International Journal of Theoretical Physics, Vol. 39, No. 1, 2000

Shock-Free Wave Propagation in Gauge Theories

J. McCarthy[1,2] and Ö. Sarioğlu[3]

Received May 14, 1999

We present the shock-free wave propagation requirements for massless fields. First, we briefly argue how the "completely exceptional" approach, originally developed to study the characteristics of hyperbolic systems in $1 + 1$ dimensions, can be generalized to higher dimensions and used to describe propagation without emerging shocks, with characteristic flow remaining parallel along the waves. We then study the resulting requirements for scalar, vector, vector-scalar, and gravity models and characterize physically acceptable actions in each case.

1. INTRODUCTION

In this work, a brief version of which appeared in ref. 1, we study the propagation of excitations of classical massless field actions. In general, criteria for physical propagation of such waves can be derived in many ways. Here, we will only consider the "completely exceptional" (CE) approach [2], originally developed for systems in $D = 1 + 1$. Roughly speaking, complete exceptionality is the property ensuring that the initial "wavefronts" evolve so as to prevent the emergence of shocks, which, in general, result when the "characteristics" propagate at different speeds. As we are not aware of a rigorous procedure extending ideas developed at $D = 2$ to higher D, we will follow steps similar to those in $D = 2$, and then outline how to generalize them to higher dimensions. In the process, we show how the CE idea can be looked at in seemingly different ways and outline a derivation that fills the gap between the two viewpoints. We apply our criteria to massless spin-0, 1, 2 nonlinear systems.

[1] Department of Physics and Mathematical Physics, University of Adelaide, SA 5005, Austraila.
[2] Present address: DSTO (ITD), Salisbury SA 5108, Australia; e-mail: Jim.McCarthy@dsto.defence.gov.au.
[3] Department of Physics, Brandeis University, Waltham, Massachusetts 02454-9110; e-mail: sarioglu@brandeis.edu.

We start in Section 2 by introducing the type of physical problems that we will study and develop the formalism that will be used throughout. Section 3 gives the analysis of characteristic surfaces, which are crucial to the CE idea. In Section 4, we give a simple example in $D = 2$, and demonstrate how shocks may be prevented for this particular problem. Motivated by this example, we next show how the introduced ideas can be extended to higher dimensions in Section 5. This naturally leads to the CE concept and we show how one can view it in two seemingly different ways, which are explained in the text. In Section 6, we study in detail the scalar field in $D = 4$ using these two separate methods, derive the CE condition on it, and argue as to how one can generalize the result to arbitrary D. Next, we turn to models of nonlinear electrodynamics in Section 7. Here we encounter particular models, the constraints on which not only automatically guarantee the CE property (as originally discussed in ref. 3), but also ensure that both polarizations of light propagate according to the same dispersion law, i.e., "no birefringence" [4, 5]. Hence we call these constraints the "strong CE" conditions. We also derive (for the first time, to our knowledge) the regular CE requirement conditions (much weaker than strong CE) in the most general $D = 4$ case. Finally, in Section 8, we find that wide classes of gravity models share with the Einstein case the null nature of their characteristic surfaces. In three appendices, we show the details of some calculations skipped in the text.

2. THE FORMALISM

In this paper we will be dealing with systems of PDEs that are Euler–Lagrange equations of relativistic actions. They will be linear in highest derivatives (quasilinear) and their coefficients will not depend on the coordinates explicitly. So they can be reduced to a set of differential equations of first derivative order. Hence for U an N-vector of fields, $\mathscr{A}$ an $N \times N$ matrix, and $\mathscr{B}$ an N-vector (both arbitrary smooth functions of U), the equations of interest can always be written in the form

$$\mathscr{A}^{\mu}(U)(\partial_{\mu} U) + \mathscr{B}(U) = 0 \tag{2.1}$$

The theory of such equations in arbitrary dimensions is quite difficult, but we will be mainly interested in the evolution of the spatial boundary of a wave propagating into some given vacuum. So, with $\overline{U}$ some smooth (say at least C^1) solution, at some initial time we have some spatial region outside of which the "state" is the "vacuum solution" $\overline{U}$, and across the boundary surface the full solution U is continuous, but its first derivative may not be. We want to consider the evolution of such initial "wavefronts."

We will follow the formalism developed for this situation in refs. 2 and 3 and the references therein. Let the hypersurface S specified by

$$\varphi(x^\mu) = 0 \tag{2.2}$$

denote the surface of evolution of the initial wavefront; i.e., the initial wavefront is the spatial surface $\varphi(0,\vec{x}) = 0$. Assume that the field U is continuous across S; so only the normal derivative can be discontinuous. Choosing a local coordinate system denoted by $x^\mu = (\varphi, \psi^i)$, the "first-order discontinuity" in a given quantity f can be defined as

$$\delta_1 f \equiv \left[\frac{\partial f}{\partial \varphi}\right] \tag{2.3}$$

where

$$[X] \equiv X\big|_{\varphi=0^+} - X\big|_{\varphi=0^-} \equiv \delta_0 X \tag{2.4}$$

Then it is easy to check that

$$\delta_0(\partial_\mu U) = [\partial_\mu U] = (\partial_\mu \varphi)\, \delta_1 U \tag{2.5}$$

Here we are considering the possibility that $\delta_1 U$ is discontinuous. Taking "first-order discontinuity" is then like differentiation,

$$\delta_1 f(U) = (\nabla_U f)\delta_1 U \tag{2.6}$$

The generalization to quasilinear systems of higher order, say q, in derivatives is straightforward now. Define

$$\delta_r f \equiv \left[\frac{\partial^r f}{\partial \varphi^r}\right] \tag{2.7}$$

and consider the case that

$$\delta_q U \neq 0, \qquad \delta_r U = 0, \qquad 0 \leq r < q \tag{2.8}$$

Notice that "taking the discontinuity" depends on the order of derivative. For example, if f has a second-order discontinuity, i.e., $\delta_2 f \neq 0$, then $\delta_1 f = 0$, but $\delta_1(\partial_\mu f) \neq 0$. Hence in general one has

$$\delta_r\, \partial_\mu = (\partial_\mu \varphi)\delta_{r+1} \tag{2.9}$$

3. ANALYSIS OF THE CHARACTERISTICS

Taking the discontinuity of (2.1), we find that, on S,

$$(\mathscr{A}^\mu \varphi_\mu)\, \delta_1 U = 0 \tag{3.1}$$

where $\mathscr{A}^\mu = \mathscr{A}^\mu(\overline{U})$. [Here $\varphi_\mu \equiv \partial_\mu \varphi$; henceforth, we will drop the subscript

1 on δ and use δU to mean the first-order discontinuity in U, i.e., $\delta_1 U$.] Since $\delta U \neq 0$, we see that S must be a characteristic surface, i.e.,

$$\det(\mathcal{A}^\mu \varphi_\mu) = 0 \tag{3.2}$$

must hold on S. Thus δU is in the kernel of $\mathcal{A}^\mu \varphi_\mu$ for a given choice of root in (3.2).

We can assume that (2.1) can always be rewritten such that $\mathcal{A}^0$ is the identity matrix and also that we have a flat metric on spacetime. We next define the unit normal to S,

$$\hat{\mathbf{n}} = \frac{\vec{\nabla}\varphi}{|\vec{\nabla}\varphi|} \tag{3.3}$$

and the "characteristic eigenvalue"

$$\lambda = -\frac{\partial_0 \varphi}{|\vec{\nabla}\varphi|} \tag{3.4}$$

So for a given choice of root, δU is always a linear combination of the right eigenvectors of $\mathcal{A}_n \equiv \hat{\mathbf{n}} \cdot \vec{\mathcal{A}}$ for the corresponding eigenvalue λ. In the hyperbolic case, the set of eigenvalues $\lambda^{(I)}$ $(I = 1, \ldots, N)$ are distinct, and the corresponding right (left) eigenvectors R^I (L^I) are real and form a linearly independent set.

For general $\mathcal{A}^\mu$ ($\mathcal{A}^0$ not necessarily equal to the identity matrix), λ are just the roots of the characteristic equation (3.2) and we have

$$-L_I \mathcal{A}^0 \lambda^{(I)} + L_I \mathcal{A}_n = 0 = \mathcal{A}_n R_J - \lambda^{(J)} \mathcal{A}^0 R_J \tag{3.5}$$

$$-L_I \mathcal{A}^0 \lambda^{(I)} R_J + L_I \mathcal{A}_n R_J = 0 = L_I \mathcal{A}_n R_J - L_I \lambda^{(J)} \mathcal{A}^0 R_J \tag{3.6}$$

Then for the hyperbolic case, $L_I \mathcal{A}^0 R_J = 0$ $(I \neq J)$, and one can always choose to normalize such that

$$L_I \mathcal{A}^0 R_J = \delta^{IJ} \tag{3.7}$$

The characteristic equation (3.2) is homogeneous of order N in $p_\mu \equiv \varphi_\mu$. By analogy, we can write it as

$$H(x, p) = 0 \tag{3.8}$$

where we may introduce the explicit coefficients

$$H(x, p) = G^{\mu_1 \cdots \mu_N} p_{\mu_1} \cdots p_{\mu_N} \tag{3.9}$$

Then, by homogeneity of H,

$$\sum_{\mu} p_{\mu} \frac{\partial H}{\partial p_{\mu}} = NH = 0 \tag{3.10}$$

which can be written as

$$\vec{p} \cdot \vec{\nabla}_p H = 0 \tag{3.11}$$

where $\vec{p}$ is the D-dimensional vector with components $(p_0, \ldots, p_{(D-1)})$ and $\vec{\nabla}_p H$ is the vector with components $(\partial H/\partial p_0, \ldots, \partial H/\partial p_{(D-1)})$.

Since $\vec{p}$ is the normal to the hypersurface S, we see that the tangential vector is parallel to $\vec{\nabla}_p H$, or that the curves

$$\frac{dx^{\mu}}{ds} = \frac{\partial H}{\partial p_{\mu}} \tag{3.12}$$

are tangential on (3.8). Notice that this is a set of curves, one for each root of the characteristic equation. In analogy to classical mechanics, the "momenta" then satisfy

$$\frac{dp_{\mu}}{ds} = -\frac{\partial H}{\partial x^{\mu}} \tag{3.13}$$

on S. This follows from $dH/ds = 0$ for a tangential deformation and from the compatibility condition $\partial_{\mu} p_{\nu} - \partial_{\nu} p_{\mu} = 0$. We can eliminate s for $t = x^0$, and then H factors as

$$H = \prod_{I=1}^{N} (p_0 - h^I) \tag{3.14}$$

and p_0 can be fixed as one of the roots. Reparametrizing, for a given root $p_0 = h^{I_0}(U, p_i)$, we can span the characteristic surface with the trajectories obeying

$$\frac{dx^{\mu}}{ds} = \frac{\partial h^{I_0}}{\partial p_{\mu}}, \qquad \frac{dp_{\mu}}{ds} = -\frac{\partial h^{I_0}}{\partial x^{\mu}} \tag{3.15}$$

4. AN EXAMPLE

Consider the following $D = 2$ example [2]. Take the simple PDE

$$\partial_t u + u\, \partial_x u = 0 \tag{4.1}$$

Then it easily follows that the characteristic curve is

$$\frac{dx}{dt} = u(x, t) \tag{4.2}$$

Clearly (4.1) and (4.2) imply $du/dt = 0$ along the characteristic curve, i.e., the characteristic curve is a line with constant u. So the "velocity" dx/dt is constant and the characteristics are nothing but straight lines. If we denote by φ the point where the given line is at initial time $t = 0$, then by integrating (4.2),

$$x(t) = u(\varphi, 0)t + \varphi \tag{4.3}$$

This implicit equation for $\varphi = \varphi(x, t)$ is just the equation for the given characteristic curve, parametrized by its initial point. So then one has

$$u(x, t) = u(\varphi(x, t), 0) \tag{4.4}$$

for the solution to (4.1), in terms of the initial value of u at $t = 0$.

For a linear PDE, the coefficient of $\partial_x u$ in (4.1) is a constant independent of u and the characteristic curves are parallel straight lines. In the general case, when the coefficient of $\partial_x u$ in (4.1) is an arbitrary, say smooth, function of u, the slope of a given characteristic curve depends on the initial value of u at the starting point. Thus, as time evolves, the characteristic curves can intersect and a shock may develop. It seems that this can be prevented only if the "velocity" dx/dt can be made independent of the coordinate normal to the characteristic.

5. EXCEPTIONALITY

Even though we have neither found a proof in the literature nor been able to prove it rigorously, this "method of characteristics" seems to extend to first-order PDEs in higher dimensions. Although there does not seem to be such a construction for the matrix system (2.1), we want to carry on our discussion and see what can be done.

Motivated by this 2-dimensional example, let us look for situations where the characteristic surfaces do not cross as they evolve, hence shock waves do not develop. Following the reasoning given above, we can demand that (locally) the characteristic eigenvalue, which after all is analogous to the "velocity" in the given example, is independent of φ in the evolution, or that

$$\frac{\partial\lambda}{\partial\varphi} = 0 = (\nabla_U\lambda)\,\frac{\partial U}{\partial\varphi} \tag{5.1}$$

Now let us look at the homogeneous case in (2.1) (i.e., $\mathcal{B} = 0$). Then taking a particular root $p_0 = h^{(I_0)}$ of the characteristic equation (3.14) defines a family of surfaces (remember $p_\mu \equiv \partial_\mu\varphi^{(I_0)}$)

$$\varphi^{(I_0)}(0, \vec{x}) = \text{const} \tag{5.2}$$

Assume that U is just a function of φ. Then (2.1) implies (by $\mathcal{B} = 0$) that

$$(\mathcal{A}^{\mu}\varphi_{\mu})\frac{dU}{d\varphi} = 0 \tag{5.3}$$

and hence

$$\frac{dU}{d\varphi} = \chi^{(J)}(\varphi)R^{(J)} \tag{5.4}$$

where $R^{(J)}$ is the corresponding right eigenvector for the given root. Since U and $R^{(J)}$ are N-vectors, we end up with N ordinary differential equations. We can always assume that $R^{(J)}$ has a nonzero component, and that a particular component U_K can always be chosen such that it is equal to φ. Since the eigenvector $R^{(J)}$ is known as a function of U and φ, the ratios of the other components determine $U_A = U_A(U_K)$ $(A \neq K)$. The particular component U_K itself can be determined from the characteristic equation. Now a solution for U obtained in this way is called a *simple wave* [2].

This brings us to the so-called exceptionality condition [2]. Let us first define what is meant by that:

The wave corresponding to a given characteristic root is called *exceptional* if it is such that

$$(\nabla_U \lambda) \cdot R = 0 \tag{5.5}$$

Moreover, when all the N wave modes are exceptional, the system is said to be CE [2].

So for simple wave solutions of the homogeneous case just discussed, the exceptionality condition is just the statement that $\nabla_U \lambda$ is orthogonal to the corresponding right eigenvector. In light of (5.1) and the discussion that led to it, the exceptionality condition does after all seem to "justify" a generalization of the (naive) idea we developed to prevent the development of shocks using the $D = 2$ example (at least for the case of simple waves).

Another way of looking at the problem may be provided by the following:

From (3.1), it follows that δU can be expressed as a linear combination of right eigenvectors as

$$\delta U = \pi^I R_I \tag{5.6}$$

for some components π^I $(I = 1, \ldots, N)$ (also called the coefficients of discontinuity). In general, one would expect these coefficients to evolve according to a nonlinear differential equation. In Appendix A, it is shown how the CE condition can also be viewed as the statement that the coefficients of discontinuity evolve according to a linear ODE, and thus the characteristic curves are prevented from intersecting locally.

We now briefly mention another alternative approach developed in ref. 3 for a "covariant formulation" of exceptionality. Let us choose a particular root $p_0 = h^{(J)}(U, p_i)$ for some J in (3.14). We also have by (3.4) that p_0 is proportional to λ. If we now take the field gradient ∇_U of the "Hamiltonian" H in (3.14) and set $p_0 = h^{(J)}$ afterward, we see that only the term which is proportional to $\nabla_U p_0 \sim \nabla_U \lambda$ does not vanish in the resultant expression. For a simple wave, then, contracting this with δU and using (2.6), we get

$$\nabla_U \lambda \cdot \delta U = \delta\lambda \tag{5.7}$$

for this particular root.

Hence, in light of (5.1) and (5.5), one arrives at the "equivalent" condition for exceptionality:

The wave corresponding to a given root is exceptional if on the characteristic surface $H = 0$, one has

$$\delta\lambda = 0 \tag{5.8}$$

Again for CE, this must hold for all roots, or that

$$\delta H = 0 \tag{5.9}$$

In the following, we apply the above-mentioned (two seemingly different) CE requirements in a variety of physical cases. In the process we supply the missing details leading to the results reported in ref. 1. We want to make it clear that (5.5) was originally developed for systems in $D = 2$ only [2]. Here, however, we apply (5.5) and (5.9) to systems in higher dimensions. Although we are not aware of a rigorous construction that generalizes the results explained so far to PDEs in higher D, it is plausible that such a general proof can be given.

After all, notice that the characteristic equation and the condition for CE are algebraic equations which must hold pointwise in any x^μ. At a fixed point on the characteristic surface at a fixed time, the normal $\hat{\mathbf{n}}$ is a fixed vector and proceeding for arbitrary $\hat{\mathbf{n}}$, and U, is the same as imposing the conditions pointwise. Furthermore, the original system is rotationally invariant, where rotations act on U as some linear matrix representation. So, the CE conditions are rotation invariant, and having chosen $\hat{\mathbf{n}}$ (i.e., working at a fixed point and time), we can just rotate it to, say, the first coordinate direction x^1 and proceed to study the eigensystem $|\mathcal{A}^1 - \lambda\mathbf{I}| = 0$, provided the system can be brought into a form such that $\mathcal{A}^0$ equals the identity matrix. Of course U changes in rotating, but the eigensystem is worked out for arbitrary U.

6. SCALAR FIELD

We now want to study in detail the CE requirement for a scalar field in $D = 4$. We first work out the problem using the requirement (5.5), then show that one finds the same answer (with considerably less effort) using condition (5.9), as was in fact done earlier in ref. 3.

6.1. The First Way

Given the covariant action $I = \int d^4x\, L(z)$, where $z \equiv \frac{1}{2}(\partial_\mu \sigma)^2$ is the only invariant (in first derivatives), $\eta = (-, +, +, +)$, the field equations can be written as

$$\partial_\mu((\partial^\mu \sigma)L') = (\partial_\mu \partial^\nu \sigma)(\partial_\nu \sigma)(\partial^\mu \sigma)L'' + (\partial_\mu \partial^\mu \sigma)L' = 0 \tag{6.1}$$

Here prime denotes differentiation with respect to z.

By defining $A \equiv \partial_0 \sigma$, $B \equiv \partial_1 \sigma$, $C \equiv \partial_2 \sigma$, and $D \equiv \partial_3 \sigma$ [hence $z = \frac{1}{2}(-A^2 + B^2 + C^2 + D^2)$], we can take $\mathbf{U} = (A, B, C, D)$ and write this system in canonical form as

$$\mathbf{I}\,\frac{\partial \mathbf{U}}{\partial t} + \mathbf{M}^i\,\frac{\partial \mathbf{U}}{\partial x^i} = \mathbf{0}$$

where each $\mathbf{M}^i$ has elements (with $i = 1, 2, 3$; $\mu = 0, 1, 2, 3$)

$$m^1_{00} = \frac{-2ABL''}{\Theta}, \qquad m^1_{01} = \frac{B^2 L'' + L'}{\Theta}, \qquad m^1_{02} = \frac{BCL''}{\Theta}, \qquad m^1_{03} = \frac{BDL''}{\Theta}$$

$$m^1_{2\mu} = m^1_{3\mu} = 0; \qquad m^1_{1\mu} = -\delta_{0\mu}$$

$$m^2_{00} = \frac{-2ACL''}{\Theta}, \qquad m^2_{01} = m^1_{02}, \qquad m^2_{02} = \frac{C^2 L'' + L'}{\Theta}, \qquad m^2_{03} = \frac{CDL''}{\Theta}$$

$$m^2_{1\mu} = m^2_{3\mu} = 0; \qquad m^2_{2\mu} = -\delta_{0\mu}$$

$$m^3_{00} = \frac{-2ADL''}{\Theta}, \qquad m^3_{01} = m^1_{03}, \qquad m^3_{02} = m^2_{03}, \qquad m^3_{03} = \frac{D^2 L'' + L'}{\Theta}$$

$$m^3_{1\mu} = m^3_{2\mu} = 0; \qquad m^3_{3\mu} = -\delta_{0\mu}$$

and $\Theta \equiv A^2 L'' - L'$. Here we have also used the compatibility conditions

$$\frac{\partial B}{\partial t} = \frac{\partial A}{\partial x^1}, \qquad \frac{\partial C}{\partial t} = \frac{\partial A}{\partial x^2}, \qquad \frac{\partial D}{\partial t} = \frac{\partial A}{\partial x^3}$$

So, by the reasoning given at the end of the last section, we proceed to impose the CE condition (5.5) using the eigensystem $|\mathbf{M}^1 - \lambda\mathbf{I}| = 0$.[4] The characteristic polynomial of $\mathbf{M}^1$ turns out to be $\lambda^2(\lambda^2 + a_1\lambda + a_2) = 0$, where $a_1 \equiv 2ABL''/\Theta$ and $a_2 \equiv (B^2L'' + L')/\Theta$. Apart from the eigenvalue at $\lambda = 0$ (with multiplicity 2), there are two distinct eigenvalues λ_3, λ_4 in the general case.[5] The eigenvectors corresponding to each can be taken as

$$\mathbf{e}_1 = \left(0, \frac{-BCL''}{B^2L'' + L'}, 1, 0\right)^T, \qquad \mathbf{e}_3 = (-\lambda_3, 1, 0, 0)^T$$
$$\mathbf{e}_2 = \left(0, \frac{-BDL''}{B^2L'' + L'}, 0, 1\right)^T, \qquad \mathbf{e}_4 = (-\lambda_4, 1, 0, 0)^T$$

which clearly form a full linearly independent set, hence our system is hyperbolic. We next apply the CE condition (5.5) to this eigensystem. Obviously, it will be trivially satisfied for $\lambda = 0$. For the remaining nontrivial eigenvalues, note that by differentiating $\lambda^2 + a_1\lambda + a_2, = 0$, we can write

$$\frac{\partial\lambda}{\partial U_s} = -\frac{\lambda\partial a_1/\partial U_s + \partial a_2/\partial U_s}{2\lambda + a_1}$$

and the CE condition $\Sigma_s\ (\partial\lambda_p/\partial U_s)\ e_{p,s} = 0$ becomes

$$\lambda^2\frac{\partial a_1}{\partial A} + \lambda\left(\frac{\partial a_2}{\partial A} - \frac{\partial a_1}{\partial B}\right) - \frac{\partial a_2}{\partial B} = 0 \qquad \text{for} \quad \lambda = \lambda_3, \lambda_4$$

by using the explicit form of the eigenvectors. However, we know that λ_3, λ_4 satisfy $\lambda_2 + a_1\lambda + a_2 = 0$. Hence these two equations must be linearly dependent, which implies that

$$a_1\frac{\partial a_1}{\partial A} + \frac{\partial a_1}{\partial B} - \frac{\partial a_2}{\partial A} = 0 \qquad \text{and} \qquad a_2\frac{\partial a_1}{\partial A} + \frac{\partial a_2}{\partial B} = 0$$

have to be satisfied simultaneously.

Substituting the explicit forms of a_1 and a_2, we find after some calculation that

$$\begin{aligned} a_1\frac{\partial a_1}{\partial A} + \frac{\partial a_1}{\partial B} - \frac{\partial a_2}{\partial A} &= \frac{L'L''' - 3(L'')^2}{\Theta^3}\,[A(3B^2 + A^2)L' + A^3(B^2 - A^2)L''] \\ &= 0 \end{aligned} \tag{6.2}$$

[4] In fact, we showed separately that taking arbitrary $\hat{\mathbf{n}}$ does not alter the final results obtained in this section.

[5] For the degenerate case $L'\,[L' - (A^2 - B^2)L''] = 0$, $\lambda_3 = \lambda_4 = -A/B$, but then there is no nontrivial covariant action which can satisfy this. Moreover in this case, the system is no longer hyperbolic.

$$a_2 \frac{\partial a_1}{\partial A} + \frac{\partial a_2}{\partial B} = L'L''' - \frac{3(L'')^2}{\Theta^3}\,[B(3A^2 + B^2)L' + BA^2(B^2 - A^2)L'']$$
$$= 0 \tag{6.3}$$

The only nontrivial covariant condition we can impose such that these two constraints are satisfied simultaneously is

$$L'L''' - 3(L'')^2 = 0 \tag{6.4}$$

6.2. The Second Way

In this part, we want to impose (5.9) using the formalism developed starting in Section 2. We now look for a surface S across which the discontinuity in σ is second order. Thus with $\delta_2\sigma \equiv Q$, we have ($\sigma_\mu \equiv \partial_\mu\sigma$, $\varphi_\mu \equiv \partial_\mu\varphi$)

$$\delta\sigma_\mu = \varphi_\mu Q$$

Taking the discontinuity of (6.1) gives

$$\varphi_\mu((\delta L')\sigma^\mu + (\delta\sigma^\mu)L') = 0 \tag{6.5}$$

which, with $\delta z = \sigma^\mu(\delta\sigma_\mu) = \sigma^\mu\varphi_\mu Q$ and $\delta L' = L''\,\delta z$, becomes

$$Q(\mathcal{G}\,L' + (\sigma^\nu\varphi_\nu)^2 L'') = 0 \tag{6.6}$$

where $\mathcal{G} \equiv \varphi^\mu\varphi_\mu$. Comparing this to the previous discussion, we have $H(x, p) = G^{\mu\nu}p_\mu\, p_\nu = 0$ with ($p_\mu = \varphi_\mu$, $Q \neq 0$)

$$G^{\mu\nu} = \eta^{\mu\nu}L' + \sigma^\mu\sigma^\nu L'' \tag{6.7}$$

Imposing (5.9) (taking the discontinuity) gives

$$Q(3\mathcal{G}\,L'' + (\sigma^\lambda\varphi_\lambda)^2 L''')(\sigma^\mu\varphi_\mu) = 0 \tag{6.8}$$

and using $\mathcal{G}L' + (\sigma^\nu\varphi_\nu)^2 L'' = 0$ in (6.8) yields

$$Q\mathcal{G}\left(-\frac{L'L'''}{L''} + 3L''\right)(\sigma^\nu\varphi_\nu) = 0 \tag{6.9}$$

This again leaves us with the condition (6.4).

Notice that throughout, we have never used the fact that $D = 4$. This suggests that (6.4) is a D-invariant ($D \geq 2$) condition. For general D, using the requirement (5.5), one ends up with $\mathbf{M}^i$, which individually have $\lambda = 0$ (with multiplicities $D - 2$) and the remaining two nontrivial eigenvalues [corresponding to the pair of canonical variables (σ, π) for the only degree of freedom of the theory] with their corresponding eigenvectors yield (6.4) when inserted into (5.5).

To find the solutions of (6.4), we first note that by defining $X \equiv L'$, we can write it as $X^4\,(X'/X^3)' = 0$, which will be satisfied nontrivially provided $X' = 0$ or $(X'/X^3)' = 0$. Integrating these simple equations, we find $X = c_1$, $L = c_1 z + c_2$ or $(1/X^2)' = -2c_3$, $1/X_2 = -2c_3 z + c_4$, $L = \pm(1/c_3)\sqrt{-2c_3 z + c_4} + c_5$ for c_q $(q = 1, \ldots, 5)$ arbitrary integration constants. Choosing these constants suitably, we note the particularly interesting cases as

$$L = -z = -\tfrac{1}{2}(\partial_\mu \sigma)^2$$

and

$$L = 1 - \sqrt{1 + 2z} = 1 - \sqrt{1 + (\partial_\mu \sigma)^2} = 1 - \sqrt{-\det[\eta_{\mu\nu} + (\partial_\mu \sigma)(\partial_\nu \sigma)]}$$

which are the scalar analogs to Maxwell and Born–Infeld electrodynamics, respectively.

7. NONLINEAR ELECTRODYNAMICS IN $D = 4$

We now come to our most physically important example, the $D = 4$ Abelian gauge vector theories. Any gauge-invariant action, depending on $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ but not its derivatives, has the form

$$I[A_\mu] = \int d^4x\, L(\alpha,\beta), \qquad \alpha \equiv \frac{1}{2} F_{\mu\nu} F^{\mu\nu},$$
$$\beta \equiv \frac{1}{4} F_{\mu\nu} {*F}^{\mu\nu}, \qquad {*F}^{\mu\nu} \equiv \frac{1}{2} \epsilon^{\mu\nu\sigma\tau} F_{\sigma\tau} \tag{7.1}$$

Here subscripts on L mean differentiation with respect to the (only possible) invariants α or β and with our conventions $\epsilon^{0123} = +1$, $\eta_{\mu\nu} = (-, +, +, +)$, $\alpha = \mathbf{B}^2 - \mathbf{E}^2$, $\beta = -\mathbf{B} \cdot \mathbf{E}$ with $E^i \equiv F^{0i}$, and $B^i \equiv \frac{1}{2}\epsilon^{ijk} F_{jk}$.

We first drop the β dependence of L, show in detail how the CE condition (5.9) is applied to $L(\alpha)$, then reinclude β and carry out the CE condition (5.9) for full $L(\alpha, \beta)$. [Again we originally studied this problem using the requirement (5.5), which is quite laborious and tedious. We show in Appendix B the general outline of how (5.5) is carried out for $L(\alpha)$. We do not show how (5.5) is applied to the most general case, $L(\alpha, \beta)$, although in this case we were able to prove at least the sufficiency of (7.14) and (7.15) using (5.5).]

We look for a hypersurface S across which the discontinuity in A_μ is second order. Hence, with $\delta_2 A_\mu = \pi_\mu$ we have

$$\delta F^{\mu\nu} = \varphi^\mu \pi^\nu - \varphi^\nu \pi^\mu, \qquad \delta\, {*F}^{\mu\nu} = \epsilon^{\mu\nu\sigma\tau} \varphi_\sigma \pi_\tau \tag{7.2}$$

and

$$\delta\alpha = 2F^{\mu\nu}\varphi_\mu\pi_\nu, \qquad \delta\beta = {}^*F^{\mu\nu}\varphi_\mu\pi_\nu \tag{7.3}$$

7.1. $L(\alpha)$ Case

For $L = L(\alpha)$ only, the field equation is simply $\partial_\nu(F^{\mu\nu}L') = 0$ with the Bianchi identity $\partial_\nu {}^*F^{\mu\nu} = 0$. (Here, prime denotes differentiation with respect to α.)

Taking the discontinuity of the field equation, we find

$$-2U^\mu \mathscr{F} L'' + (\varphi_\nu\pi^\nu)\varphi^\mu L' - \mathscr{G}\pi^\mu L' = 0 \tag{7.4}$$

where we have used $U^\mu \equiv F^{\lambda\mu}\varphi_\lambda$, $\mathscr{G} \equiv \varphi^\mu\varphi_\mu$, and $\mathscr{F} \equiv F^{\lambda\sigma}\varphi_\lambda\pi_\sigma = U^\sigma\pi_\sigma$. (Taking the discontinuity of the Bianchi identity, one can see that it follows automatically.)

Now contracting (7.4) by $-U_\mu$ (and assuming $\mathscr{F} \neq 0$ for the general case), we find

$$H = 2uL'' + \mathscr{G}L' = 0 \tag{7.5}$$

where we have defined $u \equiv U^\mu U_\mu$. Now $\delta u = 2U^\mu\delta U_\mu = 2\mathscr{G}\mathscr{F}$ and $\delta\alpha = 2\mathscr{F}$. Hence imposing (5.9) gives

$$\delta H = \mathscr{F}(4uL''' + 6\mathscr{G}L'') = 0 \tag{7.6}$$

and substituting for u using (7.5), we end up with

$$\delta H = 2\mathscr{F}\mathscr{G}\left(3L'' - \frac{L'}{L''}L'''\right) = 0 \tag{7.7}$$

Hence we again find (6.4) in a new disguise, whose solutions we can immediately copy as $L(\alpha) = k + (d + c\alpha)^{1/2}$ (for arbitrary constants k, d, c) apart from Maxwell, $L = -\frac{1}{2}\alpha$ (or $L' =$ const).

We remark that in $D = 3$, where α is the only invariant, this is also the CE result, where one also has $\sqrt{1 + \alpha} = \sqrt{-\det[\eta_{\mu\nu} + F_{\mu\nu}]}$. In $D = 2$ there is of course no propagation for any $L(\alpha)$ and correspondingly no restrictions are imposed.

7.2. $L(\alpha,\beta)$ Case

We now want to study the full Lagrangian $L(\alpha, \beta)$. For this case, the field equation is $\partial_\nu(L_\alpha F^{\mu\nu} + \frac{1}{2}L_\beta\, {}^*F^{\mu\nu}) = 0$. Taking the discontinuity, we find

$$F^{\mu\nu}\varphi_\nu(\delta L_\alpha) + \varphi_\nu(\delta F^{\mu\nu})L_\alpha + \tfrac{1}{2}{}^*F^{\mu\nu}\delta_\nu(\delta L_\beta) = 0 \tag{7.8}$$

Using (7.2) and (7.3), this becomes

$$-U^{\mu}(2\mathcal{F}L_{\alpha\alpha} + \chi L_{\alpha\beta}) + \varphi_{\nu}(\varphi^{\mu}\pi^{\nu} - \varphi^{\nu}\pi^{\mu})L_{\alpha}$$
$$-\tfrac{1}{2}V^{\mu}(2\mathcal{F}L_{\alpha\beta} + \chi L_{\beta\beta}) = 0 \tag{7.9}$$

where we have used $V^{\mu} \equiv {}^{*}F^{\lambda\mu}\varphi_{\lambda}$ and $\chi \equiv {}^{*}F^{\lambda\sigma}\varphi_{\lambda}\pi_{\sigma} = V^{\sigma}\pi_{\sigma}$.

Now contracting (7.9) by $-U_{\mu}$ and then by $-V_{\mu}$, we get, respectively,

$$\mathcal{F}(2uL_{\alpha\alpha} + \mathcal{G}L_{\alpha} + \beta\mathcal{G}L_{\alpha\beta}) + \chi(uL_{\alpha\beta} + \tfrac{1}{2}\beta\mathcal{G}L_{\beta\beta}) = 0 \tag{7.10}$$

$$\mathcal{F}(2\beta\mathcal{G}L_{\alpha\alpha} + L_{\alpha\beta}(u - \alpha\mathcal{G})) + \chi(\beta\mathcal{G}L_{\alpha\beta} + \mathcal{G}L_{\alpha} + \tfrac{1}{2}L_{\beta\beta}(u - \alpha\mathcal{G})) = 0 \tag{7.11}$$

where we have made use of the identities $U^{\mu}V_{\mu} = \beta\mathcal{G}$ and $V^{\mu}V_{\mu} = u - \alpha\mathcal{G}$.

For this system to have nontrivial $\mathcal{F}$ and χ, the determinant of the 2×2 matrix that comes from writing (7.10) and (7.11) as $(\mathcal{F}\chi)\mathbf{M} = \mathbf{0}$ must vanish. Hence we have ($K \equiv L_{\alpha\alpha}L_{\beta\beta} - L_{\alpha\beta}^{2}$)

$$H = u^{2}K + u\mathcal{G}[2L_{\alpha}(L_{\alpha\alpha} + \tfrac{1}{4}L_{\beta\beta}) - \alpha K]$$
$$+ \mathcal{G}^{2}[L_{\alpha}(L_{\alpha} + 2\beta L_{\alpha\beta} - \tfrac{1}{2}\alpha L_{\beta\beta}) - \beta^{2}K] = 0 \tag{7.12}$$

Notice that for the discriminant, one gets

$$\frac{\Delta}{\mathcal{G}^{2}} = \frac{1}{4}[-L_{\alpha}(4L_{\alpha\alpha} - L_{\beta\beta}) + 2\alpha K]^{2} + 4[-L_{\alpha}L_{\alpha\beta} + \beta K]^{2} \tag{7.13}$$

For the case $\Delta = 0$, i.e., when

$$-L_{\alpha}(4L_{\alpha\alpha} - L_{\beta\beta}) + 2\alpha[L_{\alpha\alpha}L_{\beta\beta} - L_{\alpha\beta}^{2}] = 0 \tag{7.14}$$

$$-L_{\alpha}L_{\alpha\beta} + \beta[L_{\alpha\alpha}L_{\beta\beta} - L_{\alpha\beta}^{2}] = 0 \tag{7.15}$$

H takes the form $H = K(u - h)^{2} = 0$ and for $K \neq 0$, it follows that (5.9) is satisfied automatically. Hence any L that fulfills (7.14) and (7.15) is CE.

The differential constraints (7.14) and (7.15) were actually found a long time ago in different contexts [3–5]. Bialynicki-Birula [5] discovered these equations by studying the propagation of weak electromagnetic waves on a strong, constant field background. He showed that they were necessary for both polarizations of light to propagate according to the same dispersion law; he calls these the "no-birefringence" conditions. Plebański [4] studied the theory of small perturbations and their discontinuities in nonlinear electrodynamics, and, considering all possible cases for the form of the background field (e.g., null, algebraically general) and constraining the system with physical conditions such as causality along the way, proved the necessity and sufficiency of these differential constraints for the excitations of light to propagate according to a single characteristic equation, with coinciding characteristic surfaces. Boillat [3] found these conditions using equation (5.9) explained in this paper, and demanding that it be expressible as a complete square as explained in (7.12)–(7.15).

The work of Plebański involves an extensive study of characteristic surfaces, which is what the CE formulation is all about, in nonlinear electrodynamics, so it is not surprising that he finds (7.14) and (7.15) as the conditions to have coinciding characteristic surfaces; after all, that is also what Boillat gets using the CE viewpoint. Bialynicki-Birula effectively allows the discontinuities in terms of weak disturbances about a generic background. It is not surprising to see that having the same dispersion law for both polarizations implies having a single characteristic surface for the evolution of discontinuities. Apart from these historical details, we will call the two conditions (7.14) and (7.15) the "strong CE" conditions from now on because of this extra physical constraint that they impose on the system.

The solutions of (7.14) and (7.15) are important to define physically acceptable models of electrodynamics. It is clear that the Maxwell action, $I_{\text{Max}} = -\frac{1}{2}\int d^4x\,\alpha$, is indeed a solution, and it was realized in refs. 3–5 that another is the (once again popular) Born–Infeld action [6],

$$I_{\text{BI}} = \int d^4x\,(1 - \sqrt{-\det[\eta_{\mu\nu} + F_{\mu\nu}]}) = \int d^4x\,(1 - \sqrt{1 + \alpha - \beta^2}) \tag{7.16}$$

However, these are not the only solutions unless one further requires that they reduce to I_{Max} for weak fields. Otherwise there are additional solutions such as $L = \alpha/\text{gb}$. [As shown in ref. 1, without requiring the weak-field condition, imposing strong CE with duality invariance (a property shared by both of these theories) singles out Maxwell and Born–Infeld.]

Now we continue with the general case $K \neq 0$, $\Delta \neq 0$. For convenience, we define

$$P \equiv 2L_\alpha(L_{\alpha\alpha} + \tfrac{1}{4}L_{\beta\beta}) - \alpha K, \qquad R \equiv L_\alpha(\mathrm{o}L_\alpha + 2\beta L_{\alpha\beta} - \tfrac{1}{2}\alpha L_{\beta\beta}) - \beta^2 K$$

$$p \equiv 2L_{\alpha\alpha}, \qquad q \equiv L_\alpha + \beta L_{\alpha\beta}, \qquad r \equiv L_{\alpha\beta}, \qquad s \equiv \tfrac{1}{2}\beta L_{\beta\beta}$$

and rewrite H as

$$H = u^2K + u\mathscr{G}P + \mathscr{G}^2R = 0 \tag{7.17}$$

Now imposing (5.9), we find

$$\begin{aligned}\delta H = {} & u^2(2\mathscr{F}K_\alpha + \chi K_\beta) + u\mathscr{G}(4\mathscr{F}K + 2\mathscr{F}P_\alpha + \chi P_\beta) \\ & + \mathscr{G}^2(2\mathscr{F}P + 2\mathscr{F}R_\alpha + \chi R_\beta) = 0\end{aligned} \tag{7.18}$$

where we have used $\delta u = 2\mathscr{G}\mathscr{F}$, $\delta\alpha = 2\mathscr{F}$, and $\delta\beta = \chi$.

Now using $u^2 = -(\mathscr{G}/K)(uP + \mathscr{G}R)$ [from (7.17)] and $\chi = -(pu + \mathscr{G}q/ru + \mathscr{G}s)$ [from (7.10)], we find that (7.18) is simplified into a form $\delta H = u\mathscr{G}\zeta_1 + \mathscr{G}^2\zeta_2 = 0$. Since $K \neq 0$ and $\Delta \neq 0$, this implies that ζ_1 and ζ_2 must vanish simultaneously. [This can also be seen as the requirement that (7.17) and (7.18) be linearly independent.] Finally, one finds that the CE requirements [corresponding to (5.9)] are

$$2K_\alpha(rP^2 - sPK - rRK) + K_\beta(qPK - pP^2 + pRK) + 2P_\alpha(sK^2 - rPK)$$
$$+ KP_\beta(pP - qK) + 2R_\alpha(rK^2) - R_\beta(pK^2) - 2rPK^2 + 4sK^3 = 0 \qquad (7.19)$$

$$2K_\alpha(rPK - sRK) + K_\beta(qRK - pRP) - 2P_\alpha(rRK) + P\beta(pKR)$$
$$+ 2R_\alpha(sK^2) - R_\beta(qK^2) - 4rRK^2 + 2sPK^2 = 0 \qquad (7.20)$$

In Appendix C, we give these equations in terms of L and its derivatives only. Notice that these equations are quasilinear (linear in the third-order derivatives of L) just like (6.4). Being of third order, they are of course weaker than (7.14) and (7.15). Born–Infeld, of course, satisfies these equations, but we have been able to solve them neither in the general case nor for the more restricted situation when one also demands duality invariance. For the latter, one would expect to get two (or, with a bit of luck, only one) ordinary differential equations involving only α-derivatives when one substitutes for β-derivatives by using the duality invariance constraint [7] and its (α, β) derivatives recursively.

An application of CE, rather than strong CE, comes from theories involving the (neutral) scalar plus the Abelian vector field, where possible invariants are (α, β, z [$\equiv \frac{1}{2}(\partial_\mu\sigma)^2$], $y \equiv \frac{1}{2}(F_{\mu\nu}\sigma^\nu)^2$). For a Lagrangian $L(\alpha, \beta, z)$, the CE conditions further require $L_{z\alpha} = 0 = L_{z\beta}$, which reduce it to the noninteracting $L(\alpha, \beta) + L(z)$ form. Having the "fully Born–Infeld" form $\sqrt{-\det[\eta_{\mu\nu} + F_{\mu\nu} + \sigma_\mu \sigma_\nu]}$ in mind, one can consider $L(\alpha, \beta, y, z)$. It turns out, however, that there are no CE actions with nontrivial dependence on the other possible variable $y \equiv \frac{1}{2}(F_{\mu\nu}\sigma^\nu)^2$. Thus, CE alone separates the two systems and imposes the previously stated constraints on their forms.

8. GRAVITATIONAL MODELS

Finally, we turn to gravitation. For Einstein's gravity in vacuum, as well as the linearized theory, the gravitational waves are CE, the characteristic surfaces describing discontinuities being null (see, e.g., ref. 8). It can be shown that this result holds for any $D > 4$. (For $D = 3$, there is of course no propagation and no restrictions are imposed.) One can further look at pure gravitational actions of the form $\int d^4x\,(pR^2_{\mu\nu} - qR^2)\sqrt{-g}$ in $D = 4$ and $\int d^Dx\, f(R)\sqrt{-g}$ in $D > 3$ and show that the same conclusion remains unchanged.

To reduce these theories to a first-order system would be inconvenient, but is fortunately made unnecessary by a simple extension of the previous discussion. Clearly, if we rebuilt the original higher order equations from the set (2.1), we would simply have the situation that all the derivatives of the field are assumed continuous except the highest one.

We first sketch the Einstein case to establish notation. Considering a second-order discontinuity in the metric across some characteristic surface $\varphi = 0$, $\delta_2 g_{\mu\nu} = \pi_{\mu\nu}$, we have ($\varphi_\mu \equiv \partial_\mu \varphi$)

$$\delta_1 \Gamma^\lambda_{\mu\nu} = \tfrac{1}{2}(\varphi_\mu \pi^\lambda_\nu + \varphi_\nu \pi^\lambda_\mu - \varphi^\lambda \pi_{\mu\nu})$$

$$\delta_0 R_{\mu\nu} = \varphi_\lambda(\delta_1 \Gamma^\lambda{}_{\mu\nu}) - \varphi_\nu(\delta_1 \Gamma^\lambda{}_{\lambda\mu})$$

$$= \tfrac{1}{2}(\varphi_\mu\varphi_\lambda\pi^\lambda{}_\nu + \varphi_\nu\varphi_\lambda\pi^\lambda{}_\mu - \varphi_\mu\varphi_\nu\pi^\lambda{}_\lambda - \varphi_\lambda\varphi^\lambda\pi_{\mu\nu})$$

and

$$\delta_0 R = g^{\mu\nu}(\delta_0 R_{\mu\nu}) = \varphi^\mu\varphi^\nu\pi_{\mu\nu} - \varphi_\mu\varphi^\mu\pi^\nu{}_\nu$$

which implies

$$\delta_0 G_{\mu\nu} = \delta_0(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R) = \delta_0 R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}\delta_0 R = 0$$

$$\delta_0 G_{\mu\nu} = \tfrac{1}{2}[\varphi_\mu\varphi_\lambda\pi^\lambda{}_\nu + \varphi_\nu\varphi_\lambda\pi^\lambda{}_\mu - \varphi_\mu\varphi_\nu\pi^\lambda{}_\lambda - \varphi_\lambda\varphi^\lambda\pi_{\mu\nu}$$
$$- g_{\mu\nu}(\varphi^\sigma\varphi^\tau\pi_{\sigma\tau} - \varphi_\sigma\varphi^\sigma\pi^\tau{}_\tau)] = 0 \tag{8.1}$$

In the harmonic gauge $g^{\mu\nu}\,\Gamma^\sigma_{\mu\nu} = 0$, one finds that its first discontinuity implies

$$2\pi^{\mu\nu}\varphi_\mu - \pi^\mu{}_\mu\varphi^\nu = 0 \tag{8.2}$$

Multiplying this by $g_{\nu\sigma}\varphi_\tau + g_{\nu\tau}\varphi_\sigma$, one gets

$$\varphi_\mu\varphi_\lambda\pi^\lambda{}_\nu + \varphi_\nu\varphi_\lambda\pi^\lambda{}_\mu - \varphi_\mu\varphi_\nu\pi^\lambda{}_\lambda = 0 \tag{8.3}$$

whereas contracting by φ_ν, one finds

$$\varphi^\mu\varphi^\nu\pi_{\mu\nu} = \tfrac{1}{2}\varphi_\mu\varphi^\mu\pi^\mu{}_\nu \tag{8.4}$$

Using (8.3) and (8.4) in (8.1), one ends up with

$$\delta_0 G_{\mu\nu} = \tfrac{1}{2}(\varphi_\lambda\varphi^\lambda\pi^{\mu\nu} + \tfrac{1}{2}g_{\mu\nu}\varphi_\lambda\varphi^\lambda\pi^\sigma{}_\sigma) = 0$$

Hence taking the trace,

$$\delta_0 G^\mu_\mu = \frac{(D + 2)}{4}\,\varphi_\lambda\varphi^\lambda\pi^\sigma{}_\sigma = 0$$

The discontinuity in $g_{\mu\nu}$ is arbitrary, hence $\pi^\sigma{}_\sigma \neq 0$, which implies that $\varphi_\lambda\varphi^\lambda = 0$. This says that the characteristic surfaces are null: the discontinuities travel with the speed of light in all directions. The same holds for the linearized version of the theory as well, of course.

For generic quadratic Lagrangians $(pR_{\mu\nu}R^{\mu\nu} - qR^2)\sqrt{-g}$ in $D = 4$, using similar steps [writing the field equations, choosing harmonic gauge as before, and utilizing the identities (8.3), (8.4)], one finds that ($Q \equiv \varphi^\lambda \varphi_\lambda$, $\pi \equiv \pi^\lambda{}_\lambda$)

$$Q\left(\frac{1}{2}(p - 2q)\varphi_\mu\varphi_\nu\pi - \frac{p}{2}Q\pi_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\left(\frac{p}{2} - 2q\right)Q\pi\right) = 0 \tag{8.5}$$

Taking the trace, one gets $Q^2\pi(p - 3q) = 0$. (The choice $p = 3q$ corresponds to Weyl-tensor squared; the scalar degree of freedom is absent.) For $p = 3q$, (8.5) becomes

$$qQ(\tfrac{1}{2}\varphi_\mu\varphi_\nu\pi - \tfrac{3}{2}Q\pi_{\mu\nu} + \tfrac{1}{4}g_{\mu\nu}Q\pi) = 0$$

Since $\pi_{\mu\nu}$ is arbitrary, we see that again $Q = 0$, as in the Einstein case, so $Q = 0$ characterizes both Einstein and the quadratic action.

Finally, we consider the class of actions $\int d^D x\, f(R)\sqrt{-g}$ in $D \geq 4$, whose field equations are

$$E_{\mu\nu} \equiv R_{\mu\nu} f' - \tfrac{1}{2}g_{\mu\nu} f + (g_{\mu\nu}\nabla_\sigma\nabla^\sigma - \nabla_\mu\nabla_\nu) f' = 0$$

Hence the order of highest derivatives is four. Following similar steps by taking $\delta_4 g_{\mu\nu} = \pi_{\mu\nu}$, we find the same expressions for $\delta_3\Gamma^\lambda{}_{\mu\nu}$ and $\delta_2 R_{\mu\nu}$ as for $\delta_1\Gamma^\lambda{}_{\mu\nu}$ and $\delta_0 R_{\mu\nu}$ in the Einstein case. Using these, we get

$$\delta_0 E_{\mu\nu} = (Qg_{\mu\nu} - \varphi_\mu\varphi_\nu)(\varphi^\sigma\varphi^\tau\pi_{\sigma\tau} - Q\pi) f'' = 0$$

Going to harmonic gauge with identity (8.4) and taking the trace, one gets

$$\delta_0 E^\mu{}_\mu = \frac{1 - D}{2} Q^2\pi f'' = 0$$

Here, too, $Q = 0$ is the only solution, and so for a wide class of gravitational actions the propagation obeys the Einstein behavior as well. As is well known, these systems are variants of Brans–Dicke scalar-tensor theories, so their "good propagation" is not surprising.

ACKNOWLEDGMENTS

We thank S. Deser for suggesting this problem and collaborating on its applications. Ö.S. would also like to thank I. P. Ennes and H. Rhedin for useful discussions. This work was supported in part by the NSF under grant PHY-9315811.

APPENDIX A

In this appendix, we show how the CE requirement (5.5) causes the coefficients of discontinuity to evolve according to a linear ODE.

Consider the wavefront at the boundary of a region with smooth enough solution $\overline{U}$. The following derivation fills a gap in ref. 3 and generalizes ref. 9, where the evolution of discontinuities in first derivatives of the dependent variables is studied. Choose a root of the characteristic equation, $p_0 = h^{I_0}(U, p_i)$. Differentiating (2.1) with respect to φ, and contracting with the corresponding left eigenvector, we have ($A, B, C = 1, \ldots, N$)

$$\begin{aligned}(\partial_\varphi U_C) L^I_A (\nabla_{U_C} \mathcal{A}^\mu_{AB})(\partial_\mu U_B) &+ L^I_A \mathcal{A}^\mu_{AB} (\partial_\varphi\, \partial_\mu U_B) \\ &+ (\partial_\varphi U_C) L^I_A (\nabla_{U_C} \mathcal{B}_A) = 0 \qquad \text{(A.1)}\end{aligned}$$

Now we can take the discontinuity of this equation. We have higher derivative terms, but notice for the term in the middle that ($\varphi_\mu = \partial_\mu \varphi$)

$$\begin{aligned}\partial_\varphi\, \partial_\mu U_B = \varphi_\mu(\partial^2_\varphi U_B) &+ (\partial_\varphi\, \varphi_\mu)(\partial_\varphi U_B) + (\partial_\mu \psi^i)(\partial_{\psi^i} \partial_\varphi U_B) \\ &+ (\partial_\varphi\, \partial_\mu \psi^i)(\partial_{\psi^i} U_B) \qquad \text{(A.2)}\end{aligned}$$

The first term on the right-hand side of (A.2) vanishes when contracted with $\mathcal{A}^\mu$ against the left eigenvector. Thus, there is just one φ derivative (i.e., no ∂^2_φ pieces), and the discontinuity can be taken as before. We first compute the following to use for the first term in (A.1):

$$\begin{aligned}[(\partial_\varphi U_C)(\partial_\mu U_B)] = (\delta U_C)\, \varphi_\mu(\delta U_B) &+ (\delta U_C)(\partial_\mu \overline{U}_B) \\ &+ (\partial_\varphi \overline{U}_C)\, \varphi_\mu(\delta U_B) \qquad \text{(A.3)}\end{aligned}$$

Now using (5.6) and (A.3) and taking the discontinuity of (A.1), we find

$$\begin{aligned}L^I_A \mathcal{A}^\mu_{AB}(\partial_\mu \psi^i)(\partial_{\psi^i}\, \delta U_B) &+ m^I_J \pi^J \\ &+ \varphi_\mu(\delta U_C) L^I_A (\nabla_{U_C} \mathcal{A}^\mu_{AB})(\delta U_B) = 0 \qquad \text{(A.4)}\end{aligned}$$

[Here the first term comes from the third term in (A.2), the last term comes from the first term in (A.3), and we have collected as "m" the coefficients of terms linear in π without derivatives. "m^I_J" are determined by the background solution as well as the extrinsic geometry of the characteristic surface.] Let us examine the other terms in (A.4).

The first term in (A.4) is (up to a redefinition of the coefficient matrix m)

$$(L^I_A \mathcal{A}^\mu_{AB} R^J_B)(\partial_\mu \psi^i)(\partial_{\psi^i} \pi_J) = (L^I_A \mathcal{A}^\mu_{AB} R^J_B)(\partial_\mu \pi_J) \qquad \text{(A.5)}$$

By taking the p_i derivative (i.e., applying ∂_{p_i}) of the straightforward equation $L^I_A \mathcal{A}^\mu_{AB} R^J_B p_\mu = 0$ and using (3.7), one finds

$$L^I_A \mathscr{A}^i_{AB} R^J_B = -\delta^{IJ} \frac{\partial h^{I_0}}{\partial p_i} \tag{A.6}$$

Hence using the equations for the trajectories (3.15), the first term in (A.4) reduces to $d\pi/ds$, where

$$\frac{d}{ds} = \frac{\partial}{\partial t} - \frac{dx^i}{ds} \frac{\partial}{\partial x^i}$$

For the last term in (A.4), we have (by making use of $L^I_A \mathscr{A}^\mu_{AB} R^J_B p_\mu = 0$)

$$\varphi_\mu (\delta U_C) L^I_A (\nabla_{U_C} \mathscr{A}^\mu_{AB})(\delta U_B) = -(\delta U_C) L^I_A \mathscr{A}^\mu_{AB} (\delta U_B)(\nabla_{U_C} \varphi_\mu)$$

Notice that the last factor has U dependence via the characteristic root p_0. Hence using (3.4) and (3.7) [with (5.6)],

$$\begin{aligned} \varphi_\mu (\delta U_C) L^I_A (\nabla_{U_C} \mathscr{A}^\mu_{AB})(\delta U_B) &= (\delta U_C) \delta^{IJ} \pi_J |\vec{\nabla} \varphi| (\nabla_{U_C} \lambda) \\ &= |\vec{\nabla} \varphi| \pi^I \pi_J R^J_C (\nabla_{U_C} \lambda) \end{aligned} \tag{A.7}$$

Finally, then, we have a nonlinear equation for the evolution of the coefficients of discontinuity along rays,

$$\frac{d\pi^I}{ds} + m^I_J \pi^J + |\vec{\nabla} \varphi| \pi^I \pi_J R^J_C (\nabla_{U_C} \lambda) = 0 \tag{A.8}$$

This is computable because all "U's" above are actually "$\overline{U}$'s."

Thus, we recognize that CE condition can also be viewed as the statement that the coefficients of discontinuity evolve according to a linear ODE.

APPENDIX B

In this appendix, we show the general outline of how (5.5) is carried out for models of electrodynamics that depend only on the Maxwell invariant, i.e., $L = L(\alpha)$.

By taking $\mathbf{U} = (\mathbf{E}, \mathbf{B})$ and looking only at the spatial components of the field equation $\partial_\nu (F^{\mu\nu} L') = 0$ and the Bianchi identity $\partial_\nu {}^*F^{\mu\nu} = 0$ (i.e., setting $\mu = i$), we can write this system in the form $\mathbf{H}^\mu\, \partial \mathbf{U}/\partial x^\mu = 0$, where $\mathbf{H}^\mu$ are 6×6 matrices. For this new system [as in the scalar field case when we had $2(= 4 - 2 \cdot 1)$ nontrivial eigenvalues corresponding to the pair of canonical variables for the only degree of freedom of the theory] we expect to get $\lambda = 0$ eigenvalue with multiplicity 2 ($= 6 - 2 \cdot 2$) for each individual $\mathbf{H}^i$ because of the two degrees of freedom.

Just as was done in the scalar field case, we only take $\mathbf{H}^1$ to start with. Hence we have $\mathbf{H}^0\, \partial\mathbf{U}/\partial t + \mathbf{H}^1\, \partial\mathbf{U}/\partial x^1 = \mathbf{0}$, where

$$\mathbf{H}^0 = \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \qquad \text{and} \qquad \mathbf{H}^1 = \begin{pmatrix} \mathbf{S} & \mathbf{R} \\ \sigma & \mathbf{0} \end{pmatrix}$$

which have elements (with $i, j = 1, 2, 3$)

$$\begin{aligned} p_{ij} &= 2E_iE_jL'' - \delta_{ij}L' \\ q_{ij} &= -2E_iB_jL'' \\ s_{ij} &= 2\epsilon_{1ik}E_jB_kL'' \\ r_{ij} &= -\epsilon_{1ik}(2B_jB_kL'' + \delta_{jk}L') \\ \sigma_{ij} &= -\epsilon_{1ik} \end{aligned}$$

and $\mathbf{I}$ is the 3×3 identity matrix.

Multiplying by

$$(\mathbf{H}^0)^{-1} = \begin{pmatrix} \mathbf{P}^{-1} & -\mathbf{P}^{-1}\mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

we bring this system into the canonical form $\mathbf{I}\, \partial\mathbf{U}/\partial t + \mathbf{W}\, \partial\mathbf{U}/\partial x^1 = \mathbf{0}$, where

$$\mathbf{W} = (\mathbf{H}^0)^{-1}\mathbf{H}^1 = \begin{pmatrix} \mathbf{P}^{-1}(\mathbf{S} - \mathbf{Q}\sigma) & \mathbf{P}^{-1}\mathbf{R} \\ \sigma & \mathbf{0} \end{pmatrix}$$

Then the characteristic polynomial of $\mathbf{W}$ turns out to be, just as predicted, of the form $\lambda^2(\lambda^4 + c_3\lambda^3 + c_2\lambda^2 + c_1\lambda + c_0) = 0$. The eigenvectors corresponding to each λ_s can be taken to be

$$\mathbf{e}_s = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$$

where

$$\lambda_1 = 0: \qquad \mathbf{a}_1 = \mathbf{0}, \qquad \mathbf{b}_1 = \begin{pmatrix} 1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$\text{with} \quad \begin{pmatrix} y_2 \\ y_3 \end{pmatrix} = \frac{-1}{r_{22}r_{33} - r_{23}r_{32}} \begin{pmatrix} r_{33} & -r_{23} \\ -r_{32} & r_{22} \end{pmatrix} \begin{pmatrix} r_{21} \\ r_{31} \end{pmatrix}$$

$$\lambda_2 = 0: \qquad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{b}_2 = \begin{pmatrix} 0 \\ z_2 \\ z_3 \end{pmatrix}$$

$$\text{with} \quad \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} = \frac{-1}{r_{22}r_{33} - r_{23}r_{32}} \begin{pmatrix} r_{33} & -r_{23} \\ -r_{32} & r_{22} \end{pmatrix} \begin{pmatrix} s_{21} \\ s_{31} \end{pmatrix}$$

and for $\lambda_s \neq 0$ ($s = 3, 4, 5, 6$)

$$\mathbf{a}_s = \begin{pmatrix} 0 \\ \lambda_s(\rho_{22} - \lambda_s\gamma_{23}) \\ \lambda_s(\rho_{23} + \lambda_s(-\lambda_s + \gamma_{22})) \end{pmatrix},$$

$$\mathbf{b}_s = \frac{1}{\lambda_s}\sigma\mathbf{a}_s = \begin{pmatrix} 0 \\ -(\rho_{23} + \lambda_s(-\lambda_s + \gamma_{22})) \\ (\rho_{22} - \lambda_s\gamma_{23}) \end{pmatrix}$$

where $\lambda_{ij} \equiv [\mathbf{P}^{-1}(\mathbf{S} - \mathbf{Q}\sigma)]_{ij}$ and $\rho_{ij} \equiv [\mathbf{P}^{-1}\,\mathbf{R}]_{ij}$.

Clearly these eigenvectors form a linearly independent set. By differentiating $\lambda^4 + c_3\lambda^3 + c_2\lambda^2 + c_1\lambda + c_0 = 0$, we get

$$\frac{\partial\lambda_p}{\partial U_s} = -\frac{(\lambda_p)^3\,\partial c_3/\partial U_s + (\lambda_p)^2\,\partial c_2/\partial U_s + \lambda_p\,\partial c_1/\partial U_s + \partial c_0/\partial U_s}{4(\lambda_p)^3 + 3c_3(\lambda_p)^2 + 2c_2\lambda_p + c_1}$$

Substituting this into the CE condition (5.5) $\Sigma_s\,(\partial\lambda_p/\partial U_s)e_{p,s} = 0$ gives a polynomial of order 6 in λ, but by using $\lambda^4 + c_3\lambda^3 + c_2\lambda^2 + c_1\lambda + c_0 = 0$ repeatedly, one can reduce this to a polynomial of order 3, whose coefficients must be set equal to zero simultaneously.

Doing so, we find that the only nontrivial covariant condition we can impose such that these coefficients vanish simultaneously is

$$L'L''' - 3(L'')^2 = 0 \tag{B.1}$$

APPENDIX C

Here, for completeness, we present (7.19) and (7.20) in terms of L and its derivatives only. They become ($K \equiv L_{\alpha\alpha}L_{\beta\beta} - L_{\alpha\beta}^2$)

$$\begin{aligned} &\tfrac{3}{2}L_\alpha L_{\alpha\beta\beta}(L_\alpha(16L_{\alpha\alpha}^3L_{\alpha\beta} + 8L_{\alpha\alpha}L_{\alpha\beta}^3 + L_{\alpha\beta}^3L_{\beta\beta}) \\ &\quad - K[8\alpha L_{\alpha\alpha}^2L_{\alpha\beta} + \beta(8L_{\alpha\alpha}L_{\alpha\beta}^2 + 4L_{\alpha\alpha}^2L_{\beta\beta} + L_{\alpha\beta}^2L_{\beta\beta})]) \end{aligned}$$

$$
\begin{aligned}
&+\tfrac{1}{2}L_\alpha L_{\alpha\alpha\alpha}(L_\alpha L_{\alpha\beta}(16L_{\alpha\alpha}L_{\alpha\beta}^2 + 8L_{\alpha\beta}^2 L_{\beta\beta} + L_{\beta\beta}^3)\\
&- K[8\alpha L_{\alpha\beta}^3 + \beta L_{\beta\beta}(12L_{\alpha\beta}^2 + L_{\beta\beta}^2)])\\
&- \tfrac{3}{2}L_\alpha L_{\alpha\beta}L_{\alpha\alpha\beta}(L_\alpha L_{\alpha\beta}(16L_{\alpha\alpha}^2 + 4L_{\alpha\beta}^2 + 4L_{\alpha\alpha}L_{\beta\beta} + L_{\beta\beta}^2)\\
&- K[8\alpha L_{\alpha\alpha}L_{\alpha\beta} + \beta(4L_{\alpha\beta}^2 + 8L_{\alpha\alpha}L_{\beta\beta} + L_{\beta\beta}^2)])\\
&- \tfrac{1}{2}L_\alpha L_{\beta\beta\beta}(L_\alpha(16L_{\alpha\alpha}^4 + 12L_{\alpha\alpha}^2 L_{\alpha\beta}^2 + L_{\alpha\beta}^4 - 4L_{\alpha\alpha}^3 L_{\beta\beta})\\
&- K[8\alpha L_{\alpha\alpha}^3 + \beta L_{\alpha\beta}(12L_{\alpha\alpha}^2 + L_{\alpha\beta}^2)])\\
&- \tfrac{3}{2}(4L_{\alpha\alpha} + L_{\beta\beta})K^2[L_\alpha L_{\alpha\beta} - \beta K] = 0 \qquad \text{(D.1)}
\end{aligned}
$$

and

$$
\begin{aligned}
&-\tfrac{3}{2}L_\alpha L_{\alpha\alpha\beta}((4L_{\alpha\alpha} + L_{\beta\beta})(2L_\alpha^2 L_{\alpha\beta}^2 - \alpha L_\alpha L_{\alpha\beta}^2 L_{\beta\beta})\\
&\quad + \beta L_\alpha L_{\alpha\beta}(16L_{\alpha\alpha}L_{\alpha\beta}^2 + 6L_{\alpha\beta}^2 L_{\beta\beta} - 2L_{\alpha\alpha}L_{\beta\beta}^2)\\
&\quad - \beta K[-\alpha L_{\alpha\beta}L_{\beta\beta}^2 + 2\beta(4L_{\alpha\alpha}L_{\alpha\beta}^2 + 2L_{\alpha\beta}^2 L_{\beta\beta} + L_{\alpha\alpha}L_{\beta\beta}^2)])\\
&\quad + \tfrac{3}{2}L_\alpha L_{\alpha\beta\beta}((4L_{\alpha\alpha}^2 + L_{\alpha\beta}^2)(2L_\alpha^2 L_{\alpha\beta} - \alpha L_\alpha L_{\alpha\beta}L_{\beta\beta})\\
&\quad - \beta K L_{\alpha\beta}[-\alpha L_{\alpha\beta}L_{\beta\beta} + 2\beta(4L_{\alpha\alpha}^2 + L_{\alpha\beta}^2 + 2L_{\alpha\alpha}L_{\beta\beta})]\\
&\quad + 2\beta L_\alpha(8L_{\alpha\alpha}^2 L_{\alpha\beta}^2 + 2L_{\alpha\beta}^4 + L_{\alpha\alpha}L_{\beta\beta}L_{\alpha\beta}^2 - L_{\alpha\alpha}^2 L_{\beta\beta}^2))\\
&\quad + \tfrac{1}{2}L_\alpha L_{\alpha\alpha\alpha}((4L_{\alpha\beta}^2 + L_{\beta\beta}^2)(2L_\alpha^2 L_{\alpha\beta} - \alpha L_\alpha L_{\alpha\beta}L_{\beta\beta})\\
&\quad + \beta L_\alpha(16L_{\alpha\beta}^4 + 6L_{\alpha\beta}^2 L_{\beta\beta}^2 - 2L_{\alpha\alpha}L_{\beta\beta}^3)\\
&\quad - \beta K(8\beta L_{\alpha\beta}^3 + 6\beta L_{\alpha\beta}L_{\beta\beta}^2 - \alpha L_{\beta\beta}^3))\\
&\quad - \tfrac{1}{2}L_\alpha L_{\beta\beta\beta}(2L_\alpha^2 L_{\alpha\alpha}(4L_{\alpha\alpha}^2 + 2L_{\alpha\beta}^2 - L_{\alpha\alpha}L_{\beta\beta})\\
&\quad + 2\beta L_\alpha L_{\alpha\alpha}L_{\beta\beta}(8L_{\alpha\alpha}^2 + 5L_{\alpha\beta}^2 - 3L_{\alpha\alpha}L_{\beta\beta})\\
&\quad - \beta K(8\beta L_{\alpha\alpha}^3 + 6\beta L_{\alpha\alpha}L_{\alpha\beta}^2 - \alpha L_{\alpha\beta}^3) - \alpha L_\alpha(L_{\alpha\beta}^4 + 4L_{\alpha\alpha}^3 L_{\beta\beta}))\\
&\quad - \tfrac{3}{2}K^2(4L_\alpha + 4\beta L_{\alpha\beta} - \alpha L_{\beta\beta})(L_\alpha L_{\alpha\beta} - \beta K) = 0 \qquad \text{(D.2)}
\end{aligned}
$$

respectively.

REFERENCES

1. S. Deser, J. McCarthy, and Ö. Sarioğlu, hep-th/9809153, to appear in *Class. Quant. Grav.* (1999).
2. P. D. Lax, *Ann. Math. Studies* **33** (1954) 211; *Comm. Pure Appl. Math.* **10** (1957) 537; A. Jeffrey and T. Taniuti, *Nonlinear Wave Propagation* (Academic Press, New York, 1964).

3. G. Boillat, *J. Math. Phys.* **11** (1970) 941; in *Recent Mathematical Methods in Nonlinear Wave Propagation* (Springer, 1996).
4. J. Plebański, *Lectures in Non-Linear Electrodynamics* (NORDITA, 1970).
5. I. Bialynicki-Birula, *Nonlinear electrodynamics: Variations on a theme by Born and Infeld*, in *Quantum Theory of Particles and Fields*, B. Jancewicz and J. Lukierski, eds. (World Scientific, 1983), p. 31.
6. M. Born and L. Infeld, *Proc. R. Soc. A* **144** (1934) 425.
7. G. W. Gibbons and D. A. Rasheed, *Nucl. Phys. B* **454** (1995) 185; M. Perry and J. H. Schwarz, *Nucl. Phys. B* **489** (1997) 47; S. Deser and Ö. Sarioğlu, *Phys. Lett. B* **423** (1998) 369.
8. A. Lichnerowicz, *Théories Relativistes de la Gravitation et de L'électromagnétisme* (Masson Paris, 1955).
9. E. Varley and E. Cumberbatch, *J. Inst. Math. Appl.* **1** (1965) 101.