Regular hyperbolicity, dominant energy condition and causality for Lagrangian theories of maps

Willie Wai-Yeung Wong

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, UK
E-mail: ww278@dpmms.cam.ac.uk

Received 16 May 2011, in final form 23 August 2011
Published 22 September 2011
Online at stacks.iop.org/CQG/28/215008

Abstract
The goal of this paper is threefold. First is to clarify the connection between the dominant energy condition and hyperbolicity properties of Lagrangian field theories. Second is to provide further analysis on the breakdown of hyperbolicity for the Skyrme model, sharpening the results of Crutchfield and Bell and comparing against a result of Gibbons, and provide a local well-posedness result for the dynamical problem in the Skyrme model. Third is to provide a short summary of the framework of regular hyperbolicity of Christodoulou for the relativity community. In the process, a general theorem about dominant energy conditions for Lagrangian theories of maps is proved, as well as several results concerning hyperbolicity of those maps.

PACS numbers: 04.40.−b, 03.50.−z, 02.30.Jr

1. Introduction

In the study of classical field theories, an oft-imposed ‘physical assumption’ is the dominant energy condition, which requires that the Einstein–Hilbert stress-energy tensor $T$ appearing on the ‘wood’ side [16] of Einstein’s equation to have the property that, for any future causal vectors $X$ and $Y$, the contraction $T(X, Y)$ is non-negative. A posteriori this condition seems reasonable, in view of the results that can be derived from it. Two of the most prominent examples are, of course, the singularity theorem of Penrose [48] and the positive mass theorem [52, 71]. However, these results are purely results in Lorentzian geometry: that is, Einstein’s equation is wholly unnecessary, except to transfer the dominant energy condition from the ‘wood’ side to the ‘marble’ side of the equation. In other words, those theorems could equally well have been stated without reference to general relativity, but merely with certain positivity conditions about the Einstein tensor $G = \text{Ric} - \frac{1}{2} \text{Rg}$. 
The typical a priori justification for the imposition of the dominant energy condition is some heuristic requiring that the flow of energy be at a speed less than that of gravity\(^1\), that is, a property on the classical field variously known as finite speed of propagation, domain of dependence or causality (see e.g. [70]). Perhaps the most well known in this regard is a theorem of Hawking [23] which makes precise this notion: that if a matter field satisfies the dominant energy condition and the energy is strongly coercive (that is, \(T(X, Y) = 0\) for non-zero vectors \(X, Y\) only when the matter field vanishes), then if the matter field vanishes in some spacetime region, it must also vanish in its domain of dependence. Typically, a proof of the domain of dependence property uses an energy estimate, which is the curved spacetime version of the law of energy conservation on Minkowski spacetime. And therefore, it is usually expected that a domain of dependence property comes hand in hand with the hyperbolicity of the associated matter evolution. (Hyperbolicity in this context is synonymous with well-posedness in the sense of Hadamard [21] for the equations of evolution. Physically, this corresponds to requiring not only that the dynamical evolution of the classical fields can be predicted once an initial state is specified for the system, but also that imprecisions in the initial data can be propagated in a controlled fashion: this latter property is necessary to guarantee the robustness of numerical simulations, and to allow predictions to be made from finite-precision human measurements. See section 4 for more details on the mathematical descriptions of hyperbolicity.)

This expectation is, however, not always the case. The principal obstruction is that Hawking’s theorem only guarantees the finite speed of propagation of ‘perturbations of vacuum’. That is, it essentially only guarantees that the edge of vacuum cannot recede faster than the speed of gravity. For semilinear field theories where the kinematics in the high frequency limit is always governed by the spacetime metric, there can be no distinction between perturbations of vacuum and perturbations of a given solution. And hence the argument used by Hawking to establish his domain of dependence theorem generally indicates hyperbolicity for the associated matter field. For quasilinear field theories, however, the strong self-interaction means that the kinematics close to a vacuum background can be significantly different from that around a ‘large’ solution. And for these types of theories, the dominant energy condition is not sufficient to guarantee hyperbolicity. One should consider this in terms of the linearized equations of motion around a given solution, which is what governs the propagation of errors and imprecisions. For semilinear field theories, the linearized equations always share the same principal part with the original equations. For quasilinear theories, subtle algebraic differences between the principal parts of the linearized and the original equations can allow the energy perturbation to flow with arbitrarily large and possibly imaginary speeds, leading to linear instability. See also the discussion at the end of section 4.5.

A prime example of this difficulty is illustrated by two results related to the Skyrme model of nuclear physics. Motivated by unstable numerical simulations, Crutchfield and Bell demonstrated [14] that highly boosted background solutions to the Skyrme model are linearly unstable under perturbations (that there exist exponentially growing modes with rates of growth proportional to the frequency of the perturbations; that is, there can be modes growing exponentially at arbitrarily large rates). On the other hand, it was shown by Gibbons [20] that the Skyrme model in fact enjoys the dominant energy condition. In the current paper, a more detailed analysis of the breakdown of hyperbolicity for the Skyrme model will be presented. The analysis is based on Christodoulou’s regular hyperbolicity framework [10].

(One may also ask whether the reverse implication is true: that hyperbolic matter models will always enjoy a dominant energy condition. A simple example showing its falsehood is

\(^1\) If one is willing to assume a linear or semilinear theory of electromagnetism, then also the speed of light.
the linear wave equation with a negative potential \( \Box u = -u_{tt} + \Delta u = Vu \), where the potential \( V \leq 0 \). If one were to desire theories without external potentials, one can also consider the focusing nonlinear wave equation \( \Box u = -|u|^p u \), the hyperbolicity, or local well-posedness, of which is well known (see e.g. [55]). In regions where the modulus \( |u| \) is large while approximately constant, the associated Einstein–Hilbert stress–energy tensor for either of the above examples violates the dominant energy condition.)

In this paper, the connections between the dominant energy condition and hyperbolicity will be studied in the context of Lagrangian theory of maps. An emphasis will be placed on the insufficiency for mutual implication. The reasons for the focus on such matter models are twofold. First is its general applicability. Many models of mathematical physics can be cast in the framework of Lagrangian theory of maps. Starting from the simple linear wave equation, which can be regarded as a map from Minkowski space to the real line or to the complex plane, we can modify the target space to a general Riemannian manifold and obtain what is called the nonlinear \( \sigma \)-model in the physics literature, or the wave-map system in mathematics. This system is itself interesting as models in high-energy physics (see [29] and references therein) or as symmetry reductions from Einstein’s equations in general relativity (see, e.g. [4, 8, 46]).

As a semilinear modification to the standard wave equation with a geometric interpretation, the well-posedness properties (both global and local, and in both small and large data regimes) have been well studied, see [35, 49, 50, 65–69] for some recent progress on global regularity and singularity formation in the large data regime, and [6, 7, 32–34, 54, 56, 64] for a sample of classical results in this area.

In this paper, we will focus on \textit{quasilinear} modifications to the wave-map system, whose dynamics is comparatively less well studied. Such generalizations also have wide physical applications, with examples in the nonlinear \( \sigma \)-model hierarchy including the Skyrme model [45, 57, 58], the membrane equation [24, 41], a Born–Infeld-type model in cosmology [28, 38] and models of hydro- and elasto-dynamics [1, 3, 5, 9, 31, 51, 59, 62]. The models of dynamics in a continuous medium are particularly interesting in this context, as generally a physical assumption in such models is that the particle world-lines are time-like, a condition necessary to guarantee the causality of the matter model. It will be shown in this paper that one can construct examples of equations of states for which the dominant energy condition is satisfied \textit{independently} of whether the physical constraint is imposed. This further reinforces the idea that the domain of dependence theorem of Hawking is only a statement about vacuum perturbations.

The second reason for considering Lagrangian theories of maps is purely technical. For Lagrangian field theories, the Einstein–Hilbert stress energy can be defined via a variational procedure on the Lagrangian density. This allows for general and efficient calculations to check the dominant energy condition. Furthermore, in the context of Lagrangian theories of maps, the regular hyperbolicity framework of Christodoulou [10] provides a powerful while algebraically simple characterization of local well-posedness. Therefore, we will consider only such matter models for ease of discussion.

The paper is organized as follows. In section 2, we review the Lagrangian field theory of maps and give some examples that have appeared in the literature. In section 3, a geometric method is described for computing the Einstein–Hilbert stress–energy tensor for a large class of maps which includes the physically interesting models described above. The method provides an easy way to verify the dominant energy condition for these maps; we recover the result of Gibbons [20] as a special case. In section 4, we briefly describe the philosophy and method of regular hyperbolicity of Christodoulou [10], and recall the notion of canonical stress. Here, only the basic ideas behind the theory of regular hyperbolicity will be sketched, the focus being on its application. In a forthcoming paper with Speck [61], a detailed gentle
introduction to regular hyperbolicity will be given, along with some simple extensions that were alluded to, without proof, in Christodoulou’s monograph. And in section 5, we apply the theory to the problem of hyperbolicity of the Skyrme model.

2. Lagrangian theory of maps

Throughout we let \((M, g)\) be an \((m + 1)\)-dimensional Lorentzian manifold, where sign convention is taken to be \((- , + , + , \ldots)\), and we let \((N, h)\) be an \(n\)-dimensional Riemannian manifold. \(M\) represents the physical spacetime (often taken to be 3 space and 1 time dimensions, though we make no such restrictions here), while \(N\) represents the internal structure of the field. In applications, for nonlinear \(\sigma\)-models, \(N\) is usually taken to be a Lie group or a symmetric space; in dynamics of a continuous medium, \(N\) is the material manifold\(^2\) with \(n = m\).

Denote by \(\phi : M \to N\) a continuously differentiable map. In field theories \(\phi\) gives the state of the field at a point in spacetime, whereas for continuum mechanics \(\phi\) represents the coordinate transformation between the Eulerian and Lagrangian pictures. Then, the action of \(\phi\) can be used to pull back the metric \(h\) onto \(M\) as a positive semi-definite quadratic form on \(TM\), which we write as

\[
\phi^*h(X, Y) = h(d\phi \cdot X, d\phi \cdot Y),
\]

where the left-hand side is evaluated at a point \(p \in M\) and the right-hand side at the point \(\phi(p) \in N\) for \(X, Y \in T_pM\).

We define the \((1,1)\)-tensor field \(D \phi\) by composing with the inverse metric \(g^{-1}\):

\[
D \phi = g^{-1} \circ \phi^* h.
\]

We will follow Manton and Sutcliffe [45] and call this the \textit{strain tensor} for the map, the nomenclature taken from the study of dynamics in a continuous medium, where, roughly speaking, the trace of \(D \phi\) on a space-like hypersurface describes the local deformation of the material. See [1, 31, 62] for more detailed discussions (note that our definition here agrees with that of Tahvildar–Zadeh [62] if we impose the physical assumption that the map \(\phi\) admits a space-like simultaneous space (in other words, any nonzero vector \(X\) in the kernel of \(d\phi\) has \(g(X, X) < 0\), but our definition differs from that of Kijowski and Magli [31], which in addition to the above physical assumption, also breaks the degeneracy by adding in the square of the particle velocity to make the strain tensor positive definite).

At a fixed point \(p \in M\), the tensor \(D \phi\) defines a linear transformation of the tangent space \(T_pM\). We can consider its eigenvalues. In the case that \(g\) is a Riemannian metric, \(D \phi\) is a self-adjoint operator on \(T_pM\) relative to the (positive definite) inner-product given by the metric, and hence all the eigenvalues are real. For the Lorentzian case, the eigenvalues are in general complex. Denote by \(\{\lambda_1, \ldots, \lambda_k\}\) the non-zero eigenvalues, counted with multiplicity. One easily sees that

\[
k \leq \text{rank}(d\phi) \leq \min(m + 1, n).
\]

Recall the elementary symmetric polynomials \(\sigma_j(\{\lambda_1, \ldots, \lambda_k\})\) given by

\[
\sigma_j(\{\lambda_1, \ldots, \lambda_k\}) = \sum_{1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_j \leq k} \prod_{i=1}^{j} \lambda_{\alpha_i}
\]

\(^2\) In the fluids literature, the material manifold is sometimes only required to be equipped with a volume form; but as every paracompact smooth manifold admits a Riemannian metric, and a Riemannian metric can realize any volume form through conformal rescaling, it is of no loss in generality to assume the material manifold is Riemannian.
with $\sigma_0 = 1$ and $\sigma_j = 0$ for all $j > k$. The $\sigma_j$ correspond to the coefficients of the characteristic polynomial for $D^g$, with $\sigma_1 = \text{tr}(D^g)$ and $\sigma_{m+1} = \det(D^g)$, and are algebraic invariants of the tensor field $D^g$. Furthermore, since $D^g$ is a real linear transformation, the values of the $\sigma_j(D^g)$ are all real. By an abuse of notation, we will write $\sigma_j(D^g)$ when we mean the symmetric polynomials on the eigenvalues of $D^g$.

In this paper, we will only consider Lagrangian field theories for maps $\phi : (M, g) \to (N, h)$ where the action integral is given by

$$S = \int_M L(s(\phi), \sigma_1(D^g), \sigma_2(D^g), \ldots, \sigma_{m+1}(D^g)) \, d\text{vol}_g,$$

where $s : N \to \mathbb{R}$ is a non-negative scalar function on the internal space $N$. In particular, we require the only dependence of the Lagrangian on the field $\phi$ be through the algebraic invariants $\sigma_j(D^g)$ and the value of $\phi$ itself. (Note that by (2) any $\sigma_j$ not listed above as an argument for $L$ is not dynamical.) Physically, the dependence on $\sigma_j(D^g)$ and not other components of $D^g$ corresponds to the assumption that the laws of physics are locally invariant under a Lorentzian rotation in $O(1, m)$ of the tangent space $T_{\phi}M$ that fixes the kernel of $D^g$. In relativistic elasticity where a space-like simultaneous space is imposed, the relevant subgroup of $O(1, m)$ to consider is the orthogonal group $O(m)$, and this condition on the action corresponds to the assumption that the material is homogeneous, isotropic and perfectly elastic [62]. The function $s$ corresponds to the entropy per particle in relativistic elasticity, and plays the role of the symmetry-breaking mass term in the Skyrme model.

Observe that since the action itself only depends on $D^g$, which is defined through only the first derivative of the map $\phi$, the equations of motion given by applying the Euler–Lagrange equations to (4) will be a system of quasilinear second-order partial differential equations.

Some explicit examples of Lagrangian field theories from the literature that fall in this class include the following.

Wave maps can be described by the Lagrangian function

$$L = \sigma_1(D^g) = |\nabla \phi|^2,$$

where the norm is taken relative to both $g$ and $h$ (in index notation, we may write $L = g^{ij} h_{AB} \partial_i \phi^A \partial_j \phi^B$).

Relativistic elasticity typically makes the additional assumption that $d\phi$ is onto, and its kernel time-like (the space-like simultaneous space assumption). (By the rank-nullity theorem, this requires that $m = n$.) Then, interpreting $s$ as the entropy per particle, action (4) is the general form for studying a homogeneous, isotropic and perfectly elastic solid. The case where $L = L(s, \sigma_n)$ is the special case representing relativistic fluids, $\sigma_n = |\phi^* d\text{vol}_h|^2$ being equal to the squared $g$-norm of the pullback of the volume form on $N$.

Skyrme model classically assumes $(M, g)$ to be the four-dimensional Minkowski space with $(N, h)$ the standard 3-sphere, though we can also consider the model without such constraints. The Lagrangian function is (up to rescaling) [57, 58] (see also [43, 44])

$$L = \sigma_1(D^g) + \sigma_2(D^g) + s(\phi),$$

where $\sigma_2(D^g) = |d\phi \wedge d\phi|^2$, and $s(\phi) = m \text{dist}_h(\phi, \phi_0)^2$ is a multiple of the geodesic distance of $\phi$ from a fixed point $\phi_0 \in N$, the constant $m$ being the mass parameter. In general, Lagrangians of the form $L = \sigma_1 + \ldots$ are the nonlinear $\sigma$-models.

Born–Infeld-type models fix a large constant $b > 0$. Restricting the allowed maps $\phi$ to those whose eigenvalues of $D^\phi$ have real parts at least $-b$, the action is given by

$$L = \sqrt{\det(b \cdot I d + D^\phi)} - \sqrt{\det(b \cdot I d)}.$$

This is the $\sigma$-model analog of the Maxwell–Born–Infeld model of nonlinear electrodynamics, see also [20].
Membranes as considered in [24, 41] can be described by setting \((M, g)\) to Minkowski space of some fixed dimension, and \((N, h)\) to the real line with canonical metric, and taking
\[
L = \sqrt{1 + \sigma_1(D^\phi)}
\] (8)
while restricting consideration to those \(\phi\) for which \(L\) is well defined. It is a special case of the Born–Infeld-type models above, since \(\det(Id + D^\phi) = 1 + \sigma_1(D^\phi)\) when \(D^\phi\) has rank at most 1. It can also be viewed as an analog to the minimal surface equation for embedding into Minkowski space.

3. Dominant energy condition

In this section, we describe some sufficient (but far from necessary) conditions on the Lagrangian function \(L\) of the form considered in the previous section that guarantees that the associated Einstein–Hilbert stress–energy tensor satisfies the dominant energy condition. In isolated cases (fluids and elasticity with the space-like simultaneous surface assumption, or exactly the classical Skyrme and Born–Infeld models [20]) the results are known before. The novel contribution in this paper is given below in proposition 5. It can be used in a unified algebraic framework applicable to all theories of maps described by actions of the form (4), making unnecessary the ad hoc computations through explicitly evaluating the eigenvalues used in e.g. [20]. It is worth remarking that those types of computations rely on a genericity argument to diagonalize a positive semidefinite quadratic form relative to a Minkowski metric, a procedure which cannot be carried out when \(D^\phi\) admits a null eigenvector. By working geometrically and tensorially, the computations described below avoid this difficulty.

We start by reviewing some definitions. Recall that the Einstein–Hilbert stress–energy tensor \(T \in \Gamma(T^2_M)\) for a Lagrangian field theory is given formally by a variational derivative for the Lagrangian density relative to the inverse metric:
\[
T \sqrt{|\det g|} := \frac{\delta[L \sqrt{|\det g|}]}{\delta g^{-1}} = \left(\frac{\delta L}{\delta g^{-1}} - \frac{1}{2} Lg\right) \sqrt{|\det g|}. \tag{9}
\]

**Definition 1.** The stress–energy tensor \(T\) is said to obey the dominant energy condition at a point \(p \in M\) if \(\forall X \in T_pM\) such that \(g(X, X) < 0\); the following two conditions are satisfied:
\[
T(X, X) > 0 \quad (10a)
\]
\[
[T \circ g^{-1} \circ T](X, X) \leq 0 \quad (10b)
\]

unless \(T\) vanishes identically.

**Remark 2.** The definition is equivalent to the classical statements (see e.g. section 4.3 in [23] or chapter 9 of [70]) of the dominant energy condition. Observe that (10b) gives that the vector \(g^{-1} \circ T \circ X\) is a causal vector for any time-like vector \(X\), and (10a) gives that the vector \(g^{-1} \circ T \circ X\) has opposite time orientation as the time-like vector \(X\).

The set of future-pointing time-like vectors form a convex cone; hence, we have the following technical lemma, applicable to all Lagrangian field theories, not just those described in section 2.

**Lemma 3.** Let \(F = F(x_1, \ldots, x_k)\) be a continuously differentiable function of \(k\) real variables. Assume that \(F\) is concave, \(F(0) \geq 0\) and that \(\partial_i F \geq 0\) for each \(1 \leq i \leq k\). Let \(L_i, 1 \leq i \leq k\), denote a collection of Lagrangian functions, and let \(T_i\) denote their corresponding stress–energy tensors. Suppose \(T_i\) each separately obeys the dominant energy condition, or,
equivalently, the vectors \( Y_i = g^{-1} \circ T_i \circ X \) are all past-causal for any fixed future time-like \( X \). Then, the stress–energy tensor \( T \) for the Lagrangian formed by \( L = F(L_1, \ldots, L_k) \) also obeys the dominant energy condition.

**Proof.** The stress–energy tensor \( T \) can be written, using (9), as

\[
T = \sum_{i=1}^{k} \partial_i F \cdot \delta L_i - \frac{1}{2} F g = \sum_{i=1}^{k} \partial_i F \cdot T_i - \frac{1}{2} \left( F - \sum_{i=1}^{k} \delta_i F \cdot L_i \right) g.
\]

Now considering \( g^{-1} \circ T \circ X \), the first term in the above expression contributes \( \sum \partial_i F \cdot Y_i \). By assumption, this is a convex combination of past-causal vectors, and hence is past-causal. For the second term, since \( g^{-1} \circ g \circ X = X \), to show that it is also past-causal it suffices to show that

\[
F \geq \sum_{i=1}^{k} \delta_i F \cdot L_i.
\]

But this follows from the fact that \( F \) is concave and \( F(0) \geq 0 \). \( \square \)

**Remark 4.** The fact that \( F \) is required to have nonnegative partial derivatives represents the fact that each of the \( L_i \)'s contributes nonnegatively to the energy. The fact that \( F(0) \geq 0 \) states that there is no negative vacuum energy. Both conditions are therefore natural and necessary for the total Lagrangian \( L \) to have positive energy density, if the \( L_i \)'s are taken to be independent. The concavity condition is technical. It appears naturally in the proof, but can potentially be relaxed if more is assumed on the individual \( L_i \)'s.

We will apply lemma 3 to the following proposition, which is the main computational result of this section. Observe also that for \( L = s(\phi) \), its corresponding stress–energy tensor is \( T = -\frac{1}{2} s \cdot g \), and by the assumption on the positivity of \( s \), obeys the dominant energy condition.

**Proposition 5.** For \( L = \sigma_j(D\phi), T \) obeys the dominant energy condition. Furthermore, \( T = 0 \) at a point \( p \) if and only if \( j > \text{rank}(d\phi_p) \).

Before giving the proof, we need to review some linear algebra. Consider a real vector space \( V \). Let \( A \) be a linear transformation on \( V \). Then, \( A \) naturally extends to a linear transformation, which we denote \( A^\sharp_j \), on \( A^\sharp_j(V) \), the space of alternating \( j \)-vectors over \( V \). A classical result in linear algebra is that \( \sigma_j(A) \) is proportional to \( \text{tr}_{A^\sharp_j(V)} A^\sharp_j \). Now, letting \( V = T_p M \) and \( A = D\phi = g^{-1} \circ \phi^* h \), we observe that

\[
(D\phi)^\sharp_j = (g^{-1})^\sharp_j \circ \phi^* (h^\sharp_j),
\]

or, to put it in words, \((D\phi)^\sharp_j \) is obtained from first taking the induced metric \( h^\sharp_j \) on alternating \( j \)-vectors in \( T_{\phi_p} V \), pulling it back via \( \phi \), and composing it with the induced metric \((g^{-1})^\sharp_j \) for the alternating \( j \)-forms. In index notation, this can be written as

\[
[(D\phi)^\sharp_j]_{a_1 \cdots a_j}^{b_1 \cdots b_j} = g^{b_1 c_1} \cdots g^{b_j c_j} (\phi^* h)_{a_1 c_1} (\phi^* h)_{a_2 c_2} \cdots (\phi^* h)_{a_j c_j},
\]

where the bracket notation in the indices denotes full anti-symmetrization of the \( \{c_1, \ldots, c_j\} \) indices. For a Lagrangian proportional to \( \sigma_j \), we can assume

\[
L = g^{a_1 c_1} \cdots g^{a_j c_j} (\phi^* h)_{a_1 c_1} \cdots (\phi^* h)_{a_j c_j}.
\]

It is simple to check, using \((D\phi) = \text{diag}(-1, 1, 1, \ldots)\), that the above expression has the correct sign: that \( L \) defined thus is a positive multiple of \( \sigma_j \).
One can also arrive at (11) purely from a linear algebra point of view. Let \( p_j \) be the power sum

\[ p_j(\{\lambda_1, \ldots, \lambda_k\}) = \sum_{i=1}^{k} \lambda_i^j. \]

Recall Newton’s identity

\[ j \cdot \sigma_j = \sum_{i=1}^{j} (-1)^{j-i} \sigma_{j-i} p_i \]

which allows us to express \( \sigma_j \) as a rational polynomial in \( p_i \)’s. Now, by definition, it is clear

that

\[ p_j(D\phi) = \text{tr}[(D\phi)^j], \]

where \((D\phi)^j\) is the \( j \)-fold composition of \( D\phi \). Then, it is easy to check, for some \( E \),

\[ \sigma_j = g^{a_1 b_1} \cdots g^{a_j b_j} E^{i_1 \cdots i_j}(\phi^* h)_{a_1 a_1} \cdots (\phi^* h)_{a_j a_j}. \]

Newton’s identity reduces to a generating condition for \( E \) based on the Kronecker \( \delta \) symbols:

\[ E_{a_1 \cdots a_j b_1 \cdots b_j} = \sum_{i=1}^{j} (-1)^{j-i} E^{i_1 \cdots i_j}(\phi^* h)_{a_1 a_1} \cdots (\phi^* h)_{a_j a_j} \cdot \delta_{i_1 \cdots i_j}. \]

A direct computation which we omit here shows that then in fact the invariant \( E_{a_1 \cdots a_j b_1 \cdots b_j} \) is

a positive rational multiple of the generalized Kronecker symbol \( \delta^{i_1 \cdots i_j}_{a_1 \cdots a_j} \), from which we

recover (11).

**Proof of proposition 5.** We need to show that \( g^{-1} \circ T \circ X \) is past-causal for any future

time-like vector \( X \). Since \( T \) is tensorial, we can assume that \( X \) has unit length. Using the

expression for (a positive scalar multiple of) \( \sigma_j \) given in (11), we can write \( T(X, \cdot) \) for \( L = \sigma_j \)
in index notation:

\[ T_{ab} X^b = j X^{b_1} \cdots g^{a_i b_i} \cdots (\phi^* h)_{a_j a_j} X^b = \frac{1}{2} g_{ab} X^b L. \]

(12)

Take an orthonormal basis for \( T_p M \) relative to \( g \). Since we assumed \( X \) unit, let \( e_0 = X \)

and \( \{ e_i \}_{1 \leq i \leq m} \) all space-like. We can take \( j \leq m+1 \) as otherwise \( T \) is identically 0. Then, we

note that a basis for \( \Lambda^j(T_p M) \) is given by

\[ \{ e_0 \wedge e_{a_1} \wedge \cdots \wedge e_{a_j} \} \cup \{ e_{a_1} \wedge \cdots \wedge e_{a_j} \} \}

\[ \{ e_{a_1} \wedge \cdots \wedge e_{a_j} \} \}

We write the first set as \( \Lambda^j_{\perp} \) and the second set as \( \Lambda^j_{\parallel} \). Using the normalization that

\( v \wedge w = v \otimes w - w \otimes v \), we find that each of the elements in \( \Lambda^j_{\perp} \) has norm \( -j! \) while

the elements in \( \Lambda^j_{\parallel} \) have norm \( j! \).

To show that \( T(X, X) > 0 \), we observe that under expansion (12), the first term corresponds to

\[ \sum_{\omega \in \Lambda^j_{\perp}} \phi^* (h^{\omega_j})(\omega, \omega), \]

while the second term corresponds to

\[ \frac{1}{2} \left( - \sum_{\omega \in \Lambda^j_{\perp}} \phi^* (h^{\omega_j})(\omega, \omega) + \sum_{\omega \in \Lambda^j_{\parallel}} \phi^* (h^{\omega_j})(\omega, \omega) \right). \]
So summing them gives

\[ \frac{1}{2} \left( \sum_{\omega \in \Lambda_j^1} \phi^*(h^{ij})(\omega, \omega) + \sum_{\omega \in \Lambda_j^1} \phi^*(h^{ij})(\omega, \omega) \right) \]

which is non-negative by the fact that \( \phi^*(h^{ij}) \) is a positive semi-definite quadratic form on \( \Lambda_j^i(T_p M) \). Furthermore, observe that since \( \Lambda_j^1 \cup \Lambda_j^1 \) is a basis, its push-forward \( \phi^*(\Lambda_j^1 \cup \Lambda_j^1) \) spans \( \Lambda_j^i(T_p M) \subset \Lambda_j^i(T_{\phi(p)} N) \). Thus, by the fact that \( h \) and hence the induced metric \( h^{ij} \) is positive definite, we conclude that when \( T(X, X) = 0 \), necessarily \( \Lambda_j^i(\phi^* T_p M) = \{0\} \).

This proves the assertion that \( T \) vanishes only when \( j > \text{rank}(d\phi) \).

To show (10b), we observe that

\[ X^a T_{ab} \xi^c d^{-1} T_{db} X^b = -T(X, X)^2 + \sum_{i=1}^m T(X, e_i)^2. \]

The first thing to note is that \( T(X, e_i) \) does not have any contribution from the second term in (12). For the first term, a quick computation shows that \( T(X, e_i) \) corresponds to

\[ \sum_{\eta \in \Lambda_j^{i-1}} \phi^*(h^{ij})(e_0 \wedge \eta, e_i \wedge \eta) \]

so

\[ \sum_{i=1}^m T(X, e_i)^2 \leq \left( \sum |T(X, e_i)| \right)^2 \]

\[ \leq \left( \sum_{i=1}^m \sum_{\eta \in \Lambda_j^{i-1}, e_i \wedge \eta \neq 0} |\phi^*(h^{ij})(e_0 \wedge \eta, e_i \wedge \eta)| \right)^2 \]

\[ \leq \frac{1}{4} \left( \sum_{\eta \in \Lambda_j^{i-1}} \phi^*(h^{ij})(e_0 \wedge \eta, e_0 \wedge \eta) + \sum_{i=1}^m \phi^*(h^{ij})(e_i \wedge \eta, e_i \wedge \eta) \right)^2 \]

\[ = \frac{1}{4} \left( \sum_{\eta \in \Lambda_j^{i-1}} \sum_{i=0}^m \phi^*(h^{ij})(e_i \wedge \eta, e_i \wedge \eta) \right)^2 \]

and therefore (10b) is satisfied.

As an immediate application of lemma 3 and proposition 5, we have that the Skyrme model and the Born–Infeld model described in section 2 obey the dominant energy condition: it suffices to check that \( F_{\text{Skyrme}}(s, \sigma_1, \sigma_2) = s + \sigma_1 + \sigma_2 \) and \( F_{\text{BI}}(\sigma_1, \ldots, \sigma_{m+1}) = \sqrt{\sum_{j=1}^{m+1} b^{m+1-j} \sigma_j} - \sqrt{b^{m+1}} \) are concave, satisfy \( F(0) \geq 0 \) and have positive partial derivatives, conditions which are easily seen to hold. Therefore, we recover the result of [20] in dimension \( m = n = 3 \), and also extend it to arbitrary dimensions \( m, n \). In fact, we have

**Theorem 6.** Given a Lagrangian theory of maps with action given by (4), a sufficient condition for its Einstein–Hilbert stress–energy tensor to obey the dominant energy condition is that the Lagrangian function \( L \) in (4) be continuously differentiable with nonnegative partial derivatives on its arguments, be concave and satisfy \( L(0) \geq 0 \).
Remark 7. We stress again here that our notion of Lagrangian theory of maps, as discussed in section 2, is very general, and includes the general forms of Lagrangians used in relativistic elasticity and relativistic fluids. Therefore, it is remarkable that for these theories of maps, the assumption of a time-like particle world-line is inconsequential insofar as the dominant energy condition is concerned: in particular a fluid model admitting tachyonic particles can still satisfy \((10a)\) and \((10b)\), in direct contradiction to the intuition often presented as the justification for the dominant energy condition. For a concrete example, consider, for a fixed constant \(b\), the fluid Lagrangian (as a special case of the models of relativistic elasticity described at the end of section 2)

\[
L = \sqrt{b + \sigma_3(D\phi)}.
\]

Here, we assume that \(M = \mathbb{R}^{1+3}\) is the standard Minkowski space and \(N = \mathbb{R}^3\) is the usual Euclidean space. The map \(\phi\) gives the translation of the Eulerian coordinates (coordinates in spacetime) to the Lagrangian coordinates (coordinates of the material manifold or fluid parcel). In other words, making the assumption that \(d\phi\) is surjective, the integral curves of the kernel of \(d\phi\) correspond to particle world-lines. By theorem 6, this model obeys the dominant energy condition; locally, however, a fluid moving with a constant velocity larger than that of gravity is a solution, e.g. \(\phi : \mathbb{R}^{1+3} \rightarrow \mathbb{R}^3\) where \(\phi(t, x, y, z) = (t, x, y)\) is a solution to the equations of motion, while its particle world-lines are curves of constant \((t, x, y)\), which are manifestly space-like. As we will see in section 5 for the Skyrme model, however, there is an instability (i.e. a violation of hyperbolicity) associated with these tachyonic matter fields.

4. Regular hyperbolicity and the canonical stress

The question of hyperbolicity for a system of partial differential equations is generally synonymous with whether the system admits a locally well-posed Cauchy problem for all smooth initial data. In the context of linear systems of constant coefficients, it is a well-known theorem ([25], chapter 5) that a necessary and sufficient condition for hyperbolicity of a system \(P(\partial)\phi = 0\) is the hyperbolicity for its polynomial symbol \(P(\xi)\). (The fact that the Cauchy problem is well posed for all smooth initial data is crucial for the necessity. See remark 10 below.) Two difficulties arise when trying to apply this theorem to nonlinear or variable coefficient systems. First is the general difficulty of checking the hyperbolicity for a symbol, which requires computing the singular locus of \(P(\xi) - 1\). Second is the technical difficulty that concrete quantitative estimates are required for iteration arguments used in these types of problems.

Many stronger notions of hyperbolicity exist in the mathematical literature (see [12, 37] for some samplers); the common theme to all is that they provide sufficient conditions for local well-posedness of the non-characteristic Cauchy problem. In general, however, these conditions are not necessary: the fact that Leray hyperbolicity [39] is a subset of symmetric hyperbolicity [18] is well known; the fact that the Maxwell–Born–Infeld model can only be seen as symmetric hyperbolic after an augmentation procedure [2, 53] strongly suggests that the notion of symmetric hyperbolicity does not directly capture all hyperbolic systems. (It is perhaps interesting to note that the hyperbolicity of the Maxwell–Born–Infeld system can be directly treated within the regular hyperbolicity framework [60].) In this section, we will describe the regular hyperbolicity framework, which provides another sufficient condition for local well-posedness\(^3\). To guarantee local well-posedness, the general technique common to these methods is that of the \(L^2\) energy estimate.

\(^3\) The fact that this framework is not better known in the community is perhaps due to the dense mathematical language in Christodoulou's monograph [10]. The author hopes to provide here, by way of a summary of the main
4.1. Energy estimates

The energy estimates that we will need are estimates of the following form: assume for the current exposition that $M$ is diffeomorphic to $\mathbb{R}^{1+m}$ which allows us to fix a coordinate system $(t, x)$ with $x \in \mathbb{R}^m$, and let $H^k$ denote the $L^2$ Sobolev space with $k$ derivatives on constant $t$ slices which are diffeomorphic to $\mathbb{R}^m$. The energy estimates are inequalities

$$
\| \phi \|_{H^k(t)} \leq \| \phi \|_{H^k(0)} e^{Ct} \quad \forall 0 < t < T, C = C(T, \| \phi \|_{H^k(0)}).
$$

(The fact that we are only interested in such $L^2$ energy estimates is because we look for estimates generally applicable to hyperbolic equations. In particular, we require the estimates to hold for the linear wave equation. By a theorem of Littman [42], inequalities of the form (13) for the wave equation can only hold in $L^2$-based function spaces.) Inequalities of the form (13) guarantee local existence up to time $T$.

To obtain such energy estimates, the usual method is via the divergence theorem. Let $J = J(t, x, \phi, \ldots, \partial^k \phi)$ be a vector field, depending on up to $k$ derivatives of $\phi$ which we assume solves some partial differential equation. One formulation of the divergence theorem states that

$$
-dt \circ J(0, x) dx + \int dt \circ J(t, x) dx = \int_0^t \int \text{div}(J) \, dx \, dt,
$$

where $dt \circ J$ is the scalar obtained by pairing the vector field $J$ with the coordinate one-form $dt$. Inequality (13) would hold after one application of Gronwall’s lemma, if we can guarantee the following:

$$
C^{-1}\| \phi \|_{H^k(t)} \leq \int dt \circ J(t, x) dx \leq C\| \phi \|_{H^k(t)} \quad \text{for some } C \geq 1, \quad (14a)
$$

$$
\int \text{div}(J)(t) \, dx \leq C'\| \phi \|_{H^k(t)} \quad \text{for some } C' > 0. \quad (14b)
$$

A $J$ verifying (14a) and (14b) will be called a compatible energy current. (In [10], Christodoulou uses the term ‘compatible current’ to refer to any $J$ verifying (14b). The use of the word ‘energy’ here is meant to reflect the imposition of the additional positivity condition (14a).)

For (14b) to hold, it is necessary that we apply the equation: by the chain rule, the divergence of the vector field $J$ which depends on $k$ derivatives of $\phi$ will depend on $k + 1$ derivatives of $\phi$; it will be absurd to be able to control its integral by something depending on fewer derivatives. But by suitably applying the equations of motion for $\phi$, we can convert top-order derivatives to lower-order ones and satisfy the inequality.

**Remark 8.** Observe that in the above discussion we do not require the Lorentzian structure on $M$. Indeed, the divergence theorem can be stated merely with the specification of a volume form on the manifold $M$ (for computing the divergence of a vector field). In our case we have taken the volume form associated with the coordinates $(t, x)$. In application to Lagrangian theories of maps, the Lorentzian metric on $M$ will factor into the construction of the vector field $J$; see the next section.

ideas behind the theory, an advertisement for this simple yet powerful technique that is well adapted for use in Lagrangian field theories. In a forthcoming paper [61], Speck and the author will give a more accurate and detailed introduction to regular hyperbolicity and fill in some material omitted in [10].
4.2. Regular hyperbolicity

Christodoulou provides in his monograph [10] a robust and geometric way for obtaining compatible energy currents for any given Lagrangian theory of maps, using techniques similar to those used by Hughes et al in a non-geometric context [26]. In this subsection, we quickly review the key points of the theory. The main results obtained by Christodoulou are that

(1) vector fields satisfying (14b), up to a total divergence term, and up to Noether currents coming from symmetries of the target manifold, generically arise from contractions of arbitrary vector fields against what is called the canonical stress tensor, which can be obtained algorithmically from the equation of motion;
(2) the existence of vector fields satisfying (14a) depends only on the causal properties of the canonical stress tensor.

Let us begin by summarizing the construction of the canonical stress tensor. Let \( \psi \) be a function taking values in \( \mathbb{R}^n \). We will use index notation where components of \( \mathbb{R}^n \) are indicated by capital Latin letters. Assume that \( \psi \) solves a system of second-order partial differential equations of the form

\[
m_{ab}^{\alpha\beta} \partial_{a\beta}^2 \psi^\beta = F_A(\psi, \partial \psi),
\]

(15)

where lowercase Latin letters denote indices on the spacetime manifold, and \( m_{ab}^{\alpha\beta} \) is some field of coefficients, symmetric in \( a, b \), and symmetric in \( A, B \) (see remark 9 below for a discussion on these symmetry assumptions). Consider the tensor field defined by

\[
Z_{ab} = m_{ab}^{\alpha\beta} \delta_c^d - m_{cb}^{\alpha\beta} \delta_a^d - m_{ac}^{\alpha\beta} \delta_b^d;
\]

(16)

it is a direct computation to show that the tensor field

\[
Q[\psi]_{cd} := -Z_{ab}^{\alpha\beta} \psi^\alpha \psi^\beta
\]

(17)

has the property

\[
\partial_c Q[\psi]_{cd} = -\left( \partial_c Z_{a}^{b\alpha} \right) \psi^\alpha + 2 (\partial_d \psi^B) F_B(\psi, \partial \psi).
\]

(18)

Therefore, for any vector field \( X \), the vector field \( J^c := X^c |X|^2 + \sum_{|\alpha| \leq k} Q[\nabla^\alpha \psi]_{cd} X^d \) satisfies (14b). Note that thus far the construction of \( J \) is an algebraic and algorithmic statement about Lagrangian theories of maps.

Remark 9. For an arbitrary Lagrangian theory of maps, the associated coefficients \( m_{ab}^{\alpha\beta} \) are computed by taking the second variation of the Lagrangian function relative to the field velocity; see (20) and surrounding text below. Therefore, it is symmetric under pairwise exchanges of indices \((a, A) \leftrightarrow (b, B)\). However, since the equations of motion (19) are of
second order, and that the Hessian $\nabla_{ab}^2 \phi_B$ of a function is symmetric, we see that only the portion that is symmetric in $a, b$ (and hence the portion that is symmetric in $A, B$) appears in the equations of motion. The other portion, which is antisymmetric in $a, b$ (and hence antisymmetric in $A, B$) does not contribute to the equations of motion, and in fact can be captured in total divergence terms in the energy integral. See also [10], sections 3.2, 5.0 and 5.1.

To obtain hyperbolicity, it is necessary to also satisfy (14a). It was shown by Christodoulou that given a foliation $\Sigma_t$ of $M$ by level sets of a function $t$, a sufficient condition for $J$ (defined via the vector field $X$) to satisfy (14a) is for the following to hold:

1. $m_{ab}^{\tau\nu}(d\tau)d\nu$ is a negative definite matrix (in the index $A, B$);
2. $m_{aB}^{\tau\nu}\xi^a\xi^b$ is a positive definite matrix for every non-zero co-vector $\xi$ satisfying $\xi(X) = 0$.

(Please refer to [10] for the proof of this fact.) We call functions $t$ that satisfy the first property *time functions* and vector fields $X$ that satisfy the second property *observer fields*. For semilinear equations where $m_{ab}^{\tau\nu} = g^{ab}h_{\tau\nu}$, these notions agree with those of time functions and timelike observer fields in general relativity. In the quasilinear case, the time functions and observer fields form a replacement for the usual causal structure of the Lorentzian metric for governing the kinematics in perturbative analysis of solutions. One can also analogously define the notion of global hyperbolicity, domain of dependence and maximal development of initial data relative to this replacement causal structure. (This formulation is, in particular, used in Christodoulou’s seminal work on shock instabilities of the Euler equation [11].) In [10] it is claimed with a partial proof that the existence of a time function and an observer field is sufficient to guarantee the local well-posedness of the Cauchy problem with data prescribed on a level surface of $t$; a complete proof will be supplied in [61]. For the non-geometric scenario working over a fixed coordinate system in Minkowski space, with $X = \partial_t$, a proof is available in [26].

### 4.3. Breakdowns of hyperbolicity

When the existence of time functions and observer fields fails, the regular hyperbolicity of the system breaks down. This in particular implies the non-existence of general energy estimates by the work of Christodoulou [10], and hence the impossibility of applying the usual iteration method to obtain local existence and uniqueness of solutions to the Cauchy problem. While this does not preclude the possibility that such systems can be studied within a framework with a *more general* notion of hyperbolicity, here we provide some heuristic arguments as to why the lack of regular hyperbolicity is indicative of a lack of local well-posedness.

To illustrate the different modes of breakdown, we first consider the linear equations given by

$$\epsilon_0 u_t + \epsilon_1 u_{11} + \epsilon_2 u_{22} + \epsilon_3 u_{33} = 0$$

for some scalar $\epsilon_a$ on $\mathbb{R}^4$, with $\epsilon_a$ taking values ±1. The coefficients $m_{ab}^{\tau\nu}$ for this equation take values in $1 \times 1$ matrices, i.e. scalars, and we have that

$$m_{ab}^{\tau\nu} = \text{diag}(\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3).$$

The usual case of the wave equation is given by $-\epsilon_0 = \epsilon_1 = \epsilon_2 = \epsilon_3 = 1$. It is clear that then for a given covector

$$\xi = \xi_0 dt + \xi_1 dx^1 + \xi_2 dx^2 + \xi_3 dx^3,$$

4 We should recall at this stage that we consider well-posedness in the sense of Hadamard [21]: the existence of suitably regular solutions, the uniqueness of said solutions and the continuous dependence of the solutions on given initial data.
the requirement for \( m^{ab} \xi_a \xi_b \) to be negative definite (that is, negative as a scalar) is the usual condition that \( \xi \) is time-like with respect to the metric \( m^{ab} \): \( \xi_0^2 > \xi_1^2 + \xi_2^2 + \xi_3^2 \). Similarly, for \( m^{ab} \xi_a \xi_b \) to be positive, we just reverse the preceding inequality. Hence, the wave equation is hyperbolic if we choose \( t \) to be a time-like (in the usual sense) foliation and \( X \) to be any time-like (again in the usual sense) vector.

The standard example in which one cannot construct any time functions is the case of Laplace’s equation, where \( \epsilon_* = 1 \), and for any covector \( m^{ab} \xi_a \xi_b \) is positive. For elliptic problems, it is well known [22, 63] that the Cauchy problem is ill-posed: there cannot be continuous dependence on initial data. By comparison, we say that a Lagrangian theory of maps has an elliptic-type breakdown of hyperbolicity if one cannot construct any time functions even locally. Note that elliptic-type breakdowns are not the same as the equations forming a bona fide regularly elliptic system, which requires the ‘Legendre–Hadamard condition’ [10] that for all covectors \( \xi \), \( m^{AB} \xi_a \xi_b \) be positive definite matrices. What we call elliptic-type breakdown only suffices that none of those matrices be negative definite, and in particular mixed signatures will imply breakdown.

The case where one can construct a time function but not any observer field we refer to as ultrahyperbolic-type breakdown; the name is taken from the canonical example of the ultrahyperbolic equation

\[-u_{tt} - u_{11} + u_{22} + u_{33} = 0,
\]

where any foliation with the normal covector \( \xi \) satisfying \( \xi_0^2 + \xi_1^2 > \xi_2^2 + \xi_3^2 \) contributes a time function, but the trace of \( m^{ab} \) to any three-plane is indefinite. The ultrahyperbolic equations have infinite speeds of propagation [27], and can be checked by the theorem alluded to in the beginning of this section to have a non-hyperbolic polynomial symbol, and hence cannot admit well-posed Cauchy problems. (The instability of ultrahyperbolic equations has also been considered on physical grounds in the literature, see e.g. [15].)

**Remark 10.** The ultrahyperbolic equation illustrates an important connection between hyperbolicity and finite speed of propagation. As mentioned, the ultrahyperbolic equation does not admit the well-posed Cauchy problem for all smooth initial data. It was however pointed out by Craig and Weinstein [13] that, if one were willing to impose ‘non-local constraints’ (in their case a correlation on the admissible spacetime frequencies of the waves), the Cauchy problem can be well posed in the restricted class. One should think of the non-local constraints as circumventing the instabilities caused by infinite speeds of propagation by artificially requiring cancellations. For semilinear equations, it may be possible to impose such constraints a priori and globally (provided these additional constraints are compatible with the equations of motions; the constraint given in [13], section 2, is not preserved if one were to modify the ultrahyperbolic equation by a power nonlinearity \( |u|^p u \)); the situation for quasilinear equations is much less clear.

**Remark 11.** The conditions specified on \( m^{ab}_{AB} \) for regular hyperbolicity can also be viewed as a condition on plane wave perturbations. For a fixed background solution, we can study the linearized equations of motion for a plane wave ansatz. Then, regular hyperbolicity is precisely the geometric criterion that in this ansatz one cannot find linear instabilities due to exponentially growing modes with arbitrarily large rates. This should be compared to the analysis in [14] and [62].
4.4. Application to the Lagrangian theory of maps

A direct computation shows that the coefficient tensor $m_{aB}^{AB}$ can be obtained as the second variational derivative of the Lagrangian function relative to the field velocity. That is:

$$m_{aB}^{AB}(x, \phi, \partial \phi) = \frac{\delta^2}{\delta \partial_a \phi^A \delta \partial_b \phi^B} L(x, \phi, \partial \phi).$$  \hspace{1cm} (20)

Noting that the positive (negative) definite matrices form a convex cone inside the space of matrices, we see that if $t$ and $X$ are simultaneously time functions and observer fields for a collection of Lagrangian functions $L_i$, then they will also form a pair of time function and observer field for any convex sum of the $L_i$'s.

Turning now to the individual cases $L = \sigma_j(D\phi)$ as described in section 2, we can again use (11) to facilitate computations. Directly we obtain

$$\frac{\delta^2 L}{\delta \partial_d \phi^D \delta \partial_b \phi^B} = 2 \delta g^{a_1 c_1} \cdots g^{a_j c_j} (\phi^* h)_{a_1 c_1} \cdots (\phi^* h)_{a_j c_j} \cdot \left[ (\phi^* h)_{a_1 c_1} \cdots (\phi^* h)_{a_j c_j} h_{BD} + 2(j - 1) \delta_{c_j}^a \delta_{c_{j-1} a} \phi^A \delta_{d_1 a} \phi^D \delta_{c_j}^b h_{AB} h_{CD} \right].$$

Now consider an arbitrary time-like unit covector $\xi$, and let $\hat{g}$ be the restriction of $g$ to the orthogonal complement of $\xi$ (in particular it is a positive semidefinite bilinear form); we see that we can write $m_{BD}^{bd} \xi_b \xi_d$ in the following way:

$$m_{BD}^{bd} \xi_b \xi_d = -2 \delta g^{a_1 c_1} \cdots g^{a_j c_j} (\phi^* h)_{a_1 c_1} \cdots (\phi^* h)_{a_j c_j} h_{BD} - \delta_{c_j}^a \delta_{c_{j-1} a} \phi^A \delta_{d_1 a} \phi^D \delta_{c_j}^b h_{AB} h_{CD} \right]$$

(21)

after carefully counting the various permutations of indices leaving the expression non-zero.

Using the Cauchy inequality on the term now in the brackets, we see that it is positive semidefinite; it will be positive definite whenever $d\phi|_{p \in M} : T_p M \supset T_{\phi(p)} N$ has rank at least 2 (that is, the tangent space map $d\phi$ restricted to the vectors in the kernel of $\xi$ has rank at least 2). Therefore, we have the following theorem.

**Theorem 12.** Let the Lagrangian $L$ be a convex linear combination of $\sigma_j(D\phi)$ and $s(\phi)$. And let $t$ be an arbitrary time function relative to the underlying metric $g_{ab}$. Then

1. if the coefficient for $\sigma_1$ is non-zero, then $t$ is a time function for the equations of motion;
2. if $d\phi$ restricted to the level sets of $t$ has rank $\geq \max(j - 1, 2)$, then $t$ is a time function for the equations of motion;
3. failing both of the above, $t$ is borderline degenerate, but the corresponding $m_{AB}^{ab}(dt)_a(dt)_b$ is positive semi-definite.

In particular, $L$ cannot have a bona fide elliptic-type breakdown of hyperbolicity.

Noting that the domain of dependence in the theory of hyperbolic equations is determined precisely by the admissible time functions [10, 39], we have the following interesting corollary, which implies that a Lagrangian theory of maps with the dominant energy condition (compare theorem 6), when in fact hyperbolic, cannot propagate perturbations faster than the speed of gravity.

**Corollary 13.** Let $L = L(s, \sigma_1, \ldots, \sigma_j)$ be a Lagrangian function with the property that $L$ is non-decreasing and concave in its arguments. Assume further that for some $k$, the partial derivative of $L$ with respect to $\sigma_k$ is bounded from below, and that either (a) $k = 1$ or (b) admissible solutions have rank $(d\phi) \geq \max(k, 3)$. Supposing that the dynamics derived from $L$ is regularly hyperbolic, the domain of dependence for solutions must be strictly contained within that of the linear wave equation on the spacetime $(M, g)$. 

15
Proof. Using the energy estimate, it suffices to show that any time function \( t \) of the underlying spacetime is a time function for the Lagrangian theory (see [10], section 5.3, for a discussion). Using (20) we write

\[
\begin{align*}
m_{AB}^{ab} &= \frac{\delta^2 L}{\delta \sigma_i \delta \sigma_j} \delta \sigma_i \delta \sigma_j + \frac{\delta L}{\delta \sigma_j} \delta \sigma_j \delta \sigma_j
\end{align*}
\]

After contracting with \((dt)_a (dt)_b\), the first term is positive semi-definite using the concavity of the Lagrangian functional. The second term is a convex sum of positive semi-definite matrices, and by theorem 12 and the hypotheses, at least one of the terms is positive definite. Hence, \( t \) is a time function for the Lagrangian theory of maps described by \( L \).

Remark 14. We can compare the above corollary to the situation in isentropic fluids. After choosing an equation of state, one can write the pressure \( p \) of the fluid as a function of the proper energy density \( \rho \). In this regime the well-known criterion for hyperbolicity is that the speed of sound \( \sqrt{\rho / \rho_0} \) is positive and real, while the criterion for the fluid flow to be causal is that the speed of sound is less than the speed of gravity (which we can choose equal to 1) [62]. The proper energy density is in fact the Lagrangian function for a fluid. Writing \( \tau = (\rho_n(\rho^0))^0/2 \) for the volume per particle, the pressure can be given [10], section 1.3, by

\[
\begin{align*}
\sigma_n &= \frac{1}{2} \left( \frac{\rho}{\rho_0} \right) \left( -\frac{2}{3} \right) \\
\frac{\partial L(\sigma_n(D\sigma))}{\partial \sigma} &= -\frac{1}{2} \left( \frac{\rho}{\rho_0} \right) \left( -\frac{2}{3} \right)
\end{align*}
\]

Using (20) we write

\[
\begin{align*}
\rho_n &= \frac{1}{2} \left( \frac{\rho}{\rho_0} \right) \left( -\frac{2}{3} \right)
\end{align*}
\]

It suffices to show that any time function \( t \) for the Lagrangian theory of maps described by \( L \) ensures hyperbolicity; compare with the discussion below. Note that a necessary condition for the second inequality to hold is that \( \rho_n \) is concave, agreeing with the corollary above. The second inequality guarantees hyperbolicity; compare with the discussion below. Note that a necessary condition for the second inequality to hold is that \( L \) is increasing as a function of \( \sigma_n \), which is also a part of the hypothesis in the above corollary.

The analogous statement for observer fields, however, does not hold. Let \( \eta \) now be an arbitrary space-like unit covector, and let \( X \) be a unit time-like vector orthogonal to \( \eta \). Denote now by \( \tilde{g} \) the restriction of \( g \) to the orthogonal complement of \( \eta \) and the metric dual of \( X \). We see that the analogous statement to (21) is as follows:

\[
\begin{align*}
m_{BD}^{ab} \eta_b \eta_d &= 2 \tilde{g}^{a[c_1]} \cdots \tilde{g}^{a[j\cdots k\cdots 1]}(\phi^* h)_{[a} c_1 \cdots (\phi^* h)_{j\cdots k]} c_d \\
&= \frac{2}{j-1} \tilde{g}^{a[c_1]} \cdots \tilde{g}^{a[j\cdots k]}(\phi^* h)_{[a} c_1 \cdots (\phi^* h)_{j\cdots k]} \\
&= \frac{2}{j-1} \tilde{g}^{a[c_1]} \cdots \tilde{g}^{a[j\cdots k]}(\phi^* h)_{[a} c_1 \cdots (\phi^* h)_{j\cdots k]} \\
&= \frac{2}{j-1} \tilde{g}^{a[c_1]} \cdots \tilde{g}^{a[j\cdots k]}(\phi^* h)_{[a} c_1 \cdots (\phi^* h)_{j\cdots k]}
\end{align*}
\]

where the first term on the right-hand side is again positive semi-definite by Cauchy’s inequality, while the second and third terms are a priori negative semi-definite. In particular, comparing the first and third terms, if \( |X\phi|_0^2 \) is large compared to \( \tilde{g}^{a[c}(\phi^* h)_{a]} \), the bilinear form \( m_{BD}^{ab} \eta_b \eta_d \) can easily have negative eigenvalues; one can compare this to the conclusion drawn in [14] about instabilities of the Skyrme model. We are forced to conclude that a general Lagrangian theory of maps is susceptible to ultrahyperbolic-type breakdowns of hyperbolicity.

On the other hand, note that in the case \( X\phi = 0 \), the form \( m_{BD}^{ab} \eta_b \eta_d \) is indeed positive semi-definite, and furthermore positive definite if \( d\phi \) satisfies an analogous rank condition as
in theorem 12. Also note that in the case $j = 1$ (the semi-linear case), positive definiteness of $m_{bd}^j \eta_b \eta_d$ holds without the need of any assumptions on $X \phi$, as the problematic terms do not appear.

**Remark 15.** As seen in the discussions above, the term $\sigma_1$ in the Lagrangian always introduces a factor that is regularly hyperbolic. This in fact has a stabilizing effect on the hyperbolicity of the field theory. An example will be given in section 5, where the hyperbolicity of the Skyrme model persists into a regime where the particle velocity (using a fluid interpretation) exceeds that of the speed of gravity.

### 4.5. Canonical stress versus Einstein–Hilbert stress energy

Thus far we have seen that for $\sigma_j$, the dominant energy condition must hold; on the other hand, the above computations show that the analogous statement for the canonical stress need not hold. That is, for perturbations $\psi$ over a solution $\phi$ to the equations of motions determined by $\sigma_j (D \phi)$, the canonical stress $Q[\psi]_{\xi} X^d \xi_c \geq 0$ for any time-like vector $X$ and time-like covector $\xi$ with $\xi(X) < 0$. However, it is also clear from the definitions that the canonical stress tensor and the Einstein–Hilbert stress–energy tensor in fact agree in the highest order derivative terms for the semilinear case $L = \sigma_1 (D \phi)$, in which the equations of motion are always regularly hyperbolic. What is, then, the difference between the two stress tensors?

In fact we have the following.

**Proposition 16.** For $L = \sigma_1 (D \phi)$, the canonical stress tensor $Q[\phi]$ corresponding to the solution itself agrees with $g^{-1} \circ T$, the Einstein–Hilbert stress–energy tensor.

The proof follows from a direction computation of $Q[\phi] X^d \xi_c$ using the formulae given in (20) et seq., and comparison against (12). We omit the details here. This proposition shows that for Lagrangians composed of convex linear combinations of the $\sigma_j$’s, the dominant energy condition is sufficient to guarantee the existence of a compatible energy current that controls the solution itself, but not its higher derivatives. This is, of course, unsurprising, as the Einstein–Hilbert stress–energy tensor is divergence free by definition, and hence for any time-like vector field $X$, $g^{-1} \circ T \circ X$ is a compatible current.

It is tempting to propose that, in view of proposition 16, regular hyperbolicity implies the dominant energy condition. This is, however, not necessarily the case in general. Returning to (14a) of the fundamental energy estimate, we see that it only suffices that the compatible current is comparable with the Sobolev norms in an integrated sense, and indeed, the conditions on regular hyperbolicity encoded in positivity and negativity conditions on $m_{ab}^2$ only guarantee that much. By way of an example, we consider a linear equation on $\mathbb{R}^{1+2}$ with a two-dimensional target manifold. Consider the following set of matrices:

\[
\begin{align*}
\tilde{m}_{00}^{00} &= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\
\tilde{m}_{01}^{01} &= \tilde{m}_{02}^{02} = 0 \\
\tilde{m}_{11}^{11} &= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \\
\tilde{m}_{22}^{22} &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\
\tilde{m}_{12}^{12} &= \tilde{m}_{21}^{21} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}
\end{align*}
\]
and let $m^{ab}_{AB} = \tilde{m}^{ab}_{AB} - \epsilon \delta^{ab} \delta_{AB}$ for some $\epsilon \ll 1$. Now, by construction, $m^{00}_{AB}$ is clearly negative definite, so the usual function $t$ is a time function. A direct computation shows that $\tilde{m}^{ab}_{AB} \eta_i \eta_a$ is positive definite for any non-zero $\eta$ that satisfies $\eta(\partial_i) = 0$. Hence, for sufficiently small $\epsilon$, the coefficient matrix $m^{ab}_{AB}$ is regularly hyperbolic, and the equation

$$m^{ab}_{AB} \partial^2_{ab} \psi^B = 0$$

has a well-posed initial value problem for finite energy initial data. However, the canonical stress tensor $Q[\psi^B]_0$ is not point-wise positive definite for all initial data! Take an arbitrary smooth initial data with compact spatial support on $t = 0$ such that inside the ball of radius 1, the data take the following explicit values:

$$\psi^1(0, x, y) = y, \quad (24a)$$

$$\psi^2(0, x, y) = -x, \quad (24b)$$

$$\partial_t \psi^A(0, x, y) = 0. \quad (24c)$$

It is easy to check that for this initial data, $m^{ab}_{AB} \partial_a \psi^A \partial_b \psi^B = 0$ in a spatial neighborhood of 0. And hence in that neighborhood the energy density for $Q[\psi^B]_0$ is in fact point-wise negative.

This lack of point-wise positivity is a reflection of the internal structure of the target manifold $N$. Indeed, if $N$ were one dimensional or effectively one dimensional (i.e. the coefficient matrix $m^{ab}_{AB} = g^{ab} h_{AB}$ is separable), two waves that are spatially coincident and travel in the same direction must interact, and hence a point-wise measurement can capture the correct notion of energy. But when $N$ has an internal structure, two waves that are spatially coincident and travel in the same direction do not have to interact strongly, since they may be orthogonal on $N$. This orthogonality condition cannot be detected point-wise in the energy density, but manifests itself as ‘miraculous cancellations’ when the total energy is considered. In other words, to detect and account for this orthogonality requires a spatial mode decomposition of the fields, as seen in [10], chapter 5.

The fact that the dominant energy condition cannot guarantee hyperbolicity can be seen also as a manifestation of this internal structure. Proposition 16 effectively says that perturbations $\psi$ which are parallel on $N$ to the background solution $\phi$ can also be controlled by the dominant energy condition. In the case where the dimension of $N$ is greater than 1, it is precisely those additional degrees of freedom, which cannot be accounted for merely by the dominant energy condition, that require the framework of regular hyperbolicity.

### 5. Hyperbolicity of the Skyrme model

It is easy to see that the Skyrme model can be considered as a theory of elasticity with the restrictions on the particle world-lines relaxed. This connection has also been observed by Slobodeanu [59]. From the point of view of relativistic fluids and elasticity, it is then perhaps less surprising that highly boosted Skyrmions are expected to be unstable [14]. Here, we compute exactly the symbol $m^{ab}_{AB}$ associated with the Skyrme model and exhibit an ultrahyperbolic type breakdown of hyperbolicity in the tachyonic regime.

For the following discussion, we choose the scale factors to normalize the Skyrme Lagrangian [57, 58] to

$$S = \int \frac{1}{2} g^{ab} h_{AB} \partial_a \phi^A \partial_b \phi^B + \frac{1}{4} (g^{ab} h_{AB} g^{cd} h_{CD} - g^{ab} h_{AD} g^{cd} h_{CB}) \partial_a \phi^A \partial_b \phi^B \partial_c \phi^C \partial_d \phi^D \, \text{dvol}_g,$$

(25)
We ignore the mass term $s(\phi)$ as it plays no role in the discussion for hyperbolicity; here, $\phi : \mathbb{R}^{1+3} \to S^1$ is assumed as usual. Applying theorem 6 we see that (with or without the positive mass term) the dominant energy condition is always satisfied for this model.

Using (20), the coefficients $m_{AB}^{\alpha\beta}$ can be computed to effectively (see remark 9) be

$$m_{AB}^{ab} = g^{ab} h_{AB} (1 + |\phi|^2_{g,AB} + \partial\phi g^b\phi_B - g^{ab} g^{cd} \partial_c \phi_A \partial_d \phi_B - h_{AB} h_{CD} \partial_c \phi^C \partial_d \phi^D).$$

(26)

From theorem 12 we already see that any time function of the Minkowski space is an admissible time function. Therefore, it suffices to examine the conditions for $m_{AB}^{\alpha\beta} \eta_a \eta_b$ to be positive definite.

The following computations are most illustrative in a frame adapted to $d\phi$. There are two cases: the kernel of $d\phi$ being a degenerate (null) subspace, and everything else. In the everything else category, there exists an orthonormal frame of $T_p M$ that simultaneously diagonalizes $g$ and $\phi^* h$; in the case where the kernel of $d\phi$ is degenerate, there exists an exceptional null frame (see, e.g. [47], chapter 9, exercises 18 and 19, or [36], with some cases ruled out by virtue of $\phi^* h$ being positive semi-definite). The null case can be checked to be hyperbolic much in the same way as described below, so we omit its discussion (note that for the null case the canonical stress corresponding to $\sigma_2$ is degenerate; already it requires the presence of the wave-map term $\sigma_1$ in the Lagrangian for regular hyperbolicity) and focus on the generic case of an orthonormal basis.

Let the vector basis be $e_0, e_1, e_2, e_3$. orthonormal with respect to the metric $g$, and diagonal with respect to $(\phi^* h)$. Let their corresponding covector basis be $f_0, f_1, f_2, f_3$ where $f_i(e_j) = \delta_{ij}$. And let $\lambda_0^2, \lambda_1^2, \lambda_2^2, \lambda_3^2$ be the values for $(\phi^* h)(e_i, e_j)$, respectively; at least one of which must vanish from rank considerations. Let $e_i'$ be a unit vector in the tangent space of the target manifold such that the push-forward $d\phi(e_i) = \lambda_i e_i'$ for $e_i$ such that $d\phi(e_i) = 0$, choose $e_i'$ such that one of the $e_i'$ vanishes, and the rest form an orthonormal basis for $h$. Take $f_i'$ to be their corresponding covectors.

Then, we can write

$$g^{ab} = (-e_0 \otimes e_0 + e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3)^{ab}$$

(27a)

$$h_{AB} = (f_0' \otimes f_0' + f_1' \otimes f_1' + f_2' \otimes f_2' + f_3' \otimes f_3')_{AB}$$

(27b)

$$\partial\phi^A = (\lambda_0 f_0 \otimes e_0' + \lambda_1 f_1 \otimes e_1' + \lambda_2 f_2 \otimes e_2' + \lambda_3 f_3 \otimes e_3')^A.$$  

(27c)

Let us consider, in this basis, the components $m_{AB}^{ab}$. (The computations are symmetric in $[1, 2, 3]$, and the component $m_{AB}^{00}$ is already treated by theorem 12.) A direct computation shows that

$$m_{AB}^{33} = (1 + \lambda_1^2 + \lambda_2^2)f_0' \otimes f_0' + (1 - \lambda_0^2 + \lambda_3^2)f_1' \otimes f_1'$$

$$ + (1 - \lambda_0^2 + \lambda_3^2)f_2' \otimes f_2' + (1 - \lambda_0^2 + \lambda_3^2)f_3' \otimes f_3'$$

(28a)

$$m_{AB}^{30} = -\lambda_3 \lambda_0 (f_0')_A (f_0')_B$$

(28b)

$$m_{AB}^{31} = \lambda_3 \lambda_1 (f_1')_A (f_1')_B.$$  

(28c)

Now first consider the case where $\lambda_0 = 0$. Then, we see that $m_{AB}^{33}$ is automatically positive definite. We have already discussed this case in the previous section: if $d\phi$ has a time-like element in its kernel, then that element is a valid observer field.

For the case where $\lambda_0 \neq 0$, we consider first when $\lambda_3 = 0$. Then, we see that if $\lambda_0^2 > 1 + \min (\lambda_1^2, \lambda_2^2)$, we have $m_{AB}^{33}$ no longer positive. Furthermore, we have by the above
computation that in this case, for \( \eta = sf_3 + rf_0 \), the matrix \( m_{AB}^{ab} \eta_a \eta_b = s^2 m_{AB}^{33} + r^2 m_{AB}^{00} \) is a sum of an indefinite bilinear form with a negative definite one. Therefore, we see that for any covector \( \eta \) in the span of \( f_0 \) and \( f_3 \), we cannot have \( m_{AB}^{ab} \eta_a \eta_b \) to be positive definite. In particular, this implies that in this case regular hyperbolicity must fail.

For the case where \( \lambda_0 \neq \lambda_3 \), and \( \lambda_1 = 0 \), we see that \( m_{AB}^{33} \) loses positivity as soon as \( \lambda_3^2 > 1 \). Now we consider again \( \eta = sf_3 + rf_0 \). Computing explicitly we obtain

\[
m_{AB}^{ab} \eta_a \eta_b = (s^2 - r^2) (1 + \lambda_0^2) f_0 \otimes f_0 + \left[ s^2 (1 - \lambda_0^2 + \lambda_3^2) - r^2 (1 + \lambda_0^2 + \lambda_3^2) \right] f_3' \otimes f_3' + \\
+ \left[ s^2 (1 - \lambda_0^2) - r^2 (1 + \lambda_3^2) \right] f_2' \otimes f_2' + \left( s^2 - r^2 \right) (1 + \lambda_3^2) f_3' \otimes f_3' + \\
- (s \lambda_0 f_2' + r \lambda_3 f_0) \otimes (s \lambda_0 f_3' + r \lambda_3 f_0),
\]

where we see that the coefficient of \( f_2' \otimes f_2' \) will always be negative. This implies that for any \( \eta \) of this form, when \( \lambda_1 = 0 \) and \( \lambda_3^2 > 1 \), we have that \( m_{AB}^{ab} \eta_a \eta_b \) cannot be positive definite.

From these computations, we obtain the following theorem.

**Theorem 17.** The Skyrme model is regularly hyperbolic when any of the following is true.

1. The kernel of \( \partial \phi \) contains a time-like vector.
2. The kernel of \( \partial \phi \) is a degenerate subspace.
3. The time-like eigenvector for \( D^a \) has a corresponding eigenvalue with norm less than 1.

The Skyrme model has an ultrahyperbolic breakdown of hyperbolicity when \( D^a \) admits a time-like eigenvector with eigenvalue with norm greater than 1.

**Remark 18.** The number ‘1’ appearing in the third condition depends on the normalization chosen for the Lagrangian, which dictates the interaction strength between the wave map and fluid-like terms. This result should be compared with the conclusion drawn by Crutchfield and Bell in [14], where their linear analysis shows that a sufficient criterion for breakdown of hyperbolicity is, in the above language, \( \lambda_3^2 > 1 + \lambda_0^2 + \lambda_3^2 \). So the present result sharpens the regime for which hyperbolicity fails. Note that this failure of hyperbolicity is not merely that of the regular hyperbolicity method. In parts of this regime it can be shown explicitly that the associated linear system is ill-posed; see the next remark.

**Remark 19.** Fixing \( \lambda_3 = 0 \) and \( \lambda_0^2 < \lambda_3^2 - 1 < \lambda_3^2 \) (or \( \lambda_0^2 > 1 + \lambda_3^2 + \lambda_3^2 \), which is the Crutchfield–Bell condition), it is easy to check (using Descartes’ rule of signs) that the linear, constant coefficient equation \( m_{AB}^{ab} \partial_c \phi^a \phi^B = 0 \) cannot be hyperbolic (in the sense that its polynomial symbol has the requisite number of real roots). Recall that hyperbolicity requires there to be a hyperbolic direction \( \eta \) such that, for any \( \xi \) transverse to \( \eta \), the polynomial \( M(s) = \det [m_{AB}(\xi + s \eta, \xi + s \eta)] \) has only real roots [19, 25]. We fix \( \xi = f_3 \). Using that \( \lambda_3 = 0 \), we can exploit a symmetry condition that if \( \eta \) is hyperbolic, so will \( \eta - 2g(\eta, f_3) f_3 \). Using that the hyperbolic directions form a convex cone [19], we can assume the would-be hyperbolic direction is orthogonal to \( f_3 \). For such an \( \eta \), we have that \( M(s) \) is an even polynomial. The computations given before the statement of the theorem imply that \( \lim_{s \to \pm \infty} M(s)/s^6 < 0 \), with \( M(0) < 0 \). Therefore, \( M(s) \), a sixth-degree polynomial, can have at most four real roots. The same argument, reversing the role of \( \eta, \xi \), can also be used to rule out \( f_3 \) as a hyperbolic direction, proving the claim.

If we assume the local well-posedness result claimed in [10], then the above theorem implies the following well-posedness property for the Cauchy problem of the Skyrme model.
Corollary 20. The Cauchy problem for the Skyrme model with almost stationary initial data, where almost stationary is read to satisfy the hypotheses of theorem 17, is locally well-posed. In particular, a small perturbation of a static Skyrmion configuration gives rise to a well-defined local-in-time evolution.

Remark 21. The existence of static configurations to the Skyrme model is a partially open problem. In certain symmetry classes the existence is known, see [17, 30, 40]. In the case of the perturbation of a static configuration, local well-posedness also follows, by the computations above, using the techniques of [26]. The framework of regular hyperbolicity is not necessary in that regime. The main improvement of the above corollary is that it is capable of handling initial data obtained from perturbations of highly boosted static configurations. In this case, ‘∂tφ’ can be sufficiently large that the results of [26] cannot apply directly, while the geometric methods described here are agnostic to such changes of coordinate systems.

Remark 22. It has been pointed out to the author by Dan Geba that one automatically has local well-posedness with arbitrary smooth initial data, if one were to consider the spherically symmetric Skyrme model. Indeed, once the symmetry is imposed, the target space (being a quotient of a three-dimensional manifold by a symmetry with two-dimensional orbits) is effectively one dimensional, and the ultrahyperbolic-type breakdown which is due to the internal structure of the target manifold cannot occur (see also section 4.5).

However, in this situation one has a hyperbolic illustration of the failure of the Coleman principle, analogous to that observed by Kapitanski and Ladyzhenskaya [30]. More precisely, the existence and uniqueness of solutions in the spherically symmetric class does not automatically guarantee that said solutions are in fact unique in the non-spherically symmetric class. This is because the spherically symmetric solution is only guaranteed to be stable under spherically symmetric perturbations. In our case, however, the above analysis shows that the linearized system is not necessarily stable under asymmetric perturbations, and therefore we cannot use Cauchy stability to conclude that symmetric initial data must lead to symmetric solutions!

Acknowledgments

The author is supported by the Commission of the European Communities, ERC grant agreement no 208007. The author would like to thank Nick Manton and Claude Warnick for introducing him to the problem, and to Mihalis Dafermos, Gary Gibbons and Jared Speck for useful discussions.

References

[1] Beig R and Schmidt B G 2003 Relativistic elasticity Class. Quantum Grav. 20 889–904
[2] Brenier Y 2004 Hydrodynamic structure of the augmented Born–Infeld equations Arch. Ration. Mech. Anal. 172 65–91
[3] Bressan A 1978 Relativistic theories of materials Springer Tracts in Natural Philosophy vol 29 (Berlin: Springer)
[4] Cameron J A 1991 The reduction of Einstein’s equations on spacetimes with spacelike U(1) isometry groups PhD Thesis Yale University
[5] Carter B and Quintana H 1972 Foundations of general relativistic high-pressure elasticity theory Proc. R. Soc. A 331 57–83
[6] Cazenave T, Shatah J and Tahvildar-Zadeh A S 1998 Harmonic maps of the hyperbolic space and development of singularities in wave maps and Yang–Mills fields Ann. Inst. Henri Poincaré Phys. Theor. 68 315–49 (available at http://www.numdam.org/item?id=AIHPA_1998__68_3_315_0)
[42] Littman W 1963 The wave operator and $L_p$ norms J. Math. Mech. 12 55–68
[43] Loss M 1987 The Skyrme model on Riemannian manifolds Lett. Math. Phys. 14 149–56
[44] Manton N S and Ruback P J 1986 Skyrmions in flat space and curved space, Phys. Lett. B 181 137–40
[45] Manton N S and Sutcliffe P 2004 Topological Solitons (Cambridge: Cambridge University Press)
[46] Mazur P O 1982 Proof of uniqueness of the Kerr–Newman black hole solution J. Phys. A: Math. Gen. 15 3173–80
[47] O’Neill B 1983 Semi-Riemannian Geometry: With Applications to Relativity (New York: Academic)
[48] Penrose R 1965 Gravitational collapse and space-time singularities Phys. Rev. Lett. 14 57–9
[49] Raphael P and Rodnianski I 2009 Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang–Mills problems, arXiv:0911.0692
[50] Rodnianski I and Sterbenz J 2010 On the formation of singularities in the critical $O(3)$ $\sigma$-model Ann. Math. (2) 172 187–242
[51] Salt D 1971 Action principles for the elastic solid and the perfect fluid in general relativity J. Phys. A: Gen. Phys. 4 501
[52] Schoen R and Yau S-T 1981 Proof of the positive mass theorem II Commun. Math. Phys. 79 231–60
[53] Serre D 2004 Hyperbolicity of the nonlinear models of Maxwell’s equations Arch. Ration. Mech. Anal. 172 309–31
[54] Shatah J 1988 Weak solutions and development of singularities of the SU(2) $\sigma$-model Commun. Pure Appl. Math. 41 459–69
[55] Shatah J and Struwe M 1998 Geometric Wave Equations (New York, NY: Courant Institute of Mathematical Sciences)
[56] Sideris T 1989 Global existence of harmonic maps in Minkowski space Commun. Pure Appl. Math. 42 1–13
[57] Skyrme T H R 1954 A new model for nuclear matter. Proc. R. Soc. London 226 521–30
[58] Skyrme T H R 1961 Particle states of a quantized meson field. Proc. R. Soc. London 262 237–45
[59] Slobodeanu R 2010 Perfect fluids from high power sigma-models arXiv:1006.1133
[60] Speck J 2010 The nonlinear stability of the trivial solution to the Maxwell–Born–Infeld system arXiv:1008.5018
[61] Speck J and Wong W W-Y in preparation
[62] Tahvildar-Zadeh A S 1998 Relativistic and nonrelativistic elastodynamics with small shear strains Ann. Inst. Henri Poincaré Phys. Theor. 69 275–307 (available at http://www.numdam.org/item?id=AIHPA_1998__69_3_275_0)
[63] Tarkhanov N N 1995 The Cauchy Problem for Solutions of Elliptic Equations (Berlin: Akademie)
[64] Tao T 2001 Global regularity of wave maps: II. Small energy in two dimensions Commun. Math. Phys. 224 443–544
[65] Tao T 2008 Global regularity of wave maps: III. Large energy for $R^{1+2}$ to hyperbolic spaces, arXiv:0805.4666
[66] Tao T 2008 Global regularity of wave maps: IV. Absence of stationary of self-similar solutions in the energy class arXiv:0806.3592
[67] Tao T 2008 Global regularity of wave maps: V. Large data local wellposedness and perturbation theory in the energy class arXiv:0808.0368
[68] Tao T 2009 Global regularity of wave maps: VI. Abstract theory of minimal-energy blowup solutions, arXiv:0906.2833
[69] Tao T 2009 Global regularity of wave maps: VII. Control of delocalised or dispersed solutions, arXiv:0908.0776
[70] Wald R M 1984 General Relativity (Chicago, IL: University of Chicago Press)
[71] Witten E 1981 A new proof of the positive energy theorem Commun. Math. Phys. 80 381–402