Modern Physics Letters A
Vol. 30, Nos. 3 & 4 (2015) 1540006 (15 pages)

DOI: 10.1142/S0217732315400064

Problems of massive gravities

S. Deser
Walter Burke Institute for Theoretical Physics, Caltech, Pasadena, CA 91125, USA
and
Physics Department, Brandeis University, Waltham, MA 02454, USA
deser@brandeis.edu

K. Izumi
Leung Center for Cosmology and Particle Astrophysics,
National Taiwan University, Taipei 10617, Taiwan
izumi@phys.ntu.edu.tw

Y. C. Ong
Nordita, KTH Royal Institute of Technology and Stockholm University,
Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden
yenchin.ong@nordita.org

A. Waldron
Department of Mathematics, University of California, Davis, CA 95616, USA
wally@math.ucdavis.edu

Received 8 October 2014
Accepted 10 November 2014
Published 20 January 2015

The method of characteristics is a key tool for studying consistency of equations of motion; it allows issues such as predictability, maximal propagation speed, superluminality, unitarity and acausality to be addressed without requiring explicit solutions. We review this method and its application to massive gravity (mGR) theories to show the limitations of these models' physical viability: Among their problems are loss of unique evolution, superluminal signals, matter coupling inconsistencies and micro-acausality (propagation of signals around local closed time-like curves (CTCs)/closed causal curves (CCCs)). We extend previous no-go results to the entire three-parameter range of mGR theories. It is also argued that bimetric models suffer a similar fate.

Keywords: Massive gravity; characteristics; acausality; superluminality; ghost modes.

1. Introduction

Among the many recent attempts to generalize/modify General Relativity (GR), those that break its general covariance and make its range finite, by adding a mass term to the Einstein action, are perhaps the most active current contenders. These massive gravity (mGR) models are themselves of two types: fiducial massive gravity (fmGR) proceeds by embedding the dynamical "Einstein" system in a fixed spacetime background. Instead, the so-called bimetric (f-g or biGR) models also make the second metric dynamical; the system now has a single, shared, diffeomorphism invariance. While both directions were introduced long ago,[1,2] their recent revival is due to the (partial) resolution of an earlier, fatal, flaw:[3] mGR models generically propagate a sixth, ghost-like, mode in contrast with the five physical degrees of freedom (DoF) of their linearized, Fierz–Pauli (FP)[4] counterparts. This terminated interest in bimetric and fmGR models for four decades. The resolution of this no-go impasse was based first on effective field theory reasoning that used a decoupling limit (large Planck mass, small graviton mass-squared, and a constant product of the two) to study distinguished mass terms depending on $(g_{\mu\nu}, \bar{g}_{\mu\nu})$, identifying just three "ghost-free", combinations[5] (see also Refs. 6 and 7). Later, this conjecture was confirmed (at least — overoptimistically as we shall see — assuming invertibility of certain constraints) by Arnowitt–Deser–Misner (ADM) $(3+1)$ Hamiltonian techniques.[8] Indeed, we shall show here that in the process of removing the original ghost, new — and equally fatal — flaws arise as an unavoidable part of the "cure", a result that restores its unique place to Einstein's coordinate invariant long range GR, just as the uniqueness of massless Yang–Mills theories was forced (for other — quantum — reasons).

At their core, mGR models are described by a coupled set of quasilinear partial differential equations (PDEs). These can be studied in detail using the method of characteristics, which allows the models' predictability, maximal propagation speed, superluminality and acausality to be addressed, both for explicit solutions and in full generality. Since the method determines whether initial data can be propagated into the future, it also allows quantum consistency (recall the relationship between quantum field commutators and propagators[9]) to be addressed; these are methods that have been employed since the very earliest studies of massive higher spin systems.[10,11] Therefore, we begin in Sec. 2 with a review of the characteristic method.

Armed with this mathematical technology, we shall tackle in detail the fmGR models. These are defined in Sec. 3. In order to study their characteristics, it is extremely propitious to have explicit covariant expressions for their constraints. The first analysis of this type was given in Ref. 12 where it was proven — covariantly — that (for at least one of the three mass terms) the model generically propagates five DoF. Explicit expressions for these were given in Ref. 13 and extended to a two-dimensional subspace of fmGR parameter space in Ref. 14. The remaining, unprobed, third parameter choice remained problematic for quite some time because the dynamical Weyl tensor appeared in its would-be scalar constraint — the one responsible for ghost-removal. The covariant proof that the last scalar constraint extended to the full parameter space was only very recently provided in Ref. 15.

That paper gives a definitive analysis of fmGR's constraint structure; it is briefly summarized in Sec. 3.

Characteristics for fmGR theories were first studied in Ref. 13 where superluminalities were discovered. (Earlier work also found superluminality in the models' decoupling limit.[16]) The models' full characteristic matrix was then computed, at least for a subspace of parameter space in Refs. 17 and 18; the analysis of Ref. 18 then even uncovered acausalities — closed time-like curves (CTCs). The characteristic matrix for the full fmGR parameter space was finally computed in Ref. 15. Those results are summarized in Sec. 4, where we also use them to give two explicit acausality examples, the first a very simple one and the second which invalidates mGR in flat fiducial backgrounds for any combination of the three possible mass terms. Our conclusions and further comments on the causal properties of biGR and matter coupling difficulties are given in Sec. 5.

Sections 3 and 4 rely in part on work performed in collaboration with M. Sandora and G. Zahariade, originally reported in Refs. 14, 15 and 19.

2. The Method of Characteristics

The method of characteristics is a very useful tool for analyzing field theories, both classical and quantum. For the former, it determines hypersurfaces off which the system of nonlinear PDEs determining classical evolution is not fully predictive and along which shocks propagate. In the quantum setting, the characteristic method allows quantum field commutators to be analyzed. This is because the commutator function in a quantum field theory is related to the propagator and in turn to the predictability question. In short, the characteristic method probes the kinetic structure of field theories, so we begin with a brief review (see Refs. 13, 17, 18, 20–24 for further details).

2.1. *Brief review of characteristics*

The method of characteristics allows us to locate hypersurfaces beyond which the evolution of a system of PDEs ceases to be unique. Mathematically, this is characterized along the hypersurface by the vanishing of the coefficient of the highest order derivative in the normal direction. It can be intuitively understood as follows: An nth order differential equation can be generally solved once initial conditions for the $(n-1)$th order derivatives are specified. More explicitly, the initial conditions for $(i+1)$th order derivatives with $0 \leq i \leq n-2$ fix the evolution of ith order derivatives, while the evolution for the $(n-1)$th order derivative is obtained by solving the equation for the nth order derivative. However, if its coefficient vanishes, we cannot solve this equation, and so the evolution cannot be uniquely fixed.

The detailed set-up is as follows: Suppose that we have a hypersurface Σ and a vector ξ^μ not tangent to it (given a metric, ξ^μ can be chosen normal to Σ for simplicity). If the evolution equation in the direction of ξ^μ becomes singular, the evolution cannot be unique. Such a hypersurface is called a characteristic hypersurface. We

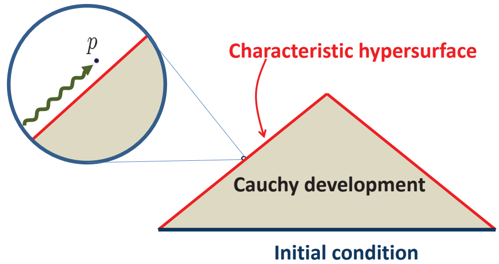

Fig. 1. The physics at a point p lying outside the Cauchy development of given initial datum S cannot be uniquely fixed. Propagation from outside S can affect p.

can decompose a spacetime into hypersurfaces Σ and normal directions ξ^μ (along the lines of the ADM formalism,[25] but for now do not stipulate whether ξ^μ is space-like, time-like or light-like with respect to some choice of metric). Now consider in this context a *quasi-linear* equation (i.e. the highest derivative order is n and that term appears linearly) for a scalar field ϕ

$$A^{\mu_1\cdots\mu_n}\partial_{\mu_1}\cdots\partial_{\mu_n}\phi + \mathcal{O}(\partial^{n-1}\phi) = 0\,. \tag{1}$$

The vanishing of the coefficient of the operator $\xi^{\mu_1}\cdots\xi^{\mu_n}\partial_{\mu_1}\cdots\partial_{\mu_n}$ is the condition that the evolution of Σ becomes singular:

$$A^{\mu_1\cdots\mu_n}\xi_{\mu_1}\cdots\xi_{\mu_n} = 0\,;$$

this is known as the *characteristic equation*. Observe that the form of the left-hand side of this equation is the same as the first term of the left-hand side of Eq. (1) upon replacing ∂_μ by ξ_μ.

The characteristic hypersurface is intimately related to the theory's *maximum propagation speed*. On it, the evolution of some physical variables becomes singular. This means that the hypersurface must be an edge of the Cauchy development from the initial data S. In GR, the characteristic hypersurface is none other than the light cone emanating from S (see Fig. 1), the speed of light being the upper bound for physical propagation. Consider a point p in a neighborhood of the characteristic hypersurface but outside of the Cauchy development; physical data at p cannot be uniquely determined given only the information on the characteristic hypersurface. Any point p that lies on the edge of the Cauchy development is affected by modes propagating along the characteristic hypersurface. Note that it is not necessary, though very convenient, to consider discontinuous modes; the discussion can be applied to any smooth (and in fact, analytic) propagation. We further comment on this point below.

2.2. *Characteristics versus shocks*

We shall now discuss the relation of characteristic hypersurfaces to shock waves (henceforth, "shocks"). Consider n first-order quasi-linear equations for n (not necessarily scalar) variables $\phi^{(a)}$:

$$\sum_b F^{\mu}_{(a,b)} \partial_\mu \phi^{(b)} + F_{(a)} = 0\,,$$

where "$a, b \in 1, \ldots, n$" are labels for the variables, and $F^{\mu}_{(a,b)}$ and $F_{(a)}$ are any functions of $\phi^{(a)}$. As shown above, the evolution from a hypersurface becomes singular if

$$\det[F^{\mu}_{(a,b)} \xi_\mu] = 0\,,$$

where the vector ξ^μ is normal to the hypersurface. Alternatively, we can consider a shock solution with discontinuities in its first-order derivatives across a hypersurface normal to ξ^μ. The shock (the jump in first-order derivatives) is denoted by

$$[\partial_\mu \phi^{(a)}] = \xi_\mu \tilde{\phi}^{(a)}\,, \tag{2}$$

which implies the junction condition

$$\sum_b F^{\mu}_{(a,b)} \xi_\mu \tilde{\phi}^{(b)} = 0\,.$$

Notice that the requirement for the existence of a nontrivial solution for $\tilde{\phi}^{(a)}$ is that the determinant of the above matrix of coefficients is zero. This is exactly the characteristic matrix. In other words, shock fronts must lie along characteristic hypersurfaces. Technically a shock analysis is easier than the more general method of characteristics, which holds even for analytic, shock-free configurations.

2.3. *Characteristics versus strong coupling*

The analysis of characteristics is completely independent of the strong coupling scale. For instance, consider a scalar field action in Minkowski spacetime:

$$S = -\frac{1}{2}\int d^4x \left[\alpha^2(\partial_\mu\varphi\partial^\mu\varphi + m^2\varphi^2) + \frac{\lambda}{6}\,\varphi^4\right].$$

Because the characteristic method analyzes propagation, it focuses on the behavior of the kinetic term. Thus here (and in many other similar situations) at energy scales $E \gg m$, a good approximation is to ignore the mass term. On the other hand, in a strong coupling analysis, rather than the mass, the magnitude of the coupling constant λ/α^4 of the canonically normalized field $\varphi_c := \alpha\varphi$ is the relevant quantity. In particular, in an mGR context, we can be confident of the characteristic analysis as long as the energy scale is larger than the — *observationally very small* — graviton mass. This energy scale can therefore be *much smaller* than the strong coupling scale $(m^2 M_P)^{1/3}$.[7] Of course, in models (such as mGR) with constraints, the mass terms can re-enter the characteristic matrix via the constraints themselves. Indeed it has been suggested in Ref. 26 that characteristic surfaces indicate a breakdown of the theory's effective description of some putative ultraviolet completion. A more detailed study of the relationship between ultraviolet completions of the effective description and low energy gravitational physics was performed recently in Ref. 27.

2.4. *Characteristics and branches*

The characteristic method neatly captures the difficulties faced by constrained systems whose solutions have multiple branches. As an example, consider the following constraint equation,

$$C[\phi^{(a)}] = C_1[\phi^{(a)}]C_2[\phi^{(a)}] = 0\,, \tag{3}$$

where C_1 and C_2 are independent functionals of dynamical fields $\phi^{(a)}$. Here we have two obvious choices for solving the constraint $C = 0$, namely $C_1 = 0$ or $C_2 = 0$. Each choice yields a different branch of solutions. Suppose we choose the constraint $C_1 = 0$. Although Eq. (3) is not quasi-linear, we can act with $\xi^\mu\partial_\mu$ to yield a, first-order, quasi-linear equation:

$$C_2[\phi^{(a)}]\xi^\mu\partial_\mu C_1[\phi^{(a)}] + C_1[\phi^{(a)}]\xi^\mu\partial_\mu C_2[\phi^{(a)}] = 0\,.$$

Given the choice $C_1 = 0$, the second term vanishes, so the corresponding characteristic equation is $C_2 = 0$. If we change branch from $C_1 = 0$ to $C_2 = 0$ along some hypersurface, then this surface must be characteristic. Thus, upon branch change, evolution ceases to be unique. In a Hamiltonian analysis, this pathology manifests itself as a singular structure of the Dirac bracket.

3. Fiducial Massive Gravity

Fiducial mGR (fmGR) consists of a dynamical massive spin-2 field on a fixed, fiducial background $\bar{g}_{\mu\nu}$, free of the bulk ghost excitations that plague mGR theories.[5] Models of this type were first uncovered[2] in a study of FP limits of one of the metrics in the original "f-g" two-metric model,[1] keeping the other as a fixed background. However the whole subject lay long dormant after it was discovered[3] that, at least for generic mass terms, a sixth, ghost, excitation appeared beyond linear order. The subject awoke from hibernation when effective field theory decoupling limit technology was employed to show that a distinguished, three-parameter family of mass terms might be ghost-free.[5] This was verified for generic regions of field space at full nonlinear order in Ref. 8. Those studies were performed in a, rather intricate, metric formulation; later, it was realized that the original, highly efficient, frame-like methods of Ref. 1 could be employed.[28–33] This allowed a covariant analysis of the models' DoF[12–14] and its characteristic matrix.[13,14,17,18] Recently these analyses were extended to the models' full parameter range;[15] the key advantage of the frame-like approach was that differential forms are exactly the correct objects for covariantly describing the hypersurfaces required for evolution and constraint analyses. In that language, suppressing wedge products whenever obvious, the action is simply

$$S[e,\omega] = -\epsilon_{mnrs}\int\bigg(\frac{1}{4}e^m e^n[d\omega^{rs} + \omega^r{}_t\omega^{ts}] \\ - m^2 e^m\left[\frac{\beta_0}{4}e^n e^r e^s + \frac{\beta_1}{3}e^n e^r f^s + \frac{\beta_2}{2}e^n f^r f^s + \beta_3 f^n f^r f^s\right]\bigg). \tag{4}$$

Here $e^m := e_\mu{}^m\, dx^\mu$, $e_\mu{}^m \eta_{mn} e_\nu{}^n = g_{\mu\nu}$ and similarly for the fiducial variables. For the model to have a linear FP limit about its fiducial background, its parameters must obey

$$m^2(\beta_0 + \beta_1 + \beta_2 + \beta_3) = \frac{\bar{\Lambda}}{6}\,, \tag{5}$$

where $\bar{\Lambda}$ is the fiducial cosmological constant. In that case, the FP mass is

$$m_{\rm FP}^2 = m^2(\beta_1 + 2\beta_2 + 3\beta_3)\,. \tag{6}$$

3.1. *Constraints*

The models' constraints are very easily understood; here is their counting, as given in Ref. 15:

(1) There are 40 equations of motion. Varying the independent spin connection ω^{mn} imposes vanishing torsion and varying the vierbein e^m sets the Einstein tensor equal to the stress tensor of the mass terms.
(2) Since the equations of motion are forms, evaluating them along a space-like hypersurface cannot give any time derivatives on dynamical fields; this gives sixteen primary constraints.
(3) The anti-symmetric part of the Einstein tensor, plus the covariant curl of the Einstein tensor equation of motion yield the so-called symmetry and vector constraints. These constitute ten secondary constraints.
(4) The covariant curls of the symmetry and vector constraints yield the four remaining tertiary constraints.

The above constraint structure leaves ten independent variables subject to ten independent first-order evolution equations, and thus five physical DoF.

4. Fiducial Massive Gravity Characteristics

Since we are interested in vanishing of the characteristic matrix's determinant, our strategy is to use Gaußian elimination to maximally simplify its associated homogeneous linear equations, and then study the determinant of the remaining, *reduced characteristic matrix* determined this way. We compute the shocks in the equations of motion and constraints to find this matrix. For that, one begins by writing the equations of motion in first-order form

$$\begin{aligned} 0 &= \nabla e^m\,, \\ 0 &= \frac{1}{2}\epsilon_{mnrs} e^n R^{rs} - m^2 t_m\,, \\ 0 &= d\omega^{mn} + \omega^m{}_t \omega^{tn} - R^{mn}\,. \end{aligned} \tag{7}$$

Here we have treated the Riemann curvature two-form as an independent field and therefore added the third equation above to the system. Also, ∇ is the covariant curl with respect to the spin connection ω^{mn} and the mass stress tensor is encoded by the three-form

$$t_m := \epsilon_{mnrs}[\beta_0 e^n e^r e^s + \beta_1 e^n e^r f^s + \beta_2 e^n f^r f^s + \beta_3 f^n f^r f^s] \,.$$

Then we denote the shock discontinuity across Σ (see Eq. (2)) in $(e^m, \omega^{mn}, R^{mn})$ by

$$[\partial_\mu e^m] = \xi_\mu \tilde{e}^m \,, \qquad [\partial_\mu \omega^{mn}] = \xi_\mu \tilde{\omega}^{mn} \,, \qquad [\partial_\mu e^m] = \xi_\mu \tilde{R}^{mn} \,.$$

To study characteristics that are space-like with respect to the dynamical metric, we assume (without loss of generality), along the shock hypersurface Σ, that

$$\xi_\mu g^{\mu\nu} \xi_\nu = -1 \,.$$

Thus we can use the projector $\Pi^\mu_\nu := \delta^\mu_\nu + \xi^\mu \xi_\nu$ to decompose tensors, and hence any differential form α, as

$$\alpha := \xi \mathring{\alpha} + \boldsymbol{\alpha} \,,$$

where $\xi := \xi_\mu \, dx^\mu$ and $\boldsymbol{\alpha}$ is the purely spatial part of α. In this notation, we can easily compute the shocks in the PDEs in Eq. (7) as well as that of the covariant curl of the third of these (which says $\nabla R^{mn} = 0$). Altogether, this forces vanishing of the spatial parts of the shock profiles $(\tilde{e}^m, \tilde{\omega}^{mn}, \tilde{R}^{mn})$:

$$\tilde{\mathbf{e}}^m = 0 = \tilde{\boldsymbol{\omega}}^{mn} = \tilde{\mathbf{R}}^{mn} \,.$$

If we restrict ourselves to the branch where the symmetry constraint (recall that symmetry of the Einstein tensor implies the same for the mass stress tensor) is solved via

$$f^m e_m = 0 \,,$$

then the analysis is simplified by using $f_{\mu\nu} = f_\mu{}^m e_{\nu m}$ as independent variable. The shock of the symmetry constraint implies that

$$\tilde{e}^m = \xi \xi^\mu l_\mu{}^m \tilde{f}_{oo} \,,$$

where

$$\tilde{f}_{oo} := \xi^\mu \xi^\nu \tilde{e}_\mu{}^m f_{\nu m}$$

and $l^\mu{}_m$ denotes the inverse fiducial vierbein.

At this point, only independent shock profiles with time-like components, namely $(\tilde{f}_{oo}, \mathring{\tilde{\omega}}^{mn}, \mathring{\tilde{R}}^{mn})$, remain. The shock in the curvature identity (which becomes an equation of motion now that the Riemann curvature is taken independent)

$$\nabla^\mu R_{\mu\nu\rho\sigma} = 2\nabla_{[\rho} G_{\sigma]\nu} + g_{\nu[\rho} \nabla_{\sigma]} G_\mu{}^\mu \,,$$

can be used to determine the profiles $\mathring{\tilde{R}}^{mn}$ (see Ref. 15 for the detailed formula which will not be important to us here). This leaves seven shock profiles $(\tilde{f}_{oo}, \mathring{\tilde{\omega}}^{mn})$ undetermined. They obey a set of homogeneous linear equations that determine the reduced characteristic matrix. These are determined from the shocks of the vector, scalar and curled symmetry constraints. Schematically the result takes the form

$$\begin{pmatrix} M & N_{mn} \\ P^{(6)} & Q^{(6)}_{mn} \end{pmatrix} \begin{pmatrix} \tilde{f}_{oo} \\ \mathring{\tilde{\omega}}^{mn} \end{pmatrix} = 0\,.$$

The label "(6)" is used to indicate a column vector built from a pair of (spatial) three-vectors, so the reduced characteristic matrix is 7×7. It depends on (i) the dynamical and fiducial vierbein, (ii) the *contorsion* tensor given by the difference between dynamical and fiducial connections, (iii) the fiducial curvature, and (iv) for models with $\beta_3 \neq 0$, the dynamical curvature.

To ensure the absence of superluminal shocks, one would need to demonstrate that the reduced characteristic matrix determinant cannot vanish anywhere. Clearly, this is highly unlikely, given its dependence on both dynamical and fiducial fields. However, there does remain the hope that, at least for some magical combination of parameters (and possibly a special choice of fiducial background), this miracle occurs. Not only is there no evidence for this, but counterexamples are rather easy to find, as we shall soon show. Superluminal shocks are therefore generic features of fmGR models.

The above argument can also be adapted to show that there are superluminal shocks with respect to the fiducial metric; this implies that the model, viewed as a quantum theory of a massive spin-2 field propagating in a background, is pathological. However, there remains the possibility that it somehow still manages to avoid CTCs — the presence of superluminal propagation is, after all, just a warning that something nasty may plague the theory. Indeed, as pointed out in Ref. 34, superluminality may not always present a problem since it does not necessarily lead to acausality[35] (the existence of closed causal curves (CCCs);[20] see Refs. 21 and 22 for examples of acausality in modified theories of gravity without diffeomorphism invariance or local Lorentz invariance). Here, on the contrary however, we will present an *explicit* construction of CCCs.[a]

4.1. *Acausality*

First note that acausality here refers to local (infinitesimal) closed propagation curves. This "micro-acausality" differs from acausal GR solutions: the (Gödel) CTCs in GR are *global*, in the sense that locally in a neighborhood of a point on the curve, an observer can always define future and past. In contrast, the fmGR acausalities are *local* so that the causal structure is broken even in an infinitesimal region. Indeed, CCCs in fmGR are *generated dynamically*, whereas in GR, starting from a spacetime without CCCs, it is at best extremely difficult, if not impossible

[a] Here we differentiate (slightly pedantically) between CTCs — closed curves in a spacetime whose tangent vector is everywhere time-like — and CCCs — loops around which initial data can be transmitted back to its starting point (typically by propagating shocks around loops along characteristic surfaces).

to form CCCs without breaking energy conditions or causing event horizons that shroud the CCCs.[36]

To construct an explicit example of acausality, we consider the case where the background and dynamical metrics (the latter viewed as a mean field along which shocks propagate) are both flat but not causally aligned:

$$d\bar{s}^2 = -dz^2 + dx^2 + dy^2 + dt^2 \qquad \text{and} \qquad ds^2 = -dt^2 + dx^2 + dy^2 + dz^2\,, \tag{8}$$

with $f^3 = e^0 = dt$, $f^1 = e^1 = dx$, $f^2 = e^2 = dy$ and $f^0 = e^3 = dz$. This configuration is an fmGR solution so long as (in addition to the linearization condition Eq. (5) at $\bar{\Lambda} = 0$) its parameters obey[15]

$$\beta_1 + 2\beta_2 + 3\beta_3 = 0\,. \tag{9}$$

Notice that this requirement implies that the linearized FP mass $m_{\rm FP} = 0$ (see Eq. (6)), so one might already rule this model out on the basis that its nonlinear and linearized DoF counts do not agree. Nonetheless, it exemplifies the generic difficulties faced by models with a field-dependent characteristic matrix. In the current "flat on flat", and hence contorsion-free, setting, the reduced characteristic matrix simplifies significantly to

$$\begin{pmatrix} M & 0 \\ 0 & \mathbf{e}^m \wedge \mathbf{f}^n \\ 0 & \epsilon^{mn}{}_{rs}[\beta_1 \mathbf{e}^r \wedge \mathbf{e}^s + 2\beta_2 \mathbf{e}^r \wedge \mathbf{f}^s + 3\beta_3 \mathbf{f}^r \wedge \mathbf{f}^s] \end{pmatrix} \begin{pmatrix} \tilde{f}_{oo} \\ \tilde{\mathring{\omega}}_{mn} \end{pmatrix} = 0\,. \tag{10}$$

Here, because this matrix is block-diagonal, the exact form of the scalar M will not concern us; it is given in Ref. 15. Note that because $\mathbf{e}^m \wedge \mathbf{f}^n$ is a spatial two-form, contracted on the shock profile $\tilde{\mathring{\omega}}_{mn}$ this gives three conditions, and the same comment applies for the last line of the matrix above, so we have, as promised, a 7×7 homogeneous linear system of equations. Now consider the two, relatively tilted, time-like vectors

$$\xi_A = \frac{\partial}{\partial t} \qquad \text{and} \qquad \xi_B = \frac{\partial}{\partial t} + \alpha \frac{\partial}{\partial z}\,, \qquad 0 < \alpha < 1\,.$$

Vector A defines space-like hypersurfaces $t = \text{const.}$, which we call Σ_A. These are characteristic[b] because the shock profile condition imposed by the second line of the reduced characteristic matrix (calling $\xi^\mu \omega_{\mu mn} := \omega_{omn}$)

$$\mathbf{e}^m \wedge \mathbf{f}^n\, \tilde{\omega}_{omn} = 2\tilde{\omega}_{o12}\, dx \wedge dy + ([\tilde{\omega}_{o10} + \tilde{\omega}_{o13}]dx + [\tilde{\omega}_{o20} + \tilde{\omega}_{o23}]dy) \wedge dz\,,$$

has solutions for nonvanishing shock profiles $\tilde{\omega}_{o01} = \tilde{\omega}_{o13}$, $\tilde{\omega}_{o02} = \tilde{\omega}_{o23}$ and $\tilde{\omega}_{o12} = 0$, along which the condition implied by the third line of the reduced characteristic

[b] Here will use this property to send signals around a CTC using superluminal shocks. We also now know that initial data along Σ_A (determined by restriction of the solution Eq. (8) to one of the characteristics surfaces Σ_A) does not uniquely determine subsequent evolution. Note that this does NOT mean that there are no solutions to the original PDEs with this initial data: rather, there are many of them, one of which is Eq. (8). In other words, the model has lost predictive power.

matrix is trivially satisfied (so long as the parameters obey the on-shell condition Eq. (9)). We now need to find a second, independent, family of space-like hypersurfaces, Σ_B (say), since we could then form a CTC by sequentially sending (i) a right-moving[c] signal along Σ_A, (ii) a right-moving signal along Σ_B, (iii) a left-moving signal along Σ_A, (iv) a left-moving signal along Σ_B back to the starting spacetime point.

Thus it only remains to establish that surfaces of constant[d] $\tau = -t + \alpha z$ are characteristic. For that we recompute the condition given by second line of the reduced characteristic matrix along Σ_A and now find

$$\mathbf{e}^m \wedge \mathbf{f}^n \, \tilde{\omega}_{omn} = 2\tilde{\omega}_{o12} \, dx \wedge dy + (1+\alpha)([\tilde{\omega}_{o10} + \tilde{\omega}_{o13}]dx + [\tilde{\omega}_{o20} + \tilde{\omega}_{o23}]dy) \wedge dz \,,$$

which is solved exactly as above. The third line of the characteristic equation again gives no new conditions, so Σ_B is characteristic (for any choice of tilting α).

It is not difficult to generate even more general acausalities valid for any choice of the model's parameters $(\beta_0, \beta_1, \beta_2, \beta_3)$: For simplicity, we again consider a flat fiducial metric with tetrads $f^0 = dt$, $f^1 = dx$, $f^2 = dy$, $f^3 = dz$ for which the fiducial connection vanishes. Now suppose that, along a constant-t hypersurface Σ, the fiducial and the dynamical tetrads are aligned but have different amplitudes:

$$e^0 = A \, dt \,, \qquad e^1 = B \, dx \,, \qquad e^2 = C \, dy \,, \qquad e^3 = -B \, dz \,,$$

where A, B and C are constants. Observe, for later, that e^1 and e^3 have the same amplitudes but different signs. We further assume that all connection components, as well as their spatial derivatives vanish along this "initial" hypersurface Σ. Thus, along Σ, the only non-vanishing derivatives of dynamical fields are with respect to the fiducial time coordinate t. The initial configuration for the Riemann tensor (when viewed as an independent field) is determined directly from the third, algebraic, equation in (7).

We must first check whether this initial configuration is consistent with the constraints (these were given explicitly for the first time in Ref. 15 and are summarized in Sec. 3.1):

(i) The 16 primary constraints follow from the spatial parts of the equations in (7). Only the second of these gives a nontrivial condition:

$$0 = 2m^2(-3\beta_0 B^2 C - \beta_1 B^2 + \beta_2 C + 3\beta_3) \,.$$

This condition can be satisfied by tuning C to

$$C = \frac{3\beta_3 - \beta_1 B^2}{3\beta_0 B^2 - \beta_2} \,.$$

[c]This slight abuse of language is designed to help the reader visualize the (t, z) projection of Minkowski spacetime.

[d]Notice $d\tau = (-1, 0, 0, \alpha)_\mu dx^\mu$ so $\xi^\mu \partial_\mu = (-1, 0, 0, 1)_\mu \eta^{\mu\nu} \partial_\nu = \frac{\partial}{\partial t} + \alpha \frac{\partial}{\partial z} = \xi_B$. We do not bother normalizing the vector ξ_B since this only contributes an overall irrelevant constant.

The singular cases $C = 0, \infty$ are avoided by requiring that the parameter B obeys neither $3\beta_3 - \beta_1 B^2 = 0$ nor $3\beta_0 B^2 - \beta_2 = 0$.

(ii) The ten secondary constraints are both trivially satisfied. The six symmetry constraints require that $e^m f_m = 0$ along the initial surface which obviously holds, while the four vector constraints are proportional to the contorsion which vanishes along the Σ for this configuration (this need not be the case upon evolving the system).

(iii) Of the four remaining tertiary constraints, three are proportional to the contorsion (they are just the covariant curl of the secondary symmetry constraint). Thus only the scalar constraint must be verified. As shown in Ref. 15 (see Eq. (15), Sec. 2.3.2), for vanishing contorsion and flat fiducial backgrounds, this constraint is rather simple

$$\epsilon_{mnrs}(\beta_1 e^m e^n + 2\beta_2 e^m f^n + 3\beta_3 f^m f^n) \wedge R^{rs} \approx 0\,.$$

Along Σ the only non-vanishing part of the Riemann tensor is $dt \wedge dx^j\, R_{tj}{}^{rs}$ so the sums over m and n in the above display can run only over values $1, 2, 3$ (because $dt \wedge dt = 0$) so the scalar constraint takes the form

$$M_{12}\, dx \wedge dy \wedge R^{03} + M_{23}\, dy \wedge dz \wedge R^{01} + M_{31}\, dz \wedge dx \wedge R^{02} \approx 0\,,$$

for constants M_{ij}. Thus only the part of $R^{0i} \propto dt \wedge dx^i$ can contribute to the scalar constraint. But these can be computed from the second equation in (7) (note that this feature is guaranteed by the covariant proof given in Ref. 15 that the above is a constraint). This gives a (somewhat complicated) expression for the scalar constraint which is linear in the variable A. The key point is, that for *any* choice of parameters $(\beta_0, \beta_1, \beta_2, \beta_3)$, a value for A can be found such that the scalar constraint holds. Note that the constant B is still free, so we have really found a one parameter family of field data for which Σ will be characteristic.

Next we demonstrate that the initial hypersurface Σ is characteristic. Because the contorsion vanishes along Σ, we can use the simplified version of the characteristic matrix given in Eq. (10). The key feature of this initial configuration is that the coefficients of e^1 and e^3 have the same amplitude but opposite signs. Then, the $dx \wedge dz$ component in the middle line of the reduced characteristic matrix gives no condition. Thus, the hypersurface Σ is characteristic.

We are now ready to generate acausalities. The above analysis shows that, in a flat Minkowski background there are configurations along constant t initial hypersurfaces whose evolution is not uniquely determined. This implies superluminal (with respect to the background) propagation along this characteristic hypersurface. To generate propagation around local loops, we first note that the above configuration has a parity symmetry: the action is invariant under a simultaneous flip of pairs of fiducial and dynamical tetrad components, for example $(e^1, f^1, e^2, f^2) \to (-e^1, -f^1, -e^2, -f^2)$. (The mass term is obviously invariant under this transformation while invariance of the Einstein–Hilbert term requires also transforming

$\omega^{12} \to \omega^{12}$, $\omega^{xy} \to \omega^{xy}$, $\omega^{1x} \to -\omega^{1x}$ and $\omega^{2x} \to -\omega^{2x}$ where $x, y \neq 1, 2$.) For our initial configuration, the map $(e^1, f^1, e^2, f^2) \to (-e^1, -f^1, -e^2, -f^2)$ produces a new solution to the constraints (because changing coordinates $(x, y) \to (-x, -y)$ yields the original configuration back again). In this way we use parity symmetries to generate new solutions. Now consider a small signal propagating in the direction $(0, a, b, c)$. The parity symmetry can be used to generate signals propagating also in the directions $(0, a, -b, -c)$, $(0, -a, b, -c)$ and $(0, a, -b, -c)$. From these, a CCC can clearly be constructed.

The above examples are damning for fmGR, but yeasayers might complain that these are only very special initial conditions. Needless to say, however, one counterexample suffices to establish inconsistency. Actually the real lesson is that generating ill-posed configurations is easy, and likely a completely generic feature of the model. In any case, what might seem to be a possible (albeit unlikely) escape route — careful tunings of the parameters and fiducial background designed to avoid space-like characteristics — is actually a major shortcoming. Absent some guiding principle for choosing the fiducial space, observational predictability seems to be completely lost in the swamp of what amounts to infinitely many tunable parameters. Indeed, another essential ingredient in completing these models is to specify, and see the consequences of, their coupling to matter. For fmGR, a variety of difficulties were already exhibited in Ref. 37 (see also Ref. 38); these were unavoidably and precisely linked to the theory's constraint structure. In a theory with two metrics it is unclear how to couple matter's stress energy without destroying biGR's physical DoF content. (See, for example, the conflicting Refs. 39 and 40 (as well as Ref. 41) regarding whether any two-metric combinations can consistently couple to matter.) Thus, attempting to leave the borders of GR forces one ineluctably back to it and the gauge principle as the unique mechanisms for coupling geometry to matter.

5. Outlook: Bimetric Gravity?

The necessary incompatibility of the dynamical and fiducial causal structures, and its attendant superluminality and acausalities, as well as the massive loss of observational predictability implied by choosing, by fiat, a fiducial metric, leaves biGR as the only conceivable resolution: perhaps a magical combination of two dynamical metrics could determine a good causal structure. While there is as yet no complete study of biGR's characteristics, there is already strong evidence against its consistency. This was garnered by considering a putative partially massless (PM) limit. Recall that for the linear FP theory in cosmological backgrounds, there is an intermediate spin-2 theory, whose excitations were light-like, but describe only helicities $(\pm 2, \pm 1)$.[42,43] The PM model was discovered by tuning the FP mass to the cosmological constant in order to achieve gauge invariant, lightcone propagation. Both fmGR and biGR possess PM free limits for special points in their parameter spaces.[44,45] However, both these models suffer from a new version of the original "fatal flaw"[3] of the generic mGR models, namely their interacting and linearized physical DoF contents do not match.[14,19,46,47] We hope to complete, and report on, the outcome of such explicit biGR consistency calculations.[48]

Acknowledgments

K.I. is supported by Taiwan National Science Council under Project No. NSC101-2811-M-002-103. A.W. thanks the Indian Institute of Science Bangalore, and the Harish-Chandra Institute, Allahabad for hospitality. S.D. was supported in part by NSF and DOE grants PHY-1266107 and #DE-SC0011632, respectively.

References

1. A. Salam and J. Strathdee, *Phys. Rev.* **184**, 1750, 1760 (1969); C. J. Isham, A. Salam and J. Strathdee, *Phys. Lett. B* **31**, 300 (1970).
2. B. Zumino, Effective Lagrangians and broken symmetries, in *Lectures on Elementary Particles and Quantum Field Theory*, Vol. 2, eds. S. Deser, M. Grisaru and H. Pendleton (MIT Press, 1970), p. 437.
3. D. Boulware and S. Deser, *Phys. Rev. D* **6**, 3368 (1972); *Phys. Lett. B* **40**, 227 (1972).
4. M. Fierz and W. Pauli, *Proc. Roy. Soc. Lond. A* **173**, 211 (1939); M. Fierz, *Helv. Phys. Acta* **12**, 3 (1939).
5. C. de Rham and G. Gabadadze, *Phys. Rev. D* **82**, 044020 (2010); C. de Rham, G. Gabadadze and A. J. Tolley, *Phys. Rev. Lett.* **106**, 231101 (2011); *Phys. Lett. B* **711**, 190 (2012); *JHEP* **1111**, 093 (2011).
6. K. Hinterbichler, *Rev. Mod. Phys.* **84**, 671 (2012).
7. C. de Rham, *Living Rev. Relativ.* **17**, 7 (2014).
8. S. F. Hassan and R. A. Rosen, *Phys. Rev. Lett.* **108**, 041101 (2012); S. F. Hassan, R. A. Rosen and A. Schmidt-May, *JHEP* **02**, 026 (2012); S. F. Hassan and R. A. Rosen, *ibid.* **04**, 123 (2012); S. Hassan, A. Schmidt-May and M. von Strauss, *Phys. Lett. B* **715**, 335 (2012).
9. J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, 1965).
10. K. Johnson and E. C. G. Sudarshan, *Ann. Phys.* **13**, 126 (1961).
11. G. Velo and D. Zwanziger, *Phys. Rev.* **186**, 1337 (1969); *ibid.* **188**, 2218 (1969).
12. C. Deffayet, J. Mourad and G. Zahariade, *J. Cosmol. Astropart. Phys.* **1301**, 031 (2013).
13. S. Deser and A. Waldron, *Phys. Rev. Lett.* **110**, 111101 (2013).
14. S. Deser, M. Sandora and A. Waldron, *Phys. Rev. D* **87**, 101501 (2013).
15. S. Deser, M. Sandora, A. Waldron and G. Zahariade, *Phys. Rev. D* **90**, 104043 (2014).
16. C. Burrage, C. de Rham, L. Heisenberg and A. J. Tolley, *J. Cosmol. Astropart. Phys.* **1207**, 004 (2012); P. de Fromont, C. de Rham, L. Heisenberg and A. Matas, *JHEP* **1307**, 067 (2013); A. Gruzinov, arXiv:1106.3972.
17. K. Izumi and Y. C. Ong, *Class. Quantum Grav.* **30**, 184008 (2013).
18. S. Deser, K. Izumi, Y. C. Ong and A. Waldron, *Phys. Lett. B* **726**, 544 (2013); arXiv:1312.1115.
19. S. Deser, M. Sandora and A. Waldron, *Phys. Rev. D* **88**, 081501 (2013).
20. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2 (Wiley-Interscience, 1962), pp. 590, 596.
21. H. Chen, J. M. Nester and H.-J. Yo, *Phys. Pol. B* **29**, 961 (1998).
22. Y. C. Ong, K. Izumi, J. M. Nester and P. Chen, *Phys. Rev. D* **88**, 024019 (2013).

23. K. Izumi, J.-A. Gu and Y. C. Ong, *Phys. Rev. D* **89**, 084025 (2014).
24. K. Izumi, *Phys. Rev. D* **90**, 044037 (2014); H. Reall, N. Tanahashi and B. Way, *Class. Quantum Grav.* **31**, 205005 (2014); arXiv:1409.3874.
25. R. L. Arnowitt, S. Deser and C. W. Misner, *Phys. Rev.* **116**, 1322 (1959); *ibid.* **117**, 1595 (1960).
26. H. Georgi, *Weak Interactions and Modern Particle Theory* (Dover Publications, 2009); N. Arkani-Hamed, H. Georgi and M. D. Schwartz, *Ann. Phys.* **305**, 96 (2003).
27. N. Kaloper, A. Padilla, P. Saffin and D. Stefanyszyn, arXiv:1409.3243.
28. C. Deffayet, J. Mourad and G. Zahariade, *JHEP* **1303**, 086 (2013).
29. A. H. Chamseddine and V. Mukhanov, *JHEP* **1108**, 091 (2011).
30. A. H. Chamseddine and M. S. Volkov, *Phys. Lett. B* **704**, 652 (2011).
31. K. Hinterbichler and R. A. Rosen, *JHEP* **1207**, 047 (2012).
32. S. F. Hassan, A. Schmidt-May and M. von Strauss, arXiv:1204.5202.
33. K. Nomura and J. Soda, *Phys. Rev. D* **86**, 084052 (2012).
34. R. Geroch, arXiv:1005.1614.
35. H. Andréka, J. X. Madarász, I. Németi, M. Stannett and G. Székely, *Class. Quantum Grav.* **31**, 095005 (2014).
36. S. Deser, R. Jackiw and G. 't Hooft, *Phys. Rev. Lett.* **68**, 267 (1992).
37. S. Deser and A. Waldron, *Phys. Rev. D* **89**, 027503 (2014).
38. İ. Güllü and B. Tekin, *Phys. Lett. B* **728**, 268 (2014).
39. Y. Yamashita, A. De Felice and T. Tanaka, arXiv:1408.0487.
40. C. de Rham, L. Heisenberg and R. H. Ribeiro, arXiv:1408.1678.
41. J. Noller and S. Melville, *J. Cosmol. Astropart. Phys.* **01**, 003 (2015); V. O. Soloviev, arXiv:1410.0048.
42. S. Deser and R. I. Nepomechie, *Ann. Phys.* **154**, 396 (1984); *Phys. Lett. B* **132**, 321 (1983).
43. S. Deser and A. Waldron, *Phys. Rev. Lett.* **87**, 031601 (2001); *Nucl. Phys. B* **607**, 577 (2001); *Phys. Lett. B* **508**, 347 (2001); *ibid.* **513**, 137 (2001); *Nucl. Phys. B* **662**, 379 (2003); *Phys. Lett. B* **603**, 30 (2004); *Phys. Rev. D* **74**, 084036 (2006); M. Grigoriev and A. Waldron, *Nucl. Phys. B* **853**, 291 (2011); *Phys. Rev. D* **87**, 087702 (2013); S. Deser, E. Joung and A. Waldron, *ibid.* **86**, 104004 (2012); *J. Phys. A* **46**, 214019 (2013).
44. C. de Rham and S. Renaux-Petel, *J. Cosmol. Astropart. Phys.* **1301**, 035 (2013).
45. S. F. Hassan, A. Schmidt-May and M. von Strauss, arXiv:1208.1797; Y. Akrami, T. S. Koivisto, D. F. Mota and M. Sandstad, *J. Cosmol. Astropart. Phys.* **1310**, 046 (2013).
46. C. de Rham, K. Hinterbichler, R. A. Rosen and A. J. Tolley, *Phys. Rev. D* **88**, 024003 (2013).
47. E. Joung, W. Li and M. Taronna, *Nucl. Phys. B* **886**, 712 (2014) ; Y. M. Zinoviev, arXiv:1405.4065.
48. S. Deser, K. Izumi, Y. C. Ong, A. Waldron and G. Zahariade, in preparation.