e_a^\lambda. \quad (2)$$

In local coordinates, the components of the torsion tensor are

$$\overset{w}{T}{}^\lambda_{\mu\nu} := \overset{w}{\Gamma}{}^\lambda_{\nu\mu} - \overset{w}{\Gamma}{}^\lambda_{\mu\nu} = e_a^\lambda (\partial_\mu e^a_\nu - \partial_\nu e^a_\mu) \neq 0. \quad (3)$$

The Weitzenböck connection can be shown to be curvature-free.<sup>4</sup>

The Weitzenböck connection differs from the Levi-Civita connection by the contortion tensor<sup>5</sup>

$$K^{\mu\nu}{}_\rho = -\frac{1}{2} (\overset{w}{T}{}^{\mu\nu}{}_\rho - \overset{w}{T}{}^{\nu\mu}{}_\rho - \overset{w}{T}{}^\mu{}_\rho{}^\nu). \quad (4)$$

The Einstein-Hilbert action of general relativity is (in units where  $c = 1$ )

$$S = \frac{1}{2\kappa} \int d^4x \sqrt{-g} R, \quad \kappa = 8\pi G. \quad (5)$$

Up to a total divergence (which does not affect the field equations), this action is equivalent to

<sup>4</sup>There is an alternative formulation for teleparallel theories that has some advantages; it uses the teleparallel connection as an *a priori* independent dynamical variable that is constrained to have vanishing curvature via a Lagrange multiplier [4,5,7,32–34]. Here, to more simply communicate our ideas to the audience we most want to reach, we will use the formulation familiar in the  $f(T)$  works.

<sup>5</sup>We remark that the proper term is contortion instead of the often used “contorsion.” See, e.g., [35].

$$S = -\frac{1}{2\kappa} \int d^4x |e| T, \quad (6)$$

where  $e = \det(e^a{}_\mu)$ ,

$$S_\rho{}^{\mu\nu} := \frac{1}{2} \left( K^{\mu\nu}{}_\rho + \delta_\rho^\mu T^{\nu\alpha}{}_\alpha - \delta_\rho^\nu T^{\alpha\mu}{}_\alpha \right), \quad (7)$$

and  $T$  is a certain quadratic-in-torsion scalar:

$$T := S_\rho{}^{\mu\nu} T^{\rho}{}_{\mu\nu}, \quad (8)$$

the latter is the so-called torsion scalar.

In this teleparallel language, (6) is called the TEGR action.  $f(T)$  gravity simply promotes the ‘‘torsion scalar’’ in the TEGR Lagrangian to a function, i.e., the action of  $f(T)$  gravity is<sup>6</sup>

$$S = \int d^4x |e| f(T). \quad (9)$$

One can then obtain the field equation for  $f(T)$  theory of modified gravity by varying with respect to the frame  $e^a{}_\nu$ , this leads to

$$X_a{}^\nu := \partial_\mu (e f_T S_a{}^{\mu\nu}) - e e_a{}^\lambda T^\rho{}_{\mu\lambda} S_\rho{}^{\nu\mu} f_T - \frac{1}{4} e e_a{}^\nu f = 0. \quad (10)$$

We remark that the (quadratic-in-)torsion scalar defined above is a very special one. We note that mathematically we can define a more general torsion ‘‘scalar’’  $\tilde{T}$  by relaxing the coefficients:<sup>7</sup>

$$\tilde{T} = a T^{\rho}{}_{\eta\mu} T^{\eta\mu}{}_\rho + b T^{\rho}{}_{\mu\eta} T^{\eta\mu}{}_\rho + c T^{\rho}{}_{\rho\mu} T^{\nu\mu}{}_\nu. \quad (11)$$

In particular one can construct the one-parameter combination

$$T[\lambda] = \frac{1 + 2\lambda}{4} T^{\rho}{}_{\eta\mu} T^{\eta\mu}{}_\rho + \frac{1 - 2\lambda}{2} T^{\rho}{}_{\mu\eta} T^{\eta\mu}{}_\rho - T^{\rho}{}_{\rho\mu} T^{\nu\mu}{}_\nu, \quad (12)$$

where  $\lambda \geq 0$ . Years ago a teleparallel theory using this as the Lagrangian was proposed as an alternative to GR. This theory passed all observational (as of, perhaps up to, 1990) [37,38] and some theoretical tests [39]. But then certain theoretical problems were found, which we will discuss in the next section.

We summarize how teleparallel theories work: We start with choosing a good local orthonormal frame  $\{e_a\}$  which is declared to be covariantly constant under parallel

<sup>6</sup>For simplicity here we follow the somewhat nonstandard convention used in [26], in which the factor  $-(2\kappa)^{-1}$  is absorbed into  $f$ , i.e., TEGR corresponds to  $f(T) = -(2\kappa)^{-1}T$ .

<sup>7</sup>In fact one could also include two more quadratic-in-torsion combinations which have odd parity; they can be found in Baekler and Hehl [36].

transport by imposing the Weitzenböck connection. This defines a global frame field that sometimes goes by the name orthoteleparallel frame or OT frame. Recall that there are infinitely many possible choices of frames that span the same tangent space at  $p \in M$ , related to each other by elements of the proper orthochronous Lorentz group. The torsion tensor, and consequently the torsion scalar, can thus be expressed in terms of any one of these infinitely many bases. Once we get a parallelization, it is defined up to global Lorentz transformation. This defines a Weitzenböck geometry. The set of all possible parallelizations of  $M$  is then partitioned into equivalence classes consisting of parallelizations that relate to each other in the same class by a global Lorentz transformation (thus, one cannot pass from one parallelizations to another by a local Lorentz transformation). Note that at this stage, nonzero torsion does not mean that we have some ‘‘distinguished frames’’ or preferred frames; any equivalence class of frames are on the same footing, in the sense that they parallelized the spacetime equally well. However, once we introduce the physics via the action, we get field equations. Now we can only admit certain classes of parallelizations, just as in general relativity where we can only admit metrics that satisfy the Einstein field equations.

We remark that due to the lack of local Lorentz invariance (see also the discussions in [40,41]), unlike in general relativity where we can change coordinate systems and use any frame fields, this is not so in generic teleparallel theories. As a consequence, this implies that we cannot directly extract the tetrad from the metric in the straightforward way. TEGR is special since its action does not determine the admissible OT frame but only the metric. From now onwards, we will suppress all the explicit overscript  $w$ 's on the torsion scalar and connection coefficients etc.

### III. EXTRA DEGREES OF FREEDOM AND THE DANGER OF NONLINEAR CONSTRAINTS

We are reminded of the problem faced by the one-parameter teleparallel theory in which the teleparallel Lagrangian has the form (12). This theory has four extra degrees of freedom [42]. Kopczyński [32] first pointed out that this theory has predictability problems; he argued that one is unable to determine uniquely the evolution of the teleparallel geometry. However, further analysis [43] showed that the problem is not generic, instead it only occurs for a special class of solutions. Cheng *et al.* verified this behavior using the Hamiltonian approach and identified the troubles as a certain effect of nonlinear constraints, there referred to as constraint bifurcation [42], in which the chain of constraints could ‘‘bifurcate’’ depending on the values of the fields. That is to say, the number or type of constraints would depend on the values of the phase space variables. One could imagine that the evolution of some system as depicted by a curve in phase space could pass through regions with a different number or type of constraints.

This would be rather strange (since this means that the number of gauges and physical degrees of freedom also change), and one will find it hard, if at all possible, to predict the evolution of the system. To put it in another way, the matrix with the Poisson brackets of the constraints as entries may not have constant rank, due to the nonlinearity of the constraints. The number or type of constraints may change as one approaches a point in the phase space where the rank changes. The more general Poincaré gauge theory also has similar problems [44–46]. We suspect that this might also be the case for  $f(T)$  theory (and probably for many teleparallel theories). It is of course conceivable that with a specific choice of  $f(T)$ , we could prevent the solutions from ever approaching such problematic regions in phase space, e.g., if they would act as a dynamical repeller.

The number of degrees of freedom of a generic  $f(T)$  theory of gravity (in  $(3 + 1)$ -dimensions) has been found to be five by Miao Li *et al.* [25] using a Hamiltonian approach.<sup>8</sup> That is, there are three extra degrees of freedom, which the authors suggest could correspond to one massive vector field or one massless vector field with one scalar field. To understand the properties of the degrees of freedom in a theory, one should first try to analyze them by a perturbative approach. However, it could happen that not all degrees of freedom show up in the perturbation level.

For the case of  $f(T)$  gravity, as remarked by Li *et al.* at the end of their paper [25], for some special backgrounds (including Minkowski space with  $e^a{}_\mu = \delta^a{}_\mu$ ), some of the extra degrees of freedom do not appear in the *linear* perturbation. It is also found that [26], on the flat FLRW background with a scalar field, linear perturbation up to second order does not reveal any extra degree of freedom. In other words,  $f(T)$  theory is highly nonlinear: the number and type of constraints in the linear theory is different from that of the full, nonlinear theory. This is also one of the problematic features discussed in the context of Poincaré Gauge Theory [44–46]. The worry of such behavior is that linearized modes which are “good” may cease to be so in the full nonlinear theory, perhaps accompanied by anomalous characteristics, as we will further comment on in the later sections.

The revealing Hamiltonian discussion of Li *et al.* for  $f(T)$  gravity in  $(3 + 1)$  dimensional spacetime is rather complicated and some of the results are only inferred implicitly. However they also worked out explicitly the analogous but somewhat simpler Hamiltonian formulation for  $f(T)$  gravity in  $(2 + 1)$ -dimensions; for this case the problem we wish to address is clearly evident. In this case there are six first class constraints<sup>9</sup> ( $H, H_i, \Pi^{a0}$ ), where

<sup>8</sup>His analysis was based on the Hamiltonian formulation of Maluf [47,48]. Additional insight could probably be found using the other more general Hamiltonian approaches to gravity theories which specialize to teleparallel theory [5,33,49,50].

<sup>9</sup>We do not explain notations of Li *et al.* [25] here since we will not need them in detail.

$a = 0, 1, 2$  and  $i = 1, 2$ , as well as four second class constraints ( $\Gamma^1, \Gamma^2, \Gamma^{12}, \pi$ ), which leads to two degrees of freedom.

They define the quantities

$$\begin{aligned} y_i &= \{H_0, \Gamma^i\}, & y_3 &= \{H_0, \Gamma^{12}\}, & x_0 &= \{H_0, \pi\}, \\ x_i &= \{\Gamma^i, \pi\}, & x_3 &= \{\Gamma^i, \pi\}, \end{aligned} \quad (13)$$

and

$$\begin{aligned} A_i &= \{\Gamma^i, \Gamma^{12}\} \\ &\approx 2e[g^{0i}(g^{01}g^{2m} - g^{02}g^{1m}) + g^{1i}(g^{0m}g^{02} - g^{2m}g^{00}) \\ &\quad - g^{i2}(g^{0m}g^{01} - g^{1m}g^{00})]\partial_m \phi. \end{aligned} \quad (14)$$

The self-consistency equation in matrix form is shown to be  $M_{3D}\Lambda_{3D} = 0$ , where  $\Lambda_{3D} = (1, \lambda_1, \lambda_2, \lambda_3, \lambda)^T$  and

$$M_{3D} = \begin{pmatrix} 0 & y_1 & y_2 & y_3 & x_0 \\ -y_1 & 0 & 0 & A_1 & x_1 \\ -y_2 & 0 & 0 & A_2 & x_2 \\ -y_3 & -A_1 & -A_2 & 0 & x_3 \\ -x_0 & -x_1 & -x_2 & -x_3 & 0 \end{pmatrix} \quad (15)$$

satisfies  $\det M_{3D} = 0$ . This matrix generically is of rank 4. However, as previously mentioned, some of the Poisson brackets could vanish in some cases, which would result in rank changes of the matrix as the system evolves. Miao Li *et al.* found the explicit formulas for the Lagrange multipliers:

$$\lambda_1 = \frac{A_2 x_0 + x_3 y_2 - x_2 y_3}{A_1 x_2 - A_2 x_1}, \quad (16)$$

$$\lambda_2 = \frac{-A_1 x_0 - x_3 y_1 + x_1 y_3}{A_1 x_2 - A_2 x_1}, \quad (17)$$

$$\lambda_3 = \frac{y_1 x_2 - y_2 x_1}{A_1 x_2 - A_2 x_1}, \quad (18)$$

$$\lambda = \frac{A_1 y_2 - A_2 y_1}{A_1 x_2 - A_2 x_1}. \quad (19)$$

While they were interested in the most general case and thus only considered the case with  $A_1 x_2 - A_2 x_1 \neq 0$ , this type of special case is precisely what we are interested in this work. Indeed one sees that in principle  $A_1 x_2 - A_2 x_1$  can vanish, yet the numerators of these expressions are not generally constrained to vanish at the same time, which allows one or more of the Lagrange multipliers to become unbounded. As we will explain below, this is signaling a superluminal propagation mode.

#### IV. THE METHOD OF CHARACTERISTICS

A viable theory of gravity should satisfy certain theoretical criteria, including the lack of tachyonic modes, or

modes that carry negative kinetic energy (i.e., ghosts).<sup>10</sup> The theory should also support a well-posed initial value problem, that is, satisfy the Cauchy-Kowalevski theorem [51–53]. Moreover any propagation mode in the theory should also be described by hyperbolic quasi-linear partial differential equations with well-behaved characteristics, i.e., the characteristic surfaces should be nonspacelike. One should also consider a theory with a “good” Minkowski limit as preferable to a theory that does not have such behavior.

The method for studying characteristics is well-known; see for example the detailed classical work of Courant and Hilbert [54], or Chapter 8 of the more recent text of [55]. A discussion of the method of characteristics in the case of general relativity can be found in Lecture 14 of [56]. For more rigorous treatment see [57], as well as in [58]. See also [59] for some useful discussions in the context of external fields in gauge theories. Here we only give a brief summary. Recall that if  $P$  is a linear differential operator of order  $k$ , then we can consider  $P$  as a polynomial in the derivative  $D$ . In multi-index notation we can write this as

$$P = \sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha. \quad (20)$$

We may ask in what directions it is really of order  $k$ . For an ODE  $a_k(x)(d/dx)^k + \dots + a_0(x) = 0$  for example, it is obvious that  $a_k(x) \neq 0$  is the condition required for the equation to be of order  $k$  everywhere. For PDEs with multiple variables, we need to be careful. Consider for example the Laplacian  $\partial^2/\partial x^2 + \partial^2/\partial y^2$  in  $\mathbb{R}^2$ . This is clearly of order 2 in both the  $x$ - and  $y$ -direction, however for an operator which is mixed, e.g.,  $\partial^2/\partial x \partial y$ , it is *not* of order 2 in either the  $x$ - or  $y$ -direction. Nevertheless, it is a second-order operator, which can be revealed by introducing new variables,  $t = x + y$ ,  $s = x - y$ , which renders  $\partial^2/\partial x \partial y = (1/4)(\partial^2/\partial t^2 - \partial^2/\partial s^2)$ . That is, this operator is of order 2 in the  $s$  and  $t$  directions. In general then, given any operator  $P$  at  $x$ , and a direction  $v$ , we can make an orthogonal change of variable so that  $v$  points along one of the new coordinate axes, say the  $x_1$  axis. Now, if the coefficient of the partial derivative  $(\partial/\partial x_1)^k$  (in the new coordinate system) is nonzero at  $x$ , then we can say that  $P$  is of order  $k$  at  $x$  in the direction  $v$ . We refer to such situation as noncharacteristic. That is, characteristic refers to the case in which the coefficient of  $(\partial/\partial x_1)^k$  vanishes at  $x$ .

The total symbol or top-order symbol of  $P$  is simply a polynomial obtained by replacing the derivative  $D$  with a variable, say  $\xi$ . The principal symbol, denoted  $\sigma_P(\xi)$ , is the highest degree component of the total symbol. That is,

$$\sigma_P(\xi) = \sum_{|\alpha|=k} a_\alpha \xi^\alpha. \quad (21)$$

<sup>10</sup>However, one may still argue that theories with such seemingly pathological features may still be acceptable. We will make further comments in Sec. VI.

The principal symbol almost completely determines the qualitative behavior of the solutions of the system. Furthermore, it is well known that for hyperbolic (as well as parabolic) partial differential equations, the zeros of the principal symbol describe the characteristics of the system. That is to say, the characteristic directions are exactly those for which the principal symbol vanishes. Indeed the “standard procedure” is to begin with the equation of motion or the field equation, keep only the highest derivative terms, and then replace the said derivatives  $\partial_\mu \partial_\nu \dots \partial_\kappa$  with components of the normal vector to the characteristics  $k_\mu k_\nu \dots k_\kappa$  and set the equation to zero. This is called the characteristic equation. We can study how the characteristics propagate by looking for what sorts of vectors are allowed as solutions to the characteristic equation.

Note that for a characteristic equation that is a matrix equation, called the characteristic matrix, it suffices to consider the vanishing of the determinant of the characteristic equation (known as the characteristic determinant) instead of the vanishing of the equation itself (the latter is of course a stronger statement).

In this paper, we are interested in one (or a few) special characteristic direction that gives the signal of a superluminal mode. By analyzing the characteristic determinant, we can examine all directions at once.

The characteristic surface which is orthogonal to the characteristic direction coincides with the edge of Cauchy development, i.e., the Cauchy horizon. This is because the higher order derivative term with respect to the characteristic direction disappears. The disappearance of the higher order derivative term causes the evolution to be singular. Thus, the Cauchy development of the characteristic surface is only the characteristic surface itself and the evolution from it is meaningless even if it is spacelike.

Note that for application in theories of gravity, we would like our characteristic directions to be null or spacelike, as a timelike characteristic direction is the signal of superluminal propagation (and could violate causality, although this is not necessarily so even in theories with local Lorentz invariance [60–62]). Correspondingly, there should not be any nontrivial solution that corresponds to any timelike vector.

### A. Example: nonlinear Proca field

The characteristic method is best illustrated via an example or two.

Consider the Lagrangian of the Proca field with a nonlinear term [63],

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{2} m^2 A^\mu A_\mu - \frac{1}{4} \lambda (A^\mu A_\mu)^2, \quad (22)$$

where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  and  $m, \lambda = \text{const}$ . The signature is  $(-, +, +, +)$ . The field equations are

$$\partial_\mu F^{\mu\nu} - m^2 A^\nu - \lambda A^\mu A_\mu A^\nu = 0. \quad (23)$$

Upon taking the divergence of this equation, because the first term vanishes by the antisymmetry of  $F^{\mu\nu}$ , we obtain an implicit constraint satisfied by the system:

$$(m^2 + \lambda A^\mu A_\mu) \partial_\nu A^\nu + 2\lambda A^\mu A^\nu \partial_\nu A_\mu = 0. \quad (24)$$

For  $\lambda = 0$ , this gives  $\partial_\nu A^\nu = 0$ , which upon substituting back into the field equation (23) gives

$$(\partial_\mu \partial^\mu - m^2) A^\nu = 0, \quad (25)$$

which is the Klein-Gordon equation. The characteristic equation for this case is just  $k_\mu k^\mu = 0$ , i.e., the characteristic is null.

To find the characteristics in the general  $\lambda$  case, replace the highest derivative terms  $\partial^N A$  by  $k^N \tilde{A}$ . Here  $\tilde{A}^\mu$  means it is not the value of the vector  $A^\mu$ , but represents the change of the vector in a certain direction.

We then get the relations

$$k \cdot k \tilde{A}^\nu - k^\nu k \cdot \tilde{A} = 0, \quad (26)$$

$$(m^2 + \lambda A \cdot A) k \cdot \tilde{A} + 2\lambda A \cdot k A \cdot \tilde{A} = 0.$$

A linear combination of these gives

$$[(m^2 + \lambda A \cdot A) k \cdot k + 2\lambda (A \cdot k)^2] k \cdot \tilde{A} = 0. \quad (27)$$

Now from Eq. (26) one can see that modes with  $k \cdot \tilde{A} = 0$  propagate with null characteristics, but from Eq. (27) one can see that modes with  $k \cdot \tilde{A} \neq 0$  generally have non-null characteristics. The normal to the characteristic surface could be timelike, indicating a superluminal characteristic. In that case there is a Lorentz frame in which  $k^\mu = (1, 0, 0, 0)$ . In such a frame the condition is

$$m^2 + \lambda(-3A_0^2 + A_i A^i) = 0. \quad (28)$$

We want to emphasize that this tachyonic mode of the nonlinear Proca field can be detected using the Hamiltonian formulation, as we now explain. From the Lagrangian, the canonical momenta are

$$\pi^\mu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu}. \quad (29)$$

That is, explicitly,  $\pi^i = F_0^i$  and  $\pi^0 = 0$ ; the latter is a primary constraint.

The Hamiltonian density, constructed according to the Dirac-Bergmann constraint procedure [64–67], is [44]

$$\begin{aligned} \frac{1}{2} \pi^i \pi_i + \frac{1}{4} F^{ij} F_{ij} - A_0 \partial_c \pi^c + \frac{1}{2} m^2 (A^i A_i - A_0^2) \\ + \frac{1}{4} \lambda (A^i A_i - A_0^2)^2 + u \pi^0, \end{aligned} \quad (30)$$

where the primary constraint  $\pi^0 \approx 0$ <sup>11</sup> has been included with an unknown Lagrange multiplier  $u$ . From the Hamiltonian evolution equation

<sup>11</sup>Here “ $\approx$ ” denotes Dirac’s weak equality, i.e., it only holds on the constraint surface within the phase space.

$$\dot{A}_0(x) = \{A_0(x), H\} = u(x), \quad (31)$$

one finds the meaning of the multiplier, it is the missing “velocity.” Preserving the primary constraint leads to the secondary constraint,

$$\chi := \partial_c \pi^c + m^2 A_0 + \lambda (A^i A_i - A_0^2) A_0 \approx 0. \quad (32)$$

The Poisson bracket of the two constraints is

$$\{\pi^0(x), \chi(y)\} = [m^2 + \lambda (A^i A_i - 3A_0^2)] \delta^3(x - y). \quad (33)$$

Generically this is nonvanishing, so they make up a 2nd class pair.

However, there is an important exception when the rhs vanishes, this is exactly the same as the anomalous characteristic condition (28). The dynamical consequence shows up when we require preservation of the  $\chi$  constraint, which is the relation that determines the “unknown multiplier”  $u$ ,

$$0 \approx \dot{\chi}(x) = \{\chi(x), H\} = u(x) [m^2 + \lambda (A^i A_i - 3A_0^2)] + \mathcal{G}, \quad (34)$$

where  $\mathcal{G}$  denotes a collection of field dependent terms which, generically, are nonvanishing. Consequently, the field “velocity”  $\dot{A}_0$ , given by the Lagrange multiplier,  $u$ , becomes unbounded at any point(s) where  $\Delta := m^2 + \lambda (A^i A_i - 3A_0^2)$  approaches 0. That is,

$$\lim_{\Delta \rightarrow 0} \dot{A}_0 = \lim_{\Delta \rightarrow 0} u(x) = \lim_{\Delta \rightarrow 0} -\frac{\mathcal{G}}{\Delta} = \infty. \quad (35)$$

This is an indication that, with respect to this constant time spacelike hypersurface, there is instantaneous propagation of the  $A_0$  mode. The nonlinear constraint has led to a field-dependent constraint Poisson bracket value, a signal for the occurrence of superluminal propagation, in complete agreement with the previous analysis obtained from the characteristic equation.

## B. Example: scalar field and Maxwell field

One issue that we have to be careful with is the possible presence of gauge degrees of freedom in a theory (which was not present in the previous nonlinear Proca field example).

In order to appreciate the issue, first we shall look at the simple example of scalar fields. We consider the Lagrangian for the single scalar field defined by

$$\mathcal{L}_S = -\frac{1}{2} \partial_\mu \phi_1 \partial^\mu \phi_1. \quad (36)$$

We know that the scalar field  $\phi_1$  propagates in the null direction. In terms of characteristics, we can understand it as follows. The equation of motion is

$$\partial_\mu \partial^\mu \phi_1 = 0. \quad (37)$$

So the characteristic equation for  $\phi_1$  is  $k_\mu k^\mu = 0$ , which means the characteristic is null.

Let us introduce another scalar field  $\phi_2$  which does not appear in the Lagrangian. Trivially, the Lagrangian does not change under the transformation  $\phi_2 \rightarrow \phi_2' = \phi_2 + f(x^\mu)$ . This property is a certain kind of gauge transformation. The equation of motion which comes from the variation of the Lagrangian (36) with respect to  $\phi_2$  is trivial, i.e.,  $0 = 0$ . Thus the characteristic equation for  $\phi_1$  and  $\phi_2$  becomes  $k_\mu k^\mu \times 0 = 0$ , which is trivially satisfied for any  $k_\mu$ . The triviality comes from the gauge mode  $\phi_2$ . Since  $\phi_2$  does not appear in the Lagrangian, we should not include it in the discussion of characteristics.

The situation for any gauge field theory is similar. For simplicity, we consider the Maxwell theory of the electromagnetic field; the Lagrangian is

$$\mathcal{L}_{\text{EM}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad (38)$$

where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ . The equation of motion is obtained to be

$$\partial_\mu F^{\mu\nu} = 0. \quad (39)$$

The principal symbol yields the characteristic equation

$$k_\mu k^\mu \tilde{A}^\nu - k_\mu k^\nu \tilde{A}^\mu = 0, \quad (40)$$

where we want to seek some nontrivial  $\tilde{A}^\lambda$ . If  $k^\mu$  is null then  $k_\mu \tilde{A}^\mu = 0$  gives a propagating mode on the null cone. However, if  $k^\mu$  is not null, then the characteristic equation can be satisfied by *any*  $\tilde{A}^\lambda = Ck^\lambda$  for arbitrary constant  $C$ . In particular  $k^\mu$  can be chosen to be timelike, which means that the characteristic direction will become timelike. However, this superluminal mode, as we shall see, actually corresponds to gauge mode and so it is not physical.

Alternatively we can consider the determinant method. If we write down the characteristic equation without removing the gauge degree of freedom, we will obtain characteristic matrix that satisfies

$$\det[k_\mu k^\mu g_{\alpha\beta} - k_\alpha k_\beta] = 0. \quad (41)$$

The left-hand side of this equation is 0 for any  $k_\mu$ , and so this equation becomes trivial. The origin of the triviality is the gauge degree of freedom. If we operate  $\partial_\nu$  to the left-hand of Eq. (39), it becomes algebraically 0. Therefore, Eq. (39) has the information of only three independent equations. We know the Maxwell field has  $U(1)$ -gauge, i.e., the Lagrangian is invariant under the  $U(1)$ -gauge transformation  $A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \Lambda$ . Since the gauge degree of freedom  $\Lambda$  does not appear in the Lagrangian, the variation of the Lagrangian with respect to the gauge degree of freedom  $\Lambda$  must be trivial, which is the origin of the above trivial equation. According to the above simple example of scalar fields, we should eliminate the contribution from gauge degrees of freedom in the

discussion of characteristics. Subtracting the contribution from the gauge degree of freedom, i.e., considering the  $3 \times 3$  matrix whose basis can be anything independent of  $k_\mu$ , we can obtain the exact characteristic equation, which is satisfied only when  $k_\mu$  is null. A more rigorous treatment of the removal of gauge degrees of freedom and the characteristics of Maxwell equations can be found in [58].

## V. CHARACTERISTICS EQUATIONS AND PROBLEMATIC TIME EVOLUTION IN $f(T)$ GRAVITY

In this section, we discuss the Cauchy problem for  $f(T)$  gravity by an analysis of the characteristics. First, we shall determine the characteristic equations of  $f(T)$  gravity. After that, we will show that a constant-time hypersurface of the FLRW metric does not provide a good initial condition and that the time evolution of the FLRW metric is not well behaved. Finally, we concretely show the problematic solution where the time evolution from a FLRW initial condition is not unique.

### A. Characteristic equation of $f(T)$ gravity

From Eq. (10), we see that the vacuum field equation in  $f(T)$  gravity is explicitly given by

$$0 = [\partial_\mu (e e_a^\rho S_\rho^{\mu\nu}) - e e_a^\lambda T^\rho{}_{\mu\lambda} S_\rho{}^{\nu\mu}] f_T - \frac{1}{4} e e_a{}^\nu f + e e_a{}^\rho S_\rho{}^{\mu\nu} (\partial_\mu T) f_{TT}. \quad (42)$$

This can be rewritten as

$$0 = f_T (\partial_\mu e) S_a{}^{\mu\nu} + 2e \left[ f_T \frac{\partial S_a{}^{\mu\nu}}{\partial T^b{}_{\alpha\beta}} + S_a{}^{\mu\nu} f_{TT} \left( S_b{}^{\alpha\beta} + T^c{}_{\kappa\rho} \frac{\partial S_c{}^{\kappa\rho}}{\partial T^b{}_{\alpha\beta}} \right) \right] \partial_\mu \partial_\alpha e^\beta{}_\beta - e e_a{}^\lambda T^\rho{}_{\mu\lambda} S_\rho{}^{\nu\mu} f_T - e \frac{1}{4} e_a{}^\nu f. \quad (43)$$

The highest-order derivative terms are just

$$2e \left[ f_T \frac{\partial S_a{}^{\mu\nu}}{\partial T^b{}_{\alpha\beta}} + S_a{}^{\mu\nu} f_{TT} \left( S_b{}^{\alpha\beta} + T^c{}_{\kappa\rho} \frac{\partial S_c{}^{\kappa\rho}}{\partial T^b{}_{\alpha\beta}} \right) \right] \partial_\mu \partial_\alpha e^\beta{}_\beta. \quad (44)$$

We recall that  $S_a{}^{\mu\nu}$  is linear in  $T^b{}_{\alpha\beta}$ , so we may write

$$S_a{}^{\mu\nu} = M_a{}^{\mu\nu}{}^{\alpha\beta} T^b{}_{\alpha\beta} \quad (45)$$

for some  $M_a{}^{\mu\nu}{}^{\alpha\beta}$  which is algebraic in the tetrad  $e^c{}_\rho$ . The explicit form of  $M_a{}^{\mu\nu}{}^{\alpha\beta}$  is

$$M_a{}^{\mu\nu}{}^{\alpha\beta} = \frac{1}{4} \eta_{ab} g^{\mu[\alpha} g^{\beta]\nu} - \frac{1}{2} e_b{}^{[\mu} g^{\nu][\alpha} e_a{}^{\beta]}, \quad (46)$$

where the bracket  $[\dots]$  means antisymmetrization, i.e.,

$$t^{a\dots[bc]\dots} := \frac{1}{2}(t^{a\dots bc\dots} - t^{a\dots cb\dots}). \quad (47)$$

Consequently we have

$$T^c{}_{\kappa\rho} \frac{\partial S_c{}^{\kappa\rho}}{\partial T^b{}_{\alpha\beta}} = T^c{}_{\kappa\rho} M_c{}^{\kappa\rho}{}_{\alpha\beta} = S_b{}^{\alpha\beta}. \quad (48)$$

Thus, we can rewrite Eq. (44) in a more simplified form as

$$2e[f_T M_a{}^{\mu\nu}{}_{\alpha\beta} + 2f_{TT} S_a{}^{\mu\nu} S_b{}^{\alpha\beta}] \partial_\mu \partial_\alpha e^b{}_\beta. \quad (49)$$

Following the standard procedure we can replace  $\partial_\nu \partial_\alpha e^b{}_\beta \rightarrow k_\nu k_\alpha \tilde{e}^b{}_\beta$  in Eq. (49) to obtain the characteristic equations for  $f(T)$  theory of gravity,

$$[f_T M_a{}^{\mu\nu}{}_{\alpha\beta} + 2f_{TT} S_a{}^{\mu\nu} S_b{}^{\alpha\beta}] k_\mu k_\alpha \tilde{e}^b{}_\beta = 0, \quad (50)$$

where

$$M_a{}^{\mu\nu}{}_{\alpha\beta} = \frac{\partial S_a{}^{\mu\nu}}{\partial T^b{}_{\alpha\beta}}. \quad (51)$$

Here the notation  $\tilde{e}^b{}_\beta$  indicates that these are not the values of the frame, but rather represent the change of the frame in a certain direction. We note that Eq. (50) thus provide 16 quadratic polynomial conditions on the four  $k_\mu$ 's.

That is, the characteristic matrix becomes

$$\mathcal{M}_a{}^{\nu\beta} \equiv 2e[f_T M_a{}^{\mu\nu}{}_{\alpha\beta} + 2f_{TT} S_a{}^{\mu\nu} S_b{}^{\alpha\beta}] k_\mu k_\alpha. \quad (52)$$

Since the tetrad  $e^a{}_\mu$  always has the inverse  $e_a{}^\mu$ , the property of the matrix  $\mathcal{M}_a{}^{\nu\beta}$  is the same as that of

$$\begin{aligned} \mathcal{M}^{\lambda\nu\gamma\beta} &\equiv \mathcal{M}_a{}^{\nu\beta} e^{a\lambda} e^{b\gamma} \\ &= 2e[f_T M^{\lambda\mu\nu\gamma\alpha\beta} + 2f_{TT} S^{\lambda\mu\nu} S^{\gamma\alpha\beta}] k_\mu k_\alpha. \end{aligned} \quad (53)$$

Therefore for convenience, we shall analyze the matrix in Eq. (53) instead.

Before discussing the property of the characteristic matrix of Eq. (53), we must first find out the gauge degrees of freedom. We have 16 equations of motion (10), but not all of them are independent due to gauge degrees of freedom. We actually have four nontrivial identities.

Indeed, let us consider the invariance of the action under an infinitesimal coordinate transformation<sup>12</sup>:

$$x^\mu \rightarrow x'^\mu = x^\mu + \xi^\mu(x). \quad (54)$$

This leads to the transformation of the tetrad  $e^a{}_\mu$  by

$$\delta e^a{}_\mu = \partial_\mu (-e^a{}_\alpha \xi^\alpha) - T^a{}_{\mu\nu} e_b{}^\nu (-e^b{}_\alpha \xi^\alpha), \quad (55)$$

and we know this is related to a gauge mode, i.e.,  $-e^a{}_\alpha \xi^\alpha = F^a$ , where  $F^a$  are four scalar fields with an internal index. Varying the action with respect to the component

$$e^a{}_\mu{}^{(\text{gauge})} = \partial_\mu F^a - T^a{}_{\mu\nu} e_b{}^\nu F^b, \quad (56)$$

one obtains the nontrivial Noether differential identities [68]

$$\partial_\mu (X_a{}^\mu) + e_a{}^\nu T^b{}_{\mu\nu} X_b{}^\mu = 0, \quad (57)$$

where  $X_a{}^\mu$  is defined in Eq. (10). We stress that Eq. (57) is an identity, i.e., it is algebraically satisfied without using the equation of motion Eq. (10).

Since the infinitesimal coordinate transformation also leads to a transformation of metric, the metric has the information about all the gauge degrees of freedom. Therefore, all the degrees of freedom of the gauge modes can be eliminated by gauge fixing of the metric components.

Now, we move on to the discussion of the characteristic determinant which is constructed by the  $12 \times 12$  part of the characteristic matrix in Eq. (53) having removed the gauge degrees of freedom. Here, we analyze the characteristic matrix  $\mathcal{M}^{\lambda\nu\gamma\beta}$  by the decomposition into symmetric and antisymmetric parts with respect to  $\nu$  and  $\lambda$  and also with respect to  $\beta$  and  $\gamma$ . Without loss of generality we can choose an appropriate basis to write down the  $12 \times 12$  matrix such that the left half  $12 \times 6$  submatrix represents the symmetric part and the right half  $12 \times 6$  submatrix represents the antisymmetric part.

First, we consider the first term of the characteristic matrix in Eq. (53). The first term of the symmetrized characteristic matrix is well behaved because it is the only contribution inTEGR. The symmetric part of the characteristic equation is related to the metric component, which *a priori* has ten degrees of freedom. However we know that 4 of the degrees of freedom are gauge modes and we have only 6 physical modes. We claim that its rank is generically 6, and indeed after removing the gauge degrees of freedom the second order derivatives are associated with all components of the symmetrized tetrad. We consider next the antisymmetrized component with respect to  $\nu$  and  $\lambda$ . The first term then gives zero contribution, which means that the rank of the first term is indeed 6 for general  $k_\mu$ .

To summarize, the characteristic matrix of  $f(T)$  gravity, after removing the gauge degrees of freedom, can be written as a  $12 \times 12$  matrix, with four  $6 \times 6$  block submatrices. The generic rank of the upper left block is 6, corresponding to the symmetrized tetrads.

Next, we consider the antisymmetrized components of the second term in Eq. (53). The form of the second term is the product of two  $S^{\lambda\nu\mu} k_\mu$  and its rank is one (since one of the eigenvalue vanishes). One can thus form a  $2 \times 2$  matrix at the very center of the  $12 \times 12$  matrix, in which each element belongs to one of the four blocks, respectively. Together with the  $6 \times 6$  block corresponding to the symmetrized tetrad, we thus get a  $7 \times 7$  matrix.

Therefore we have found that the ranks of the first term and the second term of  $\mathcal{M}^{\lambda\nu\gamma\beta}$  are 6 and 1, respectively. By the fact that  $S^{[\lambda\nu]\mu} k_\mu$  is generically nonzero, we can

<sup>12</sup>Geometrically this represents an infinitesimal diffeomorphism.

prove that the rank of matrix  $\mathcal{M}^{\lambda\nu\gamma\beta}$  is generically 7. This rests on the fact that the determinant of the  $7 \times 7$  matrix is generically nonzero, which follows from elementary linear algebra (see the Appendix).

It is important to note that the  $2 \times 2$  matrix at the center, being of rank 1, corresponds to only one of the extra degrees of freedom in  $f(T)$  gravity. However, we know that the theory contains three extra degrees of freedom, so where are the remaining two?

In fact the five antisymmetrized components corresponding to the tetrad  $e^{(i)a}_\nu$  ( $i = 1, \dots, 5$ ), which satisfy

$$S^{[\lambda\nu]\mu} e_{a\lambda} \partial_\mu e^{(i)a}_\nu = 0, \quad (58)$$

become zero. This means that in the equations for the five tetrads  $e^{(i)a}_\nu$  the second order derivative does not appear and that the characteristic equation should be at most a first order differential equation. In other words, the remaining two simply do not show up in the characteristic matrix.

Thus, in order to obtain the complete characteristic we must derive the characteristic equation for  $e^{(i)a}_\nu$  separately. This is what we will do next for the sake of completeness, although it is not a crucial part for our subsequent analysis (since it is sufficient to show that the one degree of freedom that does show up in the characteristic matrix, leads to problematic behavior).

In order to derive the characteristic equation for  $e^{(i)a}_\nu$ , we must go back to the equation of motion Eq. (10), which can be written as

$$\partial_\mu (e S_a^{\mu\nu} f_T) - \frac{1}{4} e e_a^\nu f - e e_a^\lambda T^\rho_{\mu\lambda} S_\rho^{\nu\mu} f_T = 0. \quad (59)$$

If we consider equation of motion from the variation of the action with respect to  $e^{(i)a}_\nu$ , we can easily show that the first term in Eq. (59) does not appear. The equation of motion (59) is obtained from variation of the action

$$\begin{aligned} \delta S &= \int d^4x \delta e^a_\nu \left[ \partial_\mu (e S_a^{\mu\nu} f_T) - \frac{1}{4} e e_a^\nu f \right. \\ &\quad \left. - e e_a^\lambda T^\rho_{\mu\lambda} S_\rho^{\nu\mu} f_T \right] \\ &= - \int d^4x \left[ (\partial_\mu \delta e^a_\nu) e S_a^{\mu\nu} f_T \right. \\ &\quad \left. + \delta e^a_\nu \left( \frac{1}{4} e e_a^\nu f + e e_a^\lambda T^\rho_{\mu\lambda} S_\rho^{\nu\mu} f_T \right) \right], \quad (60) \end{aligned}$$

and for  $\delta e^a_\nu = \delta e^{(i)a}_\nu$  the first term of the right-hand side becomes zero due to Eq. (58). The second term in Eq. (59) is also zero because of antisymmetrization. Eventually, only the last term in Eq. (59) gives a nonzero contribution.

Although the last term in Eq. (59) seems to be a higher-order equation of the first order derivative of  $e^{(i)a}_\nu$  (which, if true, will render the characteristic method unworkable), we can show that it is actually linear. Indeed, we can multiply it by  $-e^{a\alpha}/e$  for convenience. We then have the antisymmetrized equation with respect to  $\alpha$  and  $\nu$

$$T^{\rho\alpha} \overline{\mu S_\rho^{\nu\mu}} f_T = 0, \quad (61)$$

where

$$T^{\alpha\cdots\beta\cdots\gamma\cdots} := \frac{1}{2} (T^{\alpha\cdots\beta\cdots\gamma\cdots} - T^{\alpha\cdots\gamma\cdots\beta\cdots}). \quad (62)$$

This gives the two branches of the equation,  $f_T = 0$  and

$$T^{\rho\alpha} \overline{\mu S_\rho^{\nu\mu}} = 0. \quad (63)$$

If we choose the first branch, i.e.,  $f_T = 0$ , the five equations become degenerate, and thus, time evolution cannot be fixed uniquely. This causes trouble regarding the well-posedness of the Cauchy problem. Therefore, we choose the second branch.<sup>13</sup> The second order terms of the first derivative in Eq. (63) can be written as

$$g^{\mu\beta} \left( -\frac{1}{2} g^{\gamma[\alpha} e_a^{\nu]} e_b^{\rho} + g^{\gamma[\alpha} e_b^{\nu]} e_a^{\rho} \right) \partial_{[\beta} e_{\rho]}^a \partial_{[\mu} e^b_{\gamma]} = 0. \quad (64)$$

If both derivative factors are variations of  $e^{(i)a}_\nu$ , Eq. (64) becomes trivial. Thus at least one of the derivative term is not such a variation. This means that Eq. (63) is not second order but first order with respect to  $\partial_\mu e^{(i)a}_\nu$  and the characteristic equation for  $e^{(i)a}_\nu$  can be shown to be

$$\begin{aligned} \left( \frac{1}{4} g^{\beta[\alpha} \overline{T^{\nu]\mu\lambda} + g^{\alpha[\beta} \overline{T^{\lambda]\mu\nu}} + \frac{1}{2} T^{[\alpha\nu][\beta} g^{\lambda]\mu} \right. \\ \left. + \frac{1}{2} T^\mu g^{\alpha[\lambda} g^{\beta]\nu} + \frac{1}{2} T^{[\beta} g^{\lambda][\nu} g^{\alpha]\mu} \right) e_{a\beta} k_\mu \dot{e}^{(i)a}_\lambda = 0, \quad (65) \end{aligned}$$

where the notation  $\dot{e}^{(i)a}_\lambda$  is similar to  $\tilde{e}^a_\nu$ ; i.e., it also represents the change of the frame in a certain direction. Equation (65) gives the characteristic matrix for  $e^{(i)a}_\nu$ ,

$$\begin{aligned} \mathcal{M}_{(\text{anti})}^{\alpha\nu\beta\lambda} &= \left( \frac{1}{4} g^{\beta[\alpha} \overline{T^{\nu]\mu\lambda} + g^{\alpha[\beta} \overline{T^{\lambda]\mu\nu}} + \frac{1}{2} T^{[\alpha\nu][\beta} g^{\lambda]\mu} \right. \\ &\quad \left. + \frac{1}{2} T^\mu g^{\alpha[\lambda} g^{\beta]\nu} + \frac{1}{2} T^{[\beta} g^{\lambda][\nu} g^{\alpha]\mu} \right) k_\mu. \quad (66) \end{aligned}$$

Then, we have the exact characteristic matrix which is a combination of the  $7 \times 7$  matrix  $\mathcal{M}^{\alpha\nu\beta\lambda}$  and the  $5 \times 5$  matrix  $\mathcal{M}_{(\text{anti})}^{\alpha\nu\beta\lambda}$ .<sup>14</sup>

To summarize, the remaining two extra degrees of freedom do not show up in the characteristic matrix, since their

<sup>13</sup>There is however *a priori* neither a physical principle nor a mathematical requirement that requires us to choose the second branch. This might be another issue of  $f(T)$  gravity.

<sup>14</sup>The cross terms of the five components  $e^{(i)a}_\nu$  and the other seven do not contribute to the characteristic matrix, since their components have only one derivative. Since the nonzero components of the  $7 \times 7$  matrix  $\mathcal{M}^{\alpha\nu\beta\lambda}$  have two derivatives, the contribution of the cross terms with any smaller number of derivatives should not appear in the characteristic equation.

corresponding characteristic equation only contains one derivative term. They can nevertheless be described by another  $5 \times 5$  characteristic matrix  $\mathcal{M}_{(\text{anti})}^{\alpha\nu\beta\lambda}$ . We remark that despite the single derivative term here, this *does not* necessary mean that the dynamics of the two extra degrees of freedom is governed by first order equations. This may be analogous to the Dirac equation, which is first order and linear. However, together with its conjugate equation, one can derive a second-order Klein-Gordon equation governing the dynamics of the system. Likewise, Maxwell's equations are first-order but combining them gives the second order wave equation.

In the next section we will give concrete examples to demonstrate problematic behaviors that could arise in  $f(T)$  gravity due to bad characteristics. To appreciate what is likely to go wrong with the characteristic, let us look at the characteristic equation Eq. (50) again. We see that the first term is good since it is the only term present in TEGR, and we know that characteristics are null in general relativity. In fact the first term gives dynamics to the 6 metrical components of  $\tilde{e}^b{}_\beta$ . The second term quadratic in the field strength  $S^a{}_{\mu\nu}k^\mu$  governing *one* of the extra degrees of freedom however is almost certainly going to give a disaster if  $S^a{}_{\mu\nu}k^\mu = 0$  for timelike  $k^\mu$ . It suggests that with some field values the characteristics could leak outside the metric null cone, which in turn means that there might be problems with tachyonic modes. This nonlinear behavior is reminiscent of the nonlinearity we encountered in the generalized Proca field example.

Note that the case for TEGR is very different since the theory is locally Lorentz invariant, which allows us to choose a gauge that simplifies a lot of the calculation. In any case, TEGR corresponds to  $f_{TT} = 0$  for any value of  $T$  and there is no problem with the characteristics.

### B. Absence of Cauchy development from a constant-time hypersurface in FLRW geometry

In this section, we will show that in  $f(T)$  gravity the time direction, i.e.,  $(\partial/\partial t)^\mu$ , in the FLRW metric is the characteristic direction, and thus the time evolution of the FLRW metric is not unique even in the infinitesimal future. We analyze the characteristic equation in  $f(T)$  gravity which has been derived in the previous subsection. The characteristic determinant can be decomposed into two parts, the  $7 \times 7$  matrix  $\mathcal{M}^{\alpha\nu\beta\lambda}$  and the  $5 \times 5$  matrix  $\mathcal{M}_{(\text{anti})}^{\alpha\nu\beta\lambda}$ . In our analysis we look into the  $7 \times 7$  matrix  $\mathcal{M}^{\alpha\nu\beta\lambda}$ , which has the information about all of the six symmetric components and one of the antisymmetric components. The antisymmetric component appears only in the second term of the characteristic matrix in Eq. (53), and thus for some  $k^\mu$ , the determinant becomes zero if the value of its antisymmetric component is zero. This means that  $k^\mu$  is one of the characteristic directions. To see this, it is enough to solve  $S_{[\lambda\nu]\mu}k^\mu = 0$  for  $k^\mu$ . In the case of the FLRW metric,  $k^\mu = (\partial/\partial t)^\mu$  is a solution.

Let us see this explicitly. We consider the cotetrads

$$e^0{}_\mu dx^\mu = dt, \quad e^a{}_\mu dx^\mu = a(t)\delta_i^a dx^i. \quad (67)$$

These cotetrads give the FLRW geometry with flat spatial section:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a^2(t)\delta_{ij}dx^i dx^j. \quad (68)$$

Given the cotetrads, we can calculate the torsion tensor  $T^\lambda{}_{\mu\nu}$  and  $S^{\lambda\mu\nu}$ . They are

$$T^i{}_{tj} = H\delta_j^i, \quad (69)$$

$$S^{ijt} = -a^{-2}H\delta^{ij}, \quad (70)$$

while the other components all vanish. Here  $H$  is the Hubble parameter, i.e.,  $H = (da/dt)/a$ . Then, it is trivial that  $k^\mu = (\partial/\partial t)^\mu$  is a solution of  $S^{[\lambda\nu]\mu}k_\mu = 0$ .

### C. Concrete example of nonuniqueness of time evolution in FLRW spacetime

We have shown that the FLRW solution has a spacelike characteristic hypersurface, and thus we cannot properly discuss the time evolution from a constant time hypersurface even into the infinitesimal future. Here, we show a concrete solution where the time evolution from the FLRW initial condition without torsion is not unique in  $f(T)$  gravity with or without matter.

We introduce a matter action

$$S_m = \int d^4x \sqrt{-g} \mathcal{L}_m. \quad (71)$$

Then, equation of motion (10) is modified as

$$X_a{}^\nu = \frac{1}{4} \frac{\delta S_m}{\delta e^a{}_\nu}. \quad (72)$$

Assume that the matter action is locally Lorentz invariant, i.e., it depends on the tetrad only through the metric. Then Eq. (72) can be rewritten as

$$X_\mu{}^\nu = \frac{1}{2} g_{\mu\alpha} \frac{\delta S_m}{\delta g_{\alpha\nu}}. \quad (73)$$

Let us now consider the ansatz

$$\begin{aligned} e^0{}_\mu dx^\mu &= \cosh \theta(t) dt + a(t) \sinh \theta(t) dx, \\ e^1{}_\mu dx^\mu &= \sinh \theta(t) dt + a(t) \cosh \theta(t) dx, \\ e^2{}_\mu dx^\mu &= a(t) dy, \quad e^3{}_\mu dx^\mu = a(t) dz, \end{aligned} \quad (74)$$

which also gives the FLRW metric (68). For  $\theta = 0$ , this ansatz is equal to Eq. (67). We note that this is a good choice of tetrad despite the lack of local Lorentz invariance in the theory. In particular the vector fields are everywhere smooth and linearly independent and thus provide a parallelization of FLRW spacetime. This choice of tetrad gives a parallelization that belongs to a different equivalence class of parallelizations than the usual diagonal tetrad of FLRW spacetime (see

the penultimate paragraph of Sec. II), but it is nevertheless a solution to the field equation, as we will show below. Note that of course with the choice  $a(t) \equiv 1$ , our analysis below is in particular applicable to Minkowski spacetime.

The torsion tensor  $T^\rho{}_{\mu\nu}$  and  $S_\rho{}^{\mu\nu}$  are given by

$$\begin{aligned} T^t{}_{tx} &= -T^t{}_{xt} = a(t)\partial_t\theta(t), & T^x{}_{tx} &= -T^x{}_{xt} = H, \\ T^p{}_{tq} &= -T^p{}_{qt} = H\delta^p_q, & S_p{}^{xq} &= -S_p{}^{qx} = \frac{1}{2a(t)}\partial_t\theta(t)\delta^q_p, \\ S_x{}^{tx} &= -S_x{}^{xt} = H, & S_p{}^{tq} &= -S_p{}^{qt} = H\delta^q_p, \end{aligned} \quad (75)$$

where the other components are zero. Here  $p, q \in \{y, z\}$ . From this one readily finds that<sup>15</sup>

$$T = 6H^2. \quad (76)$$

Substituting the above frame  $e^a{}_\mu$ , metric, torsion tensor, torsion scalar and  $S_\rho{}^{\mu\nu}$  into the equation of motion (73), we can find that  $\theta(t)$  does not appear in Eq. (73). This means that if ansatz (67) (i.e.,  $\theta = 0$ ) originally satisfies Eq. (73), for any form of  $\theta(t)$  the ansatz (74) is indeed an exact solution of  $f(T)$  gravity.

Now, we can choose the form of the function  $\theta(t)$  such that  $\theta(t) = 0$  for  $t < 0$  and  $\theta(t) \neq 0$  for  $t > 0$ . In this solution, on the spacelike hypersurface  $t = 0$  torsion suddenly emerges seemingly from nothingness,<sup>16</sup> and the value of  $\theta(t) \neq 0$  for  $t > 0$  is arbitrary. This means that—even in the infinitesimal future—time evolution is not uniquely determined by the initial data, in other words the Cauchy problem is not well defined.<sup>17,18</sup>

Note that this problem does not arise in the case of TEGR due to local Lorentz invariance of the theory.

<sup>15</sup>We remind the readers that we work in  $(-, +, +, +)$  convention. The torsion scalar is  $T = -6H^2$  if one works in the often used convention  $(+, -, -, -)$  in the literature of  $f(T)$  cosmology.

<sup>16</sup>This is probably due to superluminal propagating modes that come in from infinity. One can compare this to a similar phenomenon in anti-de Sitter (AdS) spacetime. AdS spacetime is not globally hyperbolic since although its (conformal) boundary is infinitely far away as measured along any spacelike path, massless particles can propagate to the boundary and back in finite time. In particular this means that, outside the Cauchy development, the initial data does not uniquely determine the time evolution, due to information that can flow in from infinity. In our case however, the Cauchy problem is ill-defined not because of a peculiar geometry of the spacetime, but because of physical superluminal propagating degrees of freedom. In fact it is much worse than the AdS case since one can still at least have Cauchy development locally in AdS.

<sup>17</sup>This example of an unpredictable solution has some similarities with an example that demonstrated unpredictability in another teleparallel theory, which appeared in the seminal work of Kopczyński [32].

<sup>18</sup>Let us also remark that one could consider an initial tetrad which is like the one in this example only in a small region on the initial spacelike hypersurface. Such a tetrad makes an even nastier example.

## VI. DISCUSSION

To conclude, in this work we have derived the set of partial differential equations that govern the characteristics of  $f(T)$  theory of modified gravity, from which we see that there will generally be a real danger of superluminal propagation that leaks outside of the metric null cone unless  $f_{TT} = 0$  for any value of  $T$ , which is the case of TEGR.

We also took a closer look at the degrees of freedom in the theory. We find that  $f(T)$  gravity is a highly nonlinear theory in which many things could go wrong. First of all, as commented by Li *et al.* the linearized theory has a different number of degrees of freedom than the full theory, which has three extra degrees of freedom compared to TEGR (and thus GR). Such nonlinear behavior was further confirmed by careful linear perturbation on a flat FLRW background [26], in which none of the three extra degrees of freedom appear even in the second order linear perturbation. In this work we note that the characteristic equation of  $f(T)$  gravity, i.e., Eq. (50), contains a highly nonlinear second term.

Indeed, we see that generically it excites the six pieces of  $\tilde{e}^b{}_\beta$ , essentially  $e^{[\alpha\beta]}$ .<sup>19</sup> Of course, a more careful analysis by Li *et al.* showed that there are only three extra modes instead of six. Regardless, we see that the antisymmetric modes are excited in a highly nonlinear fashion that is very prone to causing troubles. In fact, these extra modes are certain to have a very singular weak field limit, since the linearized theory has only the metric degrees of freedom. The number of degrees of freedom and the number of constraints thus seem to be likely to depend on the amplitude of the field. As the tetrad fields approach some specific values, some of the Poisson brackets are approaching zero, which in turn means that their corresponding “velocity” Lagrange multipliers tend to become unbounded, which signals the occurrence of instantaneous, i.e., faster-than-light propagation, much like the situation we have exhibited for the nonlinear Proca field (See also the similar discussion in [44]). Moreover these extra dynamic parts are likely to propagate outside the metric null cone and are also likely to allow for negative energy propagating waves. By further analyzing the characteristic matrix of  $f(T)$  gravity, we have shown that there is indeed a physical superluminal mode that could arise from one of the three extra degrees of freedom. Closely related to this issue, we have also demonstrated that the Cauchy problem is ill posed in flat FLRW spacetime and even in Minkowski spacetime. We feel that this is a very bad property for any theory of gravity.

This echoes the remark (Lesson 4) in [69] that extra degrees of freedom in modified gravities are prone to give rise to complications that are very hard to control.

<sup>19</sup>In more general theories with tetrads being the fundamental variables, not only  $e^{[\alpha\beta]}$  but also  $e^0{}_\beta$  can be dynamical.

Although that remark was made under the context of  $f(R)$  gravity, it apparently also holds true for  $f(T)$  gravity. We recall that  $f(T)$  gravity was often thought of as being more well-behaved and easier to deal with compared to  $f(R)$  gravity, since its equation of motion is of second order instead of fourth order. However it now seems that this advantage comes with many tradeoffs. For example, unlike its  $f(R)$  counterpart,  $f(T)$  gravity is not locally Lorentz invariant, and thus it is much more difficult to impose an ansatz for a solution *a priori* except in the simplest situations, such as for a flat spacetime. Furthermore, as we pointed out in this work, many problems that do not arise in  $f(R)$  theory could arise in  $f(T)$  gravity due to extra degrees of freedom and nonlinearity of the constraints.

We recall that in [26], it is pointed out that the extra degrees of freedom do not appear at the linear perturbation level on a FLRW background, and that this behavior is similar to that of nonlinear massive gravity [70] in which, while nonlinear analysis shows that there are generically five gravitational degrees of freedom [71], in the second order action on open FLRW background there are only two tensor degrees of freedom [27,72]. Since the hidden degrees of freedom in fact cause nonlinear instability in the case of nonlinear massive gravity [73], it is a concern that  $f(T)$  theory could exhibit a similar pathology. Here we would like to point out another similarity between nonlinear massive gravity and  $f(T)$  gravity: they both probably exhibit superluminal propagation modes. The case for nonlinear massive gravity was studied in [74], in which the authors investigated the effect of the helicity-0 mode of the theory and found that energy can probably be emitted superluminously on the self-accelerating background. In the case of the ghost-free Wess-Zumino massive gravity [75], Deser and Waldron recently showed that it also exhibits superluminal propagation modes [76].

Here we should also make some comments on the possibility of negative energy solutions in  $f(T)$  gravity. In [3] it was shown, directly in terms of the teleparallel variables, that TEGR only permits positive energy. In general relativity, this is a well-known theorem proved initially by Schoen and Yau [77,78] and subsequently by Witten [79]. In the TEGR formulation the argument was that one can arrange for the Hamiltonian to be dominated by positive terms. A key step in the argument was using the local frame gauge freedom to remove or at least control the nonpositive terms in the TEGR Hamiltonian. But in  $f(T)$  gravity (and teleparallel theory in general) one does not have local frame gauge freedom. Therefore there is no way to kill or control some of the nonpositive terms.

In the positive energy test developed in [80], it is argued that while it is hard to prove that the energy for some theory is positive for every solution to the initial value constraints, it is not so hard to show that a negative energy solution exists. Indeed it is sufficient to show that the initial value constraints admit a nontrivial zero energy solution. If so the

theory should be discarded. In principle this is a very strong test of a theory (i.e., it could exclude many theories which pass other tests). Unfortunately in practice it was not so easy to find even one bad solution explicitly, so it is generally hard work getting results beyond those of the linearized theory. In any case, to us it seems likely that  $f(T)$  gravity would be very vulnerable to having a nontrivial, nonpositive energy solution for most choices of  $f$ , although it is not expected to be very easy to show this.

Although our work is purely classical, we would also like to make a brief comment regarding strong coupling problem once we consider quantizing the theory: Naively, the lack of extra degrees of freedom at the linear level is caused by accidental disappearance of certain kinetic terms at the linear level. Vanishing of kinetic terms at the linear level corresponds to the small limit of their coefficients, which in turn means the nonlinear coupling terms of the canonically normalized modes are becoming very strong. Consequently the perturbative approach cannot be applied and the theory is out of control at the quantum level. Then there is no reason to trust the classical effective action any more [81]. This is expected to be a problem if one attempts to naively quantize  $f(T)$  gravity. It is also worth mentioning that existence of ghosts in a classical theory of gravity may be acceptable if the ghost can be pushed to the Planck scale, to be dealt with by quantum theory of gravity [82].

Finally, it is worth emphasizing that due to the fact that  $f(T)$  theory is highly nonlinear, a full rigorous analysis is at best a very difficult task. One could have a viable theory if one manages to find a particular choice of  $f$  that could avoid the problems raised in this work. In other words, these seemingly serious “problems” could in fact be a blessing in disguise, since they may provide a guide for narrowing down the viable forms of  $f$ .

We conclude with the following remark: we agree that nonlinear theories merit investigation, but, as far as we can see,  $f(T)$  gravity does not seem to have the right kind of nonlinearity; its nonlinearity may be too simple. As commented in [44], it is possible that this could be an indication that an even more nonlinear theory is required. Recall that the linearized spin-2 theory of gravity has problems that are only cured by nonlinearities of the full theory of general relativity (see, e.g., Route 5 in Box 17.2 in MTW [83], or Feynman’s Lectures on Gravitation [84]). Furthermore, in Fierz-Pauli massive gravity [85], the Vaishtein mechanism [86] (introduced to avoid the vDVZ [87,88] discontinuity) excites the Boulware-Deser ghost [89] due to nonlinearity. However, by introducing additional nonlinearity, this ghost mode can be killed (this is the so-called “nonlinear massive gravity) [90–93]. Not surprisingly, this newly introduced nonlinearity seems to give rise to other problems [76]. While inherently nonlinear theories certainly merit serious consideration, finding one that is free of problems, in particular of the kind we have discussed, seems not so easy.

**ACKNOWLEDGMENTS**

Y.C.O. would like to thank Brett McInnes and Wu-Hsing Huang for fruitful discussions about various aspects of differential geometry. Y.C.O. and J.M.N. would also like to thank Friedrich Hehl for providing comments on the early version of the draft. K.I. is supported by Taiwan National Science Council (TNSC) under Project No. NSC101-2811-M-002-103. P.C. is supported by TNSC under Project No. NSC 97-2112-M-002-026-MY3, by Taiwan's National Center for Theoretical Sciences (NCTS), and by the U.S. Department of Energy under Contract No. DE-AC03-76SF00515. J.M.N. is also supported by TNSC under Projects No. 100-2119-M-008-018 and No. 101-2112-M-008-006 and in part by NCTS. Y.C.O. is supported by the Taiwan Scholarship from Taiwan's Ministry of Education.

**APPENDIX**

Let us write  $M = (M)_{ij}$  to denote a matrix  $M$ , while  $M_{ij}$  denotes the matrix element at the  $i$ th row and  $j$ th column of  $M$ .

*Lemma 1.* Let  $M = (M)_{ij}$  be an  $n \times n$  matrix with non-vanishing determinant. Let  $A = (A)_{ij}$  be a  $2 \times 2$  matrix with vanishing determinant. Let  $N = (N)_{ij}$  be the  $(n + 1) \times (n + 1)$  matrix constructed from  $M$  and  $A$ , such that

- (1)  $N_{nn} = M_{nn} + a_{11}$ ,
- (2)  $N_{n(n+1)} = a_{12}$ ,
- (3)  $N_{(n+1)n} = a_{21}$ ,
- (4)  $N_{(n+1)(n+1)} = a_{22}$ ,
- (5)  $N_{(n+1)i} = 0, i \in \{1, 2, \dots, n\}$ ,
- (6)  $N_{i(n+1)} = 0, i \in \{1, 2, \dots, n\}$ ,
- (7)  $N_{ij} = M_{ij}$ , otherwise.

Then,  $\det N = a_{22} \det M$ .

*Proof.* The determinant of  $M$  can be calculated straightforwardly by Laplace expansion

$$\det N = a_{22} \det \left[ \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{11} \end{pmatrix} + (M)_{nn} \right] - a_{21} \det [(M)_{n(n-1)} | r]_{n \times n}, \tag{A1}$$

where  $r = (0 \cdots 0 a_{12})^T$ .

The first term is

$$a_{22} [(a_{11} + M_{nn}) \det (M)_{(n-1) \times (n-1)} + \sum(\cdot)], \tag{A2}$$

where  $\sum(\cdot)$  is the remaining terms in the Laplace expansion. The second term is

$$-a_{21} \det [(M)_{n(n-1)} | r]_{n \times n} = -a_{21} a_{12} \det (M)_{(n-1) \times (n-1)}. \tag{A3}$$

Vanishing of determinant of  $A$  means that  $a_{21} a_{12} = a_{11} a_{22}$ . Therefore we see that Eq. (A1) simply gives

$$\det N = a_{22} [M_{nn} \det [(M)_{(n-1) \times (n-1)}] + \sum(\cdot)] = a_{22} \det [M]. \tag{A4}$$

□

It immediately follows that if  $a_{22} \neq 0$ , then  $\det N \neq 0$  since  $\det M \neq 0$ . This establishes the claim in Sec. VA that the determinant for the  $7 \times 7$  matrix is generically nonzero.

---

[1] A. Einstein, *Berichte Preuss. Akad. Wiss.*, Berlin **17**, 217 (1928); A. Einstein, *ibid.* 224 (1928).  
 [2] T. Sauer, *Historia mathematica* **33**, 399 (2006).  
 [3] J. M. Nester, *Int. J. Mod. Phys. A* **04**, 1755 (1989).  
 [4] M. Blagojevic, *Gravitation and Gauge Symmetries* (IoP Publishing, Bristol, England, 2002).  
 [5] M. Blagojevic and M. Vasilic, *Classical Quantum Gravity* **17**, 3785 (2000).  
 [6] F. W. Hehl, [arXiv:1204.3672](https://arxiv.org/abs/1204.3672).  
 [7] F. W. Hehl, J. D. McCrea, E. W. Mielke, and Y. Ne'eman, *Phys. Rep.* **258**, 1 (1995).  
 [8] Y. Itin, *Int. J. Mod. Phys. D* **10**, 547 (2001).  
 [9] Y. M. Cho, *Phys. Rev. D* **14**, 2521 (1976).  
 [10] C. Moller, *Ann. Phys. (N.Y.)* **12**, 118 (1961).  
 [11] V. C. de Andrade, L. C. T. Guillen, and J. G. Pereira, *Phys. Rev. Lett.* **84**, 4533 (2000).  
 [12] J. C. Baez and D. K. Wise, [arXiv:1204.4339v2](https://arxiv.org/abs/1204.4339v2).  
 [13] K. Hayashi and T. Shirafuji, *Phys. Rev. D* **19**, 3524 (1979).  
 [14] H. I. Arcos and J. G. Pereira, *Int. J. Mod. Phys. D* **13**, 2193 (2004).  
 [15] H. Kleinert, *Electron. J. Theor. Phys.* **24**, 287 (2010).  
 [16] L. L. So and J. M. Nester, in *The Tenth Marcel Grossmann Meeting: On Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories*, edited by M. Novello, S. P. Bergliaffa, R. Ruffini (World Scientific, Hackensack, NJ, 2005), p. 1498.  
 [17] G. R. Bengochea and R. Ferraro, *Phys. Rev. D* **79**, 124019 (2009).  
 [18] E. V. Linder, *Phys. Rev. D* **81**, 127301 (2010).  
 [19] V. F. Cardone, N. Radicella, and S. Camera, [arXiv:1204.5294v1](https://arxiv.org/abs/1204.5294v1).  
 [20] A. Aviles, A. Bravetti, S. Capozziello, and O. Luongo, [arXiv:1302.4871v1](https://arxiv.org/abs/1302.4871v1).  
 [21] H. Wei, X. P. Ma, and H. Y. Qi, *Phys. Lett. B* **703**, 74 (2011).

- [22] B. Li, T.P. Sotiriou, and J.D. Barrow, *Phys. Rev. D* **83**, 104017 (2011).
- [23] H. Wei, H. Y. Qi, and X. P. Ma, *Eur. Phys. J. C* **72**, 2117 (2012).
- [24] L. Iorio and E. N. Saridakis, *Mon. Not. R. Astron. Soc.* **427**, 1555 (2012).
- [25] M. Li, R. X. Miao, and Y. G. Miao, *J. High Energy Phys.* **07** (2011) 108.
- [26] K. Izumi and Y.C. Ong, *J. Cosmol. Astropart. Phys.* **06** (2013) 029.
- [27] A.E. Gumrukcuoglu, C. Lin, and S. Mukohyama, *J. Cosmol. Astropart. Phys.* **11** (2011) 030.
- [28] J. M. Nester and H.-J. Yo, *Chin. J. Phys.* **37**, 113 (1999).
- [29] M. Adak, M. Kalay, and O. Sert, *Int. J. Mod. Phys. D* **15**, 619 (2006).
- [30] R. Weitzenböck, *Invariantentheorie* (Noordhoff, Groningen, 1923).
- [31] R. Weitzenböck, *Differentialinvarianten in der Einsteinschen Theorie des Fernparallelismus*, Preussische Akademie der Wissenschaften, Physikalisch-mathematische Klasse, Sitzungsberichte, (1928), p. 466.
- [32] W. Koczyński, *J. Phys. A* **15**, 493 (1982).
- [33] M. Blagojevic and I. A. Nikolic, *Phys. Rev. D* **62**, 024021 (2000).
- [34] Y. N. Obukhov and J. G. Pereira, *Phys. Rev. D* **67**, 044016 (2003).
- [35] F. W. Hehl, P. von der Heyde, G. D. Kerlick, and J. M. Nester, *Rev. Mod. Phys.* **48**, 393 (1976).
- [36] P. Baekler and F. W. Hehl, *Classical Quantum Gravity* **28**, 215017 (2011).
- [37] M. Schweizer, N. Straumann, and A. Wipf, *Gen. Relativ. Gravit.* **12**, 951 (1980).
- [38] J. Nitsch and F. W. Hehl, *Phys. Lett.* **90B**, 98 (1980).
- [39] R. Kuhfuss and J. Nitsch, *Gen. Relativ. Gravit.* **18**, 1207 (1986).
- [40] B. Li, T.P. Sotiriou, and J.D. Barrow, *Phys. Rev. D* **83**, 064035 (2011).
- [41] T.P. Sotiriou, B. Li, and J.D. Barrow, *Phys. Rev. D* **83**, 104030 (2011).
- [42] W.-H. Cheng, D.-C. Chern, and J. M. Nester, *Phys. Rev. D* **38**, 2656 (1988).
- [43] J. M. Nester, *Classical Quantum Gravity* **5**, 1003 (1988).
- [44] H. Chen, J. M. Nester, and H.-J. Yo, *Acta Phys. Pol. B* **29**, 961 (1998).
- [45] H.-J. Yo and J. M. Nester, *Int. J. Mod. Phys. D* **08**, 459 (1999).
- [46] H.-J. Yo and J. M. Nester, *Int. J. Mod. Phys. D* **11**, 747 (2002).
- [47] J. W. Maluf and J. F. da Rocha-Neto, [arXiv:gr-qc/0002059](https://arxiv.org/abs/gr-qc/0002059).
- [48] J. W. Maluf and J. F. da Rocha-Neto, *Phys. Rev. D* **64**, 084014 (2001).
- [49] C.-M. Chen and J. M. Nester, *Classical Quantum Gravity* **16**, 1279 (1999).
- [50] J. M. Nester, *Mod. Phys. Lett. A* **06**, 2655 (1991).
- [51] A. Cauchy, Reprinted in *Oeuvres Completes*, 1 serie, Tome VII, 17 (1842).
- [52] S. Kowalewski, *Journal für die reine und angewandte Mathematik* **80**, 1 (1875).
- [53] A. M. Nakhushev, in *Encyclopaedia of Mathematics*, edited by M. Hazewinkel (Springer, New York, 2001).
- [54] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1962), p. 590.
- [55] R. Strichartz, *A Guide to Distribution Theory and Fourier Transforms* (World Scientific, Singapore, 2008).
- [56] H. A. Buchdahl, *Seventeen Simple Lectures on General Relativity Theory* (Wiley, New York, 1981).
- [57] F. G. Friedlander, *The Wave Equation on a Curved Space-Time*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2010).
- [58] D. Christodoulou, *Mathematical Problems of General Relativity I* (European Mathematical Society Publishing House, Zürich, 2008).
- [59] D. Zwanziger, *Lect. Notes Phys.* **73**, 143 (1978).
- [60] J.-P. Bruneton, *Phys. Rev. D* **75**, 085013 (2007).
- [61] N. Afshordi, D. J. H. Chung, and G. Geshnizjani, *Phys. Rev. D* **75**, 083513 (2007).
- [62] R. Geroch, [arXiv:1005.1614v1](https://arxiv.org/abs/1005.1614v1).
- [63] G. Velo and D. Zwanziger, *Phys. Rev.* **188**, 2218 (1969).
- [64] P. A. M. Dirac, *Can. J. Math.* **2**, 129 (1950).
- [65] P. A. M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School of Science Monographs Series Vol. 2 (Belfer Graduate School of Science, New York, 1964).
- [66] J. L. Anderson and P. G. Bergman, *Phys. Rev.* **83**, 1018 (1951).
- [67] A. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Systems* (Accademia Nazionale Dei Lincei, Rome, 1976).
- [68] E. Noether, *Gott. Nachr.* **1918**, 235 (1918); For an English translation see M. A. Tavel, *Transp. Theory Stat. Phys.* **1**, 186 (1971).
- [69] T. P. Sotiriou, *J. Phys. Conf. Ser.* **189**, 012039 (2009).
- [70] C. de Rham, G. Gabadadze, and A. J. Tolley, *Phys. Rev. Lett.* **106**, 231101 (2011).
- [71] S. F. Hassan and R. A. Rosen, *Phys. Rev. Lett.* **108**, 041101 (2012).
- [72] A. E. Gumrukcuoglu, C. S. Lin, and S. Mukohyama, *Phys. Lett. B* **717**, 295 (2012).
- [73] A. De Felice, A. E. Gumrukcuoglu, and S. Mukohyama, *Phys. Rev. Lett.* **109**, 171101 (2012).
- [74] C.-I. Chiang, K. Izumi, and P. Chen, *J. Cosmol. Astropart. Phys.* **12** (2012) 025.
- [75] B. Zumino, *Effective Lagrangians and Broken Symmetries*, Brandeis Univ. Lectures on Elementary Particles and Quantum Field Theory (MIT Press, Cambridge, MA, 1970).
- [76] S. Deser and A. Waldron, [arXiv:1212.5835v2](https://arxiv.org/abs/1212.5835v2).
- [77] R. Schoen and S.-T. Yau, *Commun. Math. Phys.* **65**, 45 (1979).
- [78] R. Schoen and S.-T. Yau, *Commun. Math. Phys.* **79**, 231 (1981).
- [79] E. Witten, *Commun. Math. Phys.* **80**, 381 (1981).
- [80] H.-H. Chen, D.-C. Chern, and J. M. Nester, *Chin. J. Phys.* **25**, 481 (1987).
- [81] N. Arkani-Hamed, H. Georgi, and M. D. Schwartz, *Ann. Phys. (N.Y.)* **305**, 96 (2003).
- [82] P. Chen, *Mod. Phys. Lett. A* **25**, 2795 (2010).
- [83] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

- [84] R. P. Feynman, F. B. Morinigo, and W. G. Wagner, *Eur. J. Phys.* **24**, 330 (2003).
- [85] M. Fierz and W. Pauli, *Proc. R. Soc. A* **173**, 211 (1939).
- [86] A. I. Vainshtein, *Phys. Lett.* **39B**, 393 (1972).
- [87] H. van Dam and M. J. G. Veltman, *Nucl. Phys.* **B22**, 397 (1970).
- [88] V. I. Zakharov, *JETP Lett.* **12**, 312 (1970).
- [89] D. Boulware and S. Deser, *Phys. Rev. D* **6**, 3368 (1972).
- [90] C. de Rham and G. Gabadadze, *Phys. Rev. D* **82**, 044020 (2010).
- [91] C. de Rham, G. Gabadadze, and A. J. Tolley, *Phys. Rev. Lett.* **106**, 231101 (2011).
- [92] C. de Rham, G. Gabadadze, and A. J. Tolley, *Phys. Lett. B* **711**, 190 (2012).
- [93] C. de Rham, G. Gabadadze, and A. J. Tolley, *J. High Energy Phys.* **11** (2011) 093.