Non-linear electromagnetism and special relativity

Denis Serre
École Normale Supérieure
de Lyon*†

To my friend Li Ta-tsien (Li Da-qian), on the occasion of his 70th anniversary

Abstract

We continue the study of nonlinear Maxwell equations for electromagnetism in the formalism of B. D. Coleman & E. H. Dill. We exploit here the assumption of Lorentz invariance, following I. Białinicki-Barula. In particular, we show that nonlinearity forbids the convexity of the electromagnetic energy density. This justifies the study of rank-one convex and of polyconvex densities, begun in [8, 16]. We also show the alternative that either electrodynamics is linear, or dispersion is lost as the electromagnetic field becomes intense.

1 Introduction

In an ideal medium like vacuum, the Maxwell's equations of electromagnetism are obtained as the Euler–Lagrange equations of an action $\mathcal{L}$ over the space of closed 2-forms. It admits a natural energy density $h(B, D)$, namely the Legendre–Fenchel transform of $\mathcal{L}$ with respect to the electric field E, at constant magnetic induction B. The classical theory of systems of conservation laws tells that if h is strongly convex, that is D^2h is positive definite, then the system is hyperbolic and the Cauchy problem is locally well-posed in the Sobolev space $H^s(\mathbb{R}^3)^6$ with $s > 5/2$. It was proved actually in [16] that strong *polyconvexity*, when viewing (D, B) as a 3×2 matrix-valued field, is sufficient for hyperbolicity and local well-posedness. This was illustrated by Y. Brenier [8] in his analysis of the Born–Infeld model.

We specialize here to the vaccum case, having in mind that the Maxwell's equation must be compatible with special relativity. By this, we mean that the action is frame indifferent, or invariant under Lorentz transformations. In the next section, we recall the characterization of this invariance in terms of the energy density (Proposition 2.2), already established by Białynicki-Barula [1]. The result involves a Hamilton–Jacobi equation that we integrate in

*UMPA (UMR 5669 CNRS), ENS de Lyon, 46, allée d'Italie, F–69364 Lyon, cedex 07, FRANCE.

†Part of this research has been done while the author was visiting the Institute of Applied Mathematics of Fudan University at Shanghai, P. R. of China. I thank also Michel Kiessling and Shadi Tahvildar-Zadeh for bringing the works by Mie, Plebansky and Białynicki-Barula to my attention.

Section 3. The surprizing fact is that either the Maxwell's equations are linear (h quadratic), or h is *not convex* (Theorem 3.1). This justifies *a posteriori* the analysis of [16] and questions the relevance of [5], since the latter assumes the convexity of the energy.

Section 4 is dedicated to the wave velocities. A natural requirement is that the system be hyperbolic and that its wave speeds be not larger than the fundamental velocity defined by the Minkowski metric, the so-called *causality*[1]. We first show that the weaker condition that h be strongly rank-one convex implies hyperbolicity, although we do not know how to deduce local well-posedness from this result. Then we turn to causality of the wave propagation. We prove that it holds true in the Born–Infeld model, a fact of which we failed to find a reference in the literature. Next, we give a new proof that in every model, and for every background state (B, D) there exists at least one (and in general two) space directions in which the wave speed equals c, the largest velocity compatible with causality; this result was previously established by Boillat in [2]. Finally, we find the amazing fact that if the model is causal and not linear, then it behaves as a pure transport (with velocity $P/|P|$ where $P := D \times B$) when the field (D, B) is intense; at the leading order, the dispersion effect disappears. This suggests that the linearity or the nonlinearity of electrodynamics could be decided by an appropriate experiment.

Our work raises open questions:

- Which one among the Lorentz invariant energies are (strongly) rank-one convex, or polyconvex ?
- Is it possible that a model be hyperbolic without h being rank-one convex ?

We point out finally that a rather accurate definition of hyperbolicity of nonlinear electrodynamics is given by D. Christodoulou (see [9], Chapter 6), where he states the action principle in terms of the potential of the electromagnetic field, instead of the field itself.

2 Relativistic nonlinear models of electromagnetism

We recall that the Maxwell equations of an electromagnetic field are obtained as the Euler–Lagrange equations of a Lagrangian, called the *Action*

$$\mathcal{L}[\omega] = \iint L(\omega)\, dx\, dt,$$

where the argument ω runs over the closed 2-forms. See for instance M. Born's memoir [6], following the seminal paper by G. Mie [13]. The components of ω are denoted by E_j and B_k, with

$$\omega = \sum_{j=1}^{3} E_j dx_j \times dt + \sum_{\epsilon(i,j,k)=1} B_i dx_j \times dx_k.$$

[1]This problem was studied by G. Boillat in several papers, for instance in [2, 4, 5] property. He employs the terminology *sub-luminal* instead, but we think that it is ambiguous. Light is nothing but an electromagnetic wave. As such, it cannot travel at sub-luminal speed!

The whole system writes

$$\partial_t B + \mathrm{curl} E = 0, \quad \partial_t D - \mathrm{curl} H = 0, \qquad D := \frac{\partial L}{\partial E},\ H := -\frac{\partial L}{\partial B}. \tag{1}$$

We warn the reader that the above system stands in vacuum. In other words, we do not consider interaction with other fields, like matter or gravitation.

In special relativity, we postulate that the physics is invariant under the linear transformations that preserve a distinguished Minkowski metric. These transformations are called Lorentz-, or Poincaré-, transformations in the literature. In suitable coordinates, the metric writes $dt^2 - dx_1^2 - dx_2^2 - dx_3^2$, which means that all velocities have to be compared with a fundamental one $c = 1$. For instance, kinematics postulates that particles may not travel at speed larger than c. It looks natural to postulate that any kind of information, including electromagnetic wave, may propagate only at speed lower than c. This is commonly refered to as *causality*.

Because electromagnetism is intimately related to special relativity, it is natural to assume that the density L be invariant under Lorentz transformations. As is well known, Lorentz invariance amounts to saying that

$$L = L_0\left(\frac{1}{2}(|B|^2 - |E|^2), E \cdot B\right) \tag{2}$$

for some function L_0 of two real variables. See [6], or go to Exercise 53 in [15].

In the sequel, we shall always assume that $E \mapsto L(B, E)$ is strictly convex. Somehow, it characterizes the orientation of the time arrow. Then we can make the Legendre transform

$$h(D, B) := \sup_{E \in \mathbb{R}^3} (D \cdot E - L(B, E)).$$

The Hamiltonian h is now convex with respect to D. The fields E and H are given by

$$E = \frac{\partial h}{\partial D}, \qquad H = \frac{\partial h}{\partial B}.$$

The system (1) has thus the form proposed by Coleman & Dill[2] in [10]:

$$\partial_t B + \mathrm{curl} E = 0, \quad \partial_t D - \mathrm{curl} H = 0, \qquad E := \frac{\partial h}{\partial D},\ H := \frac{\partial h}{\partial B}. \tag{3}$$

We recall that (3) is compatible (so long as B and D are Lipschitz continuous) with the *Poynting identity*, which plays the role of energy conservation, where h is the energy density and $E \times H$ the energy flux:

$$\partial_t h(D, B) + \mathrm{div} E \times H = 0. \tag{4}$$

[2]The fact that B, D, E, H are also the initials of the first and middle names of Coleman and Dill is fortuitous. Coleman and Dill had in mind the description of the electromagnetic field in a nonlinear material medium. Thus their models are not Lorentz-invariant.

It has long been observed that Lorentz invariance (2) implies the identity

$$E \times H = D \times B \, (=: P). \tag{5}$$

Actually, let us define

$$\sigma := \frac{1}{2}\left(|B|^2 - |E|^2\right), \quad \pi := E \cdot B.$$

Then

$$D = -(L_0)_\sigma E + (L_0)_\pi B, \quad H = -(L_0)_\sigma B - (L_0)_\pi E,$$

whence

$$E \times H = (L_0)_\sigma B \times E = D \times B.$$

Our first observation is that the converse is true:

Proposition 2.1 *Let us give ourselves a smooth function $(D, B) \mapsto h$, with the properties that*

1. *h is strictly convex with respect to D,*
2. *the identity (5) holds true, where $E := \frac{\partial h}{\partial D}$, $H := \frac{\partial h}{\partial B}$.*

Then the partial Legendre transform

$$L(B, E) := \sup_{D \in \mathbb{R}^3} (D \cdot E - h(D, B))$$

is Lorentz invariant.

Proof
By construction, we have

$$\frac{\partial L}{\partial B} = -H, \quad \frac{\partial L}{\partial E} = D.$$

Therefore the assumption (5) becomes a linear first-order differential system in L:

$$E \times \frac{\partial L}{\partial B} - B \times \frac{\partial L}{\partial E} = 0. \tag{6}$$

This system can be recast as $R_i L = 0$ for $1 \leq i \leq 3$, with

$$R_i := E_j \partial_{B_k} - E_k \partial_{B_j} - B_j \partial_{E_k} + B_k \partial_{E_j} \qquad (\epsilon(i, j, k) = 1).$$

The Lie algebra spanned by R_1, R_2 and R_3 is the vector space spanned by the R_j's and by S_1, S_2 and S_3, given by

$$S_i := B_j \partial_{B_k} - B_k \partial_{B_j} + E_j \partial_{E_k} - E_k \partial_{E_j} \qquad (\epsilon(i, j, k) = 1).$$

When $(B, E) \neq (0, 0)$ is fixed, the vectors $(R_1, \ldots, S_3)$ span a 4-dimensional vector space. This shows that a function L such that $R_1 L = \cdots = S_3 L \equiv 0$ is a function of two elementary solutions only. Since σ and π are functionally independent away from the origin, we deduce that L is a function of σ and π, whence the claim.

■

We drop from now on the Lagrangian $\mathcal{L}$ and work only with the energy density h, assuming the properties 1 and 2. Thanks to Proposition 2.1, we say that h is *Lorentz invariant*[3] if it satisfies

$$\frac{\partial h}{\partial D} \times \frac{\partial h}{\partial B} = D \times B. \tag{7}$$

We point out that the identities $S_j L \equiv 0$ write $E \times D + H \times B = 0$, that is

$$\frac{\partial h}{\partial D} \times D + \frac{\partial h}{\partial B} \times B = 0. \tag{8}$$

The identities (8) are thus consequences of (7), under the assumption that $D^2 h$ is non-singular. This derivation was not noted by Born [6], who sees (7,8) as the constraints for relativistic invariance.

What is interesting in (8) is that these are *linear* first-order PDEs in h. Since S_1, S_2 and S_3 form a Lie algebra of dimension three, the solutions of (8) are arbitrary functions of $6 - 3 = 3$ independent solutions, say of $\alpha := (\beta, \gamma, \delta)$ with

$$\beta := \frac{1}{2}|B|^2, \qquad \gamma := B \cdot D, \qquad \delta := \frac{1}{2}|D|^2.$$

Of course, the R_j's do not belong to the Lie algebra spanned by the S_i's and we still have to exploit (7). Writing that there exists a function $(\beta, \gamma, \delta) \mapsto m(\beta, \gamma, \delta)$ such that

$$h(B, D) = m\left(\frac{1}{2}|B|^2, B \cdot D, \frac{1}{2}|D|^2\right), \tag{9}$$

we find

$$\frac{\partial h}{\partial D} = m_\gamma B + m_\delta D, \qquad \frac{\partial h}{\partial B} = m_\beta B + m_\gamma D$$

and therefore

$$\frac{\partial h}{\partial D} \times \frac{\partial h}{\partial B} = (m_\beta m_\delta - m_\gamma^2) D \times B.$$

This yields the following conclusion, already given in [1]:

Proposition 2.2 *The energy density h is Lorentz invariant if, and only if, it is given by (9) for some function m satisfying the first-order PDE*

$$m_\beta m_\delta - m_\gamma^2 = 1. \tag{10}$$

[3] Of course, h is not a scalar invariant under the action of the Lorentz group, strictly speaking, since it is not a function of σ and π. We only mean that it is an energy density associated to a Lorentz invariant action density L.

Remark: Observe that if h is Lorentz-invariant, then $\hat{h}(\beta,\gamma,\delta) := h(\delta,\gamma,\beta)$ is invariant too, a fact that was not visible on the formula (34) of [16].

Let $X \in \mathbb{R}^3$ be a unit vector orthogonal to B and D. Since h is convex with respect to D, the map $s \mapsto h(B, D+sX) = m(\beta,\gamma,\delta+s^2/2)$ is convex. This implies

$$\left.\frac{d^2}{ds^2}\right|_{s=0} h(B, D+sX) = m_\delta \geq 0.$$

Using then (10), we deduce a seemingly new result. At least, it is not contained in the papers by Born & Infeld.

Proposition 2.3 *The symmetric matrix*

$$M_0(\alpha) := \begin{pmatrix} m_\beta & m_\gamma \\ m_\gamma & m_\delta \end{pmatrix}$$

is positive definite with unit determinant.

Remark. Contrary to some authors, we do not make an assumption of self-duality. Self-duality has been motivated in the past by the apparent symmetry of the roles played by (B, E) on the one hand, and by (D, H) on the other hand. This symmetry was very strong in the classical Maxwell system. However we do not see any physical reason why this symmetry would hold true in general. This symmetry is actually contradicted by the fact that one has never observed any magnetic charge: even in presence of matter, one still have $\mathrm{div} B = 0$, which is a part of $\mathrm{d}\omega = 0$. On the contrary, $\mathrm{div} D$ does not vanish in presence of matter. In terms of differential forms, there is no way to build a closed form with (D, H) in presence of matter, and therefore one cannot exchange the roles of (B, E) and of (D, H).

3 Integration of (10)

Since (10) is a Hamilton–Jacobi equation, it can be integrated by the method of characteristics. This gives us informations about the values of ∇m and $D^2 m$. We assume that h is a smooth function over $\mathbb{R}^6$. As a by-product, m is defined over the domain

$$\{\alpha = (\beta,\gamma,\delta) \in \mathbb{R}^3 \,|\, \beta > 0 \text{ and } 4\beta\delta - \gamma^2 \geq 0\},$$

and smooth in the interior. This domain can be identified with the cone of positive semi-definite 2×2 symmetric matrices by

$$\alpha \leftrightarrow \begin{pmatrix} 2\beta & \gamma \\ \gamma & 2\delta \end{pmatrix}.$$

As such, it is a *convex* subset of $\mathbb{R}^3$.

Equation (10) can be written as $\nabla m^T S \nabla m = 1$, where

$$S := \begin{pmatrix} 0 & 0 & 1/2 \\ 0 & -1 & 0 \\ 1/2 & 0 & 0 \end{pmatrix}.$$

A characteristic curve is an integral curve of the vector field $V := \frac{1}{2}(m_\delta, -2m_\gamma, m_\beta)^T = S\nabla m$. In the identification above, V is positive definite because of Proposition 2.3.

Let us denote by d/ds the corresponding derivative $V \cdot \nabla_\alpha$. Differentiating (10), we have

$$\frac{d}{ds}\nabla m = 0,$$

a property that writes also

$$D^2 m V \equiv 0. \tag{11}$$

Thus ∇m remains constant along a characteristics. This shows that characteristics are straight lines, of the form

$$s \mapsto \alpha(s) := \alpha_0 + sV(\alpha_0).$$

The values of s under consideration are those of the maximal interval $I(\alpha_0)$, containing $s = 0$ and for which $4\beta\delta - \gamma^2$ remains non-negative. Since α_0 is positive semi-definite and V is positive definite, one has $I(\alpha_0) = [s_+(\alpha_0), +\infty)$, where $s_+(\alpha_0)$ is the larger root of the polynomial

$$(4\beta\delta - \gamma^2 =)\, P(s) := s^2 + 2(\beta_0 m_\beta + \gamma_0 m_\gamma + \delta_0 m_\delta)_0 s + 4\beta_0\delta_0 - \gamma_0^2.$$

An elementary differentiation gives

Proposition 3.1 *Characteristic curves of Equation (10) are level curves of the relavistic invariants* $\sigma = \frac{1}{2}(|B|^2 - |E|^2)$ *and* $\pi = E \cdot B$.

We now write the ODE for $D^2 m$ along a characteristics. Differenting once more (10), we have

$$\frac{d}{ds}D^2 m + D^2 m S D^2 m = 0.$$

The solution of this ODE is given by the formula

$$D^2 m(\alpha(s)) = D^2 m(\alpha_0)(I_3 + sSD^2 m(\alpha_0))^{-1}. \tag{12}$$

The fact that m is a smooth function on the whole domain $4\beta\delta - \gamma^2 > 0$ tells us that the matrix $I_3 + sSD^2 m(\alpha_0)$ is non-singular for every s interior to $I(\alpha_0)$. In other words, the spectrum of $SD^2 m(\alpha_0)$ must not meet the real subset

$$(-\infty, 0) \cup (-1/s_+, +\infty).$$

We shall prove in the next paragraph that m is a concave function. Thus D^2m is negative semi-definite. Therefore the product SD^2m has real eigenvalues. The condition above is thus that

$$\mathrm{Sp}(SD^2m(\alpha)) \subset \left[0, -\frac{1}{s_+(\alpha_0)}\right]. \tag{13}$$

When $\mathrm{rk}D^2m(\alpha_0) = 1$, that is $D^2m(\alpha_0) = \epsilon ww^T$ with $\epsilon = \pm 1$, then the Hessian is given along the characteristic curve by

$$D^2m(\alpha(s)) = \frac{\epsilon}{1 + \epsilon s w^T S w} ww^T. \tag{14}$$

In particular, it remains of rank one. Likewise, if $D^2m(\alpha_0) = 0_3$, then $D^2m(\alpha(s)) \equiv 0_3$. As a consequence, we see that whatever the rank of D^2m at α_0, it remains constant along the caracteristic curve.

In formula (14), we shall see in Lemma 3.1 that w^TSw is negative. Since the formula makes sense for every $s > s_+(\alpha_0)$ and in particular for every $s > 0$, we must have $\epsilon = -1$. This proves $D^2m \leq 0_3$ in this special case. This is a part of the concavity result.

3.1 Convexity *vs* relativistic invariance

We now examine the spectrum of D^2h when $\alpha = (|B|^2/2, B \cdot D, |D|^2)$ runs along the characteristic line Γ passing through a point α_0. Our goal is to check whether or not h is a convex funcion of (B, D).

We have

$$D^2h(B, D) = M_0 \otimes I_3 + \sum_{\rho,\sigma \in \{\beta,\gamma,\delta\}} m_{\rho\sigma} X^\rho X^\sigma,$$

where the subscripts denote derivatives, and

$$X^\beta := \begin{pmatrix} B \\ 0 \end{pmatrix}, \; X^\gamma := \begin{pmatrix} D \\ B \end{pmatrix}, \; X^\delta := \begin{pmatrix} 0 \\ D \end{pmatrix}.$$

Let us assume that $P := D \times B \neq 0$, that is $4\beta\delta - \gamma^2 > 0$. Then $\mathbb{R}^6$ splits into $E_\perp \otimes^\perp E_\|$, with $E_\perp = \mathbb{R}P \times \mathbb{R}P$ and $E_\|$ spanned by

$$f^1 := X^\beta = \begin{pmatrix} B \\ 0 \end{pmatrix}, \; f^2 := \begin{pmatrix} D \\ 0 \end{pmatrix}, \; f^3 := \begin{pmatrix} 0 \\ B \end{pmatrix}, \; f^4 := X^\delta = \begin{pmatrix} 0 \\ D \end{pmatrix}.$$

Both subspaces $E_{\perp,\|}$ are invariant under D^2h. Thus the spectrum of D^2h is the union of the spectra of its restrictions to these subspaces. The restriction to $E_\perp$ has matrix M_0 in the obvious basis. From Proposition 2.3, it has two positive eigenvalues $\lambda_\pm$ such that $\lambda_-\lambda_+ = 1$ and $\lambda_- + \lambda_+ = m_\beta + m_\delta$.

In the basis $\{f^1, \dots, f^4\}$, the matrix of the restriction of D^2h to $E_\|$ is

$$M_1(\alpha) = M_0 \otimes I_2 + \begin{pmatrix} m_{\beta\beta} & m_{\beta\gamma} & m_{\beta\gamma} & m_{\beta\delta} \\ m_{\beta\gamma} & m_{\gamma\gamma} & m_{\gamma\gamma} & m_{\gamma\delta} \\ m_{\beta\gamma} & m_{\gamma\gamma} & m_{\gamma\gamma} & m_{\gamma\delta} \\ m_{\beta\delta} & m_{\gamma\delta} & m_{\gamma\delta} & m_{\delta\delta} \end{pmatrix} \begin{pmatrix} 2\beta & \gamma & 0 & 0 \\ \gamma & 2\delta & 0 & 0 \\ 0 & 0 & 2\beta & \gamma \\ 0 & 0 & \gamma & 2\delta \end{pmatrix}.$$

We notice that since this basis is not orthonormal, the resulting matrix is not symmetric. Its eigenvalues are nevertheless real, since they are eigenvalues of D^2h. We also warn the reader that the space $E_\|$ varies with (B, D).

Lemma 3.1 *As $s \to +\infty$, the matrix $D^2m(\alpha(s))$ behaves as follows (recall that $\mathrm{rk}D^2m \leq 2$ because of $D^2mV = 0$):*

Case $\mathrm{rk}D^2m(\alpha_0) = 2$**.** *We have*

$$D^2m(\alpha(s)) \sim \frac{1}{s}\left(S^{-1} - (S^{-1}V)(S^{-1}V)^T\right) = \frac{1}{s}\left(S^{-1} - \nabla m \otimes \nabla m\right),$$

where we notice that $V^TS^{-1}V = 1$ because of (10).

Case $\mathrm{rk}D^2m(\alpha_0) = 1$**.** *Let $w \in \mathbb{R}^3$ be such that $D^2m(\alpha_0) = \pm ww^T$. We have*

$$D^2m(\alpha(s)) \sim \frac{1}{s\, w^TSw} ww^T,$$

and $w^TSw < 0$.

Case $D^2m(\alpha_0) = 0_3$**.** *Then $D^2m(\alpha(s)) \equiv 0_3$.*

Proof
Just use (12) to get the formulæ. The matrix $D^2m(\alpha(s))$ is equivalent to $\frac{1}{s}\Sigma$, where Σ is symmetric, has the same kernel as $D^2m(\alpha_0)$ and otherwise behaves like S^{-1}. To see that w^TSw is negative in the second case, we argue as follows. If w^TSw was non-negative, that is $w_\beta w_\delta - w_\gamma^2 \geq 0$, the symmetric matrix

$$W := \begin{pmatrix} w_\delta & -w_\gamma \\ -w_\gamma & w_\beta \end{pmatrix}$$

would be semi-definite. However, (11) tells that $w \cdot V = 0$, which can be recast as $\mathrm{Tr}(WM_0) = 0$. Since M_0 is positive definite, this would imply $w = 0$, which negates $\mathrm{rk}D^2m = 1$.

■

Corollary 3.1 *As $s \to +\infty$, the matrix $M_1(\alpha(s))$ admits a limit $M_\infty(\alpha_0)$. In particular, the eigenvalues of $D^2h(B, D)$ admit limits, which are on the one hand those of M_0, and on the other hand those of $M_\infty(\alpha_0)$.*

Corollary 3.2 *The Hessian matrix D^2m is everywhere negative semi-definite. In other words, m is a concave function of α.*

Proof

It is enough to examine the cases $D^2m(\alpha_0) \neq 0_3$. The rank of $sD^2m(\alpha(s))$ does not change with $s \in [s_+(\alpha_0), \infty)$. According to Lemma 3.1, its limit as $s \to +\infty$,

$$\text{either } S^{-1} - \nabla m \otimes \nabla m, \text{ or } \frac{1}{w^T S w} w w^T,$$

is obviously negative semi-definite of same rank. Since the signature of a symmetric matrix can change only when its rank changes, we conclude that this signature is constant, equal to that of the limit. Thus $D^2m \leq 0_3$.

∎

Remark. The concavity of m is a consequence of the fact that m is a *global* smooth solution of (10) over the convex domain $4\beta\delta - \gamma^2 > 0$. We can restate Corollary 3.2 in the following way: *Every $\mathcal{C}^2$ solution of the Hamilton–Jacobi equation (10) over the convex cone $\beta > 0$, $4\beta\delta - \gamma^2 > 0$ is concave.*

We now analyze the eigenvalues of $M_\infty(\alpha_0)$.

Case $\mathrm{rk} D^2m(\alpha_0) = 2$**.** In this situation, M_∞ does not at all depend on the special form of $D^2m(\alpha_0)$, but only on the first derivatives of m, that are constant along Γ:

$$M_\infty(\alpha_0) = M_0 \otimes I_2 + \left(S^{-1} - \nabla m \otimes \nabla m\right)^\sharp \begin{pmatrix} m_\delta & -m_\gamma & 0 & 0 \\ -m_\gamma & m_\beta & 0 & 0 \\ 0 & 0 & m_\delta & -m_\gamma \\ 0 & 0 & -m_\gamma & m_\beta \end{pmatrix},$$

where the 'sharp' means that the second line and row have been repeated once, so that the resulting matrix is 4×4. Simplifications occur when performing the product and the sum. At the end, we have

$$M_\infty(\alpha_0) = \begin{pmatrix} 0_2 & K \\ -K & 0_2 \end{pmatrix}, \qquad K := \begin{pmatrix} -m_\gamma & m_\beta \\ -m_\delta & m_\gamma \end{pmatrix}.$$

From (10), we have $K^2 = -I_2$. Thus the eigenvalues of $M_\infty(\alpha_0)$ are ± 1, each one with multiplicity two.

In particular, $D^2h(B, D)$ has two negative eigenvalues when $\alpha = \alpha(s)$ and s is large enough.

Case $\mathrm{rk} D^2m(\alpha_0) = 1$**.** We have here

$$M_\infty(\alpha_0) = M_0 \otimes I_2 + \frac{1}{w^T S w} \left(w w^T\right)^\sharp \begin{pmatrix} m_\delta & -m_\gamma & 0 & 0 \\ -m_\gamma & m_\beta & 0 & 0 \\ 0 & 0 & m_\delta & -m_\gamma \\ 0 & 0 & -m_\gamma & m_\beta \end{pmatrix},$$

with the same notations as above. This is a rank-one perturbation of $M_0 \otimes I_2$. Since the eigenvalues $\lambda_\pm$ of the latter have multiplicity two, there follows that $\lambda_\pm$ are still eigenvalues of M_∞, though with multiplicity one only, in general.

There remains to identify the two other eigenvalues $\mu_\pm$. It is possible to prove that

$$\mathrm{Tr} M_\infty(\alpha_0) = m_\beta + m_\delta, \tag{15}$$

but this will not be needed here. Instead, we calculate the determinant. Using the formula $\det(A + wz^T) = (\det A)(1 + z^T A^{-1} w)$ and the fact that $\det M_0 = 1$, we obtain

$$\det M_\infty = 1 + \frac{1}{m_{\beta\delta} - m_{\gamma\gamma}} \left(4V^T D^2 m V + 2(m_{\beta\delta} - m_{\gamma\gamma})(m_\beta m_\delta - m_\gamma^2)\right).$$

With (10) and (11), there remains

$$\det M_\infty(\alpha_0) = -1. \tag{16}$$

We conclude that $\mu_- \mu_+ = -1$: these eigenvalues have opposite signs. Again, $D^2 h(B, D)$ has a negative eigenvalue when $\alpha = \alpha(s)$ and s is large enough.

Since a negative eigenvalue of $D^2 h$ is an indication that convexity fails, we see that h can be convex in (B, D) only if $D^2 m(\alpha_0) \equiv 0_3$. We thus obtain the following result:

Theorem 3.1 *If h is Lorentz invariant and convex over $\mathbb{R}^6$, then h is quadratic, of the form*

$$h = \frac{a_0}{2}|B|^2 + a_1 B \cdot D + \frac{a_2}{2}|D|^2.$$

This important result justifies *a posteriori* the analysis of non-linear models of Maxwell's equations when the energy density is not convex, but only poly-convex in the sense of Ball. See [16] for a general context and [8] for the Born–Infeld model. We recall that the latter is defined, after [7], by the energy density

$$h_{BI}(B, D) = -1 + \sqrt{1 + |B|^2 + |D|^2 + |D \times B|^2}.$$

4 Wave velocities

The system (1) is a first-order system of conservation laws. Its Cauchy problem is linearly well-posed if, and only if it is hyperbolic. This means that for every direction $\xi \in \mathbb{R}^3$, the symbol matrix $A(\xi)$ of the linearized problem

$$\partial_t U + \sum_{\alpha=1}^{3} A^\alpha(B, D)\partial_\alpha U = 0, \qquad A(\xi) := \sum_\alpha \xi_\alpha A^\alpha(B, D),$$

has to have real eigenvalues and is diagonalizable. We recall (see [16]) that this holds true at least when h is strictly polyconvex. Actually, the following stronger result is valid:

Proposition 4.1 *Strict rank-one convexity of h, in the sense that $X^T D^2 h X$ is positive for every non-zero rank-one matrix $X \in \mathbb{R}^6 \sim \mathbf{M}_{3\times 2}(\mathbb{R})$, is a sufficient condition for hyperbolicity.*

Proof

Hyperbolicity is a property to be checked for planar waves. Without loss of generality, we consider waves in the direction of x_1. The system thus reduces to

$$\partial_t B_2 - \partial_x \frac{\partial h}{\partial D_3} = 0, \qquad \partial_t B_3 + \partial_x \frac{\partial h}{\partial D_2} = 0$$
$$\partial_t D_2 + \partial_x \frac{\partial h}{\partial B_3} = 0, \qquad \partial_t D_3 - \partial_x \frac{\partial h}{\partial B_2} = 0.$$

The symbol of the linearized system is the matrix $A := J_1 D^2 h$ where we understand the derivatives with respect to $B_2, \dots, D_3$ only, and

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

Strict rank-one convexity is the fact that the restriction of $D^2 h$ over the cone defined by $\det X = 0$ (rank-one 2×2 matrices) is positive. It is well-known that this is equivalent to the existence of a number a such that

$$X \mapsto X^T D^2 h X + \frac{a}{2} \det X$$

is positive definite. This amounts to saying that the matrix $D^2 h - a J_1$ is positive definite. Then the product with J_1 is diagonalizable with real eigenvalues (of same signs as those of J_1). Finally,

$$A = J_1 (D^2 h - a J_1) + a I_4$$

is diagonalizable with real eigenvalues.

∎

The proof above suggests that the relevant symbol matrix must be 4×4 instead of 6×6. This is reminiscent to the constraints $\mathrm{div} B = \mathrm{div} D = 0$. For every $\xi \in \mathbf{S}^2$, $A(\xi)$ has an invariant subspace $\Pi(\xi) \subset \mathbb{R}^6$ of dimension four, namely the space $\xi^\perp \times \xi^\perp$. What we really need is thus that the restriction of $A(\xi)$ to $\Pi(\xi)$ be diagonalizable with real eigenvalues. The latter will be denoted by $\lambda_1(B, D) \leq \cdots \leq \lambda_4(B, D)$. The expressions $\lambda_j(\xi)/|\xi|$ are the velocities $c_j(\xi)$ of infinitesimal waves. In the classical, linear isotropic, Maxwell's equations, one has $c_{1,2}(B, D) = -1$ and $c_{3,4}(B, D) = 1$.

For more general, non-linear models, we postulate that the wave velocities belong to the interval $[-1, 1]$: *No information should travel faster than the speed defined by the fundamental Lorentz form.* We say that such waves are *causal* (Boillat says *sub-luminal*). This is the analogue of the causality condition in kinematics, mentionned above. A fundamental problem is thus to describe as accurately as possible the Lorentz invariant energy densities that satisfy this postulate.

4.1 The Born–Infeld model

The Born–Infeld model is given by the Lorentz invariant energy

$$h_{BI}(B,D) := \sqrt{1+|B|^2+|D|^2+|D\times B|^2}.$$

It has been studied in details by many authors. Brenier [8] found a way to embed it in a larger system of ten equations with ten unknowns, providing an easy calculation of the eigenvalues:

$$\frac{p}{h}, \quad \frac{p\pm Z}{h}, \qquad \text{where } p:=\det(\xi,D,B) \text{ and } Z:=\sqrt{|\xi|^2+(B\cdot\xi)^2+(D\cdot\xi)^2}\,.$$

The first of these three numbers is irrelevant; it corresponds to eigenvectors that do not belong to $\Pi(\xi)$. We thus have

$$\lambda_1=\lambda_2=\frac{p-Z}{h}, \qquad \lambda_3=\lambda_4=\frac{p+Z}{h}.$$

We notice that these eigenvalues are not necessarily of opposite signs. This phenomenon arises every time a Lorentz-invariant model has wave velocities distinct from ± 1, since then a Lorentz transformation can bring them to any subinterval of $(-1,+1)$.

The causality condition that $-|\xi|\le\lambda_1$ and $\lambda_4\le|\xi|$ is then equivalent to

$$|p|+Z\le|\xi|h. \tag{17}$$

We do not know whether the property (17) has ever been checked. At least it was not in Brenier's paper. We now prove it.

Proposition 4.2 *In the Born–Infeld model, the wave velocities are bounded by that of the fundamental form. In other words, the wave propagation is causal: (17) is satisfied.*

We shall see however from the proof that relativistic propagation (thus at unit speed) may occur.

Proof

We must verify inequality (17). It is equivalent to show that

$$p^2+Z^2+2\epsilon pZ\le|\xi|^2h^2, \qquad \epsilon=\pm 1.$$

Without loss of generality, we fix $\xi=\vec{e}_1$. The inequality to be proved simplifies into

$$0\le -2\epsilon(D_2B_3-D_3B_2)Z+B_2^2+B_3^2+D_2^2+D_3^2+(D_1B_2-D_2B_1)^2+(D_1B_3-D_3B_1)^2. \tag{18}$$

Let us take $\epsilon=+1$, meaning that we are dealing with $\lambda_{3,4}$. The right-hand side of (18) is nothing but

$$(b^T,d^T)S\begin{pmatrix} b\\ d\end{pmatrix}, \qquad \text{with } b:=\begin{pmatrix} B_2\\ B_3\end{pmatrix},\ d:=\begin{pmatrix} D_2\\ D_3\end{pmatrix}.$$

The matrix S is given as

$$S=\begin{pmatrix} (1+D_1^2)I_2 & s\\ s^T & (1+B_1^2)I_2\end{pmatrix}, \qquad s:=\begin{pmatrix} -B_1D_1 & Z\\ -Z & -B_1D_1\end{pmatrix}.$$

The matrix s is a direct similitude of homothety ratio $\sqrt{(1+D_1^2)(1+B_1^2)}$. The symmetric matrix S is thus positive semi-definite and the condition (18), thus also (17), is satisfied.

The case $\epsilon = -1$ is similar.

■

Because $ss^T = (1+D_1^2)(1+B_1^2)I_2$, S is only semi-definite. It is actually of rank two only. We thus have an equality case in (17). For $\epsilon = +1$, it is given by

$$(1+D_1^2)b + sd = 0, \tag{19}$$

or equivalently $s^T b + (1+B_1^2)d = 0$. Of course this is the equality case for $\xi = \vec{e}_1$ only. We thus see that for every $\xi \in \mathbf{S}^2$ (unit vectors of $\mathbb{R}^3$), the velocity $\lambda_{3,4}(\xi)$ equals one for every state (B, D) in a submanifold $\mathcal{F}(\xi)$ of codimension two in $\mathbb{R}^6$.

For a general unit vector ξ, the condition (19) translates as

$$\sqrt{1+(D\cdot\xi)^2}\, b = \sqrt{1+(B\cdot\xi)^2}\, R_\xi(\theta) d, \qquad b := B\times\xi,\ d := D\times\xi, \tag{20}$$

with $R_\xi(\theta)$ the rotation of angle θ in the plane $\xi^\perp$, θ being given by

$$\cos\theta = \frac{(B\cdot\xi)(D\cdot\xi)}{\sqrt{(1+(B\cdot\xi)^2)(1+(D\cdot\xi)^2)}}, \qquad \sin\theta = -\frac{\sqrt{1+(B\cdot\xi)^2+(D\cdot\xi)^2}}{\sqrt{(1+(B\cdot\xi)^2)(1+(D\cdot\xi)^2)}}.$$

4.2 Other models

For a general, Lorentz invariant model, with energy $h = m(\beta, \gamma, \delta)$, the matrix symbol is given by

$$A(\xi) = J(\xi) D^2 h, \qquad J(\xi) := \begin{pmatrix} 0_3 & j(\xi) \\ -j(\xi) & 0_3 \end{pmatrix}, \qquad j(\xi) = \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix}.$$

The role of $j(\xi)$ is to describe the product of vectors: $j(\xi)v = \xi \times v$. Because of the special form of h, we split $A(\xi)$ into two parts:

$$A(\xi) = A_0(\xi) + J(\xi) \sum_{\rho,\sigma} m_{\rho\sigma} X^\rho (X^\sigma)^T,$$

with $A_0(\xi)(M_0 \otimes I_3)$. The following result is elementary:

Lemma 4.1 *For every unit vector $\xi \in \mathbf{S}^2$, the spectrum of $A_0(\xi)$ consists in $\{-1, 0, 1\}$, each eigenvalue being semi-simple of multiplicity two. The eigenspace associated to $+1$ is the space of vectors*

$$\begin{pmatrix} -m_\delta v \\ (m_\gamma + j(\xi))v \end{pmatrix}, \qquad \forall v \perp \xi.$$

Lemma 4.1 tells us nothing but the fact that for a *linear*, Lorentz invariant model, the wave velocities equal one in every direction. We are now interested in the wave velocities for an arbitrary Lorentz invariant model. In general, these velocities are not identically equal to one (see for instance the Born–Infeld case). We have however the following result.

Theorem 4.1 *Let h be Lorentz invariant. Then, for every given state (B, D), there exists two directions $\xi \in \mathbf{S}^2$ (in some degenerate cases, they may coincide), such that one of the wave velocities in the direction ξ be equal to one.*

To prove this result, we need the following lemma.

Lemma 4.2 *Given $(B, D) \in \mathbb{R}^6$, there exists $w, \xi \in \mathbf{S}^2$ such that*

$$v \cdot B = v \cdot \xi = 0, \quad \det(\xi, v, D) = 0, \quad v \cdot D + \det(v, \xi, B) = 0.$$

Proof

This statement is invariant under isometries. We thus assume that $B = b\vec{e_1}$. If $b = 0$, just take ξ parallel to D and v orthogonal. Otherwise, $b \neq 0$ and we write

$$D = \begin{pmatrix} D_1 \\ d \end{pmatrix}, \qquad v = \begin{pmatrix} 0 \\ w \end{pmatrix}, \qquad \xi = \begin{pmatrix} \xi_1 \\ \eta \end{pmatrix}.$$

We wish to solve

$$w \cdot \eta = 0, \quad \xi_1 w \times d + D_1 \eta \times w = 0, \quad w \cdot d + bw \times \eta = 0, \quad |\xi| = 1.$$

Since this problem is homogeneous in w, we only need to find a solution with $w \neq 0$.

For every $w \in \mathbb{R}^2$, we define

$$f(w) := b^2 (w \times d)^2 |w|^2 - (w \cdot d)^2 \left((w \times d)^2 + D_1^2 |w|^2 \right).$$

Let us assume first that $d \neq 0$. From $f(d) = -|d|^6 D_1^2 \leq 0$ and $f(w) = b^2 |w|^4 |d|^2 \geq 0$ if $w \perp d$, and from the connectedness of $\mathbf{S}^1$, we deduce the existence of a non-zero w such that $f(w) = 0$. We build $\eta \in \mathbb{R}^2$ in the following way: It is a vector orthogonal to w, with norm

$$|\eta| = \frac{|w \times d|}{\sqrt{(w \times d)^2 + D_1^2 |w|^2}} \quad \left(= \frac{|w \cdot d|}{|b|\,|w|} \right).$$

There remains to fix the orientation of η. By construction, and because of $f(w) = 0$, we have $(w \cdot d)^2 = b^2 (\eta \times w)^2$ (notice that $|\eta \times w| = |\eta|\,|w|$ because of orthogonality). We thus choose η in such a way that

$$w \cdot d = b \eta \times w.$$

If $D_1 \neq 0$, we have $f(d) < 0$ and therefore $w \times d \neq 0$. We may define

$$\xi_1 := -D_1 \frac{\eta \times w}{w \times d}.$$

We then have

$$|\xi|^2 = \xi_1^2 + |\eta|^2 = |\eta|^2 \left(D_1^2 \frac{|w|^2}{(w \times d)^2} + 1 \right) = |\eta|^2 \frac{b^2 |w|^2}{(w \cdot d)^2} = 1,$$

where we have used twice the property $f(w) = 0$.

Thus v, ξ solve our problem.

There remains to investigate the special cases. Let us begin with $d \neq 0$, $D_1 = 0$. Then

$$f(w) = (w \times d)^2 \left(b^2|w|^2 - (w \cdot d)^2\right).$$

There are two subcases, depending on the sign of $|b| - |d|$:

$|d| \leq |b|$. One sets $w = d$ and define η by $\eta \perp w$ with $|\eta| = |d|/|b| (\leq 1)$. Then we take $\xi_1 = \pm\sqrt{1 - |\eta|^2}$.

$|b| \leq |d|$. There exists a $w \in \mathbf{S}^2$ such that $w \cdot d = b|w|$. One then sets $\eta \perp w$ with $\eta \in \mathbf{S}^1$, and $\xi_1 = 0$.

The final case where $d = 0$ is the simplest one. Just take $\eta = 0$, $\xi_1 = \pm 1$ and $w \neq 0$.

■

We remark that the equation $f = 0$ admits two independent solutions in general. In coordinates $(x, y) := (w \times d, w \cdot d)$, it is a bi-quadratic equation

$$b^2 x^4 + (b^2 - 1 - D_1^2) x^2 y^2 - D_1^2 y^4 = 0.$$

There are two roots for x^2/y^2, of opposite signs. Only the positive root may be used. But then there are two opposite values for the slope x/y. The only case where both solutions coincide is when $D_1 = 0$, meaning that $B \cdot D = 0$. When $d = 0$ (meaning that $B \times D = 0$), there is only one direction $\xi \parallel B$, for which the velocity one is of double multiplicity.

We now proceed to the proof of Theorem 4.1.

Proof

Let us define $D' := m_\gamma B + m_\delta D$. We apply Lemma 4.2 to the pair (B, D'). There exists $v, \xi \in \mathbf{S}^2$ such that

$$v \cdot B = v \cdot \xi = 0, \quad \det(\xi, v, D') = 0, \quad v \cdot D' + \det(v, \xi, B) = 0.$$

After elementary manipulations, this is equivalent to

$$v \cdot B = v \cdot \xi = 0, \quad m_\gamma v \cdot D + \det(\xi, v, D) = 0, \quad m_\delta v \cdot D + \det(v, \xi, B) = 0.$$

Let us define a vector in $\mathbb{R}^6$ by

$$R := \begin{pmatrix} -m_\delta v \\ (m_\gamma + j(\xi))v \end{pmatrix}.$$

On the one hand, $v \cdot \xi = 0$ tells us that $A_0(\xi)R = R$. On the other hand, the other identities above tell us that

$$R \perp X^\beta,\ X^\gamma,\ X^\delta,$$

whence R belongs to the kernel of

$$\sum_{\rho,\sigma} m_{\rho\sigma} X^\rho (X^\sigma)^T.$$

Finally, we obtain $A(\xi)R = R$, as desired.

■

Remarks.

- Theorem 4.1 shows in particular that Proposition 4.2 is sharp: the union of the manifolds $\mathcal{F}(\xi)$ as ξ runs over $\mathbf{S}^2$ covers $\mathbb{R}^6$.
- One finds in general two directions in which $c_j(\xi) = 1$. They coincide when $m_\gamma B + m_\delta D$ is either parallel or orthogonal to B, that is if either $D \parallel B$ or $2\beta m_\gamma + \gamma m_\delta = 0$.
- These directions ξ and the corresponding eigenvectors R do not depend at all of the Hessian $D^2 m$.

The eigenform. Let $\xi_0 \in \mathbf{S}^2$ be a direction as in Theorem 4.1, and R_0 the eigenvector: we have $A(\xi_0)R_0 = R_0$, that is

$$J(\xi_0)D^2hR_0 = R_0. \tag{21}$$

Since R_0 belongs to the range of $J(\xi_0)$, the vector $S_0 := J(\xi_0)R_0$ has the property that $J(\xi_0)S_0 = R_0$. In addition, we have

$$D^2hR_0 = (M_0 \otimes I_3)R_0 = \begin{pmatrix} -v + m_\gamma \xi_0 \times v \\ m_\delta \xi_0 i \times v \end{pmatrix},$$

where the right-hand side belongs to $\xi^\perp \times \xi^\perp$, that is to the range of $J(\xi_0)$. Multiplying (21) by $J(\xi_0)$, we thus obtain

$$D^2hJ(\xi_0)S_0 = S_0.$$

In other words, we have $A(\xi_0)^T S_0 = S_0$. We deduce that the eigenform associated to the unit eigenvalue is nothing but S_0^T, that is $L_0 := R_0^T J(\xi_0)$.

We point out that

$$L_0R_0 = R_0^T J(\xi_0)R_0 = \begin{pmatrix} -m_\delta v \\ m_\gamma v + \xi_0 \times v \end{pmatrix} \cdot \begin{pmatrix} m_\gamma \xi \times v + \xi_0 \times (\xi_0 \times v) \\ m_\delta \xi_0 \times v \end{pmatrix} = 2m_\delta |\xi_0 \times v|^2 = 2m_\delta |v|^2$$

is positive.

4.3 First- and second-order conditions

Let $\xi_0 \in \mathbf{S}^2$ and R_0 be as above. We assume for the sake of simplicity that $+1$ is a simple eigenvalueat ξ_0. Therefore the symbol $A(\xi)$ admits a simple eigenvalue $\lambda(\xi)$ for every ξ in some neighbourhood of ξ_0, with the properties that $\xi \mapsto \lambda(\xi)$ is analytic and $\lambda(\xi_0) = 1$. Likewise, there is an analytic eigenfield $R(\xi)$ with $R(\xi_0) = R_0$, and an analytic eigenform $L(\xi)$ such that $L(\xi_0) = L_0$.

Since we wish that the wave velocities remain in the interval $[-1, 1]$ for every direction $\xi \in \mathbf{S}^2$, we ask whether the inequality $\lambda(\xi) \leq |\xi|$ remains valid. If so, then we must have the first- and second-order conditions

$$\begin{aligned} d\lambda(\xi_0) &= \xi_0, & (22)\\ D^2\lambda(\xi_0)\theta \otimes \theta &\leq |\theta|^2 - (\xi_0 \cdot \theta)^2. & (23) \end{aligned}$$

Proposition 4.3 *The first-order condition (22) is trivially satisfied.*

Proof
Differentiating either of the identities

$$(A(\xi) - \lambda(\xi)I_6)R(\xi) = 0, \qquad L(\xi)(A(\xi) - \lambda(\xi)I_6) = 0$$

and using the other one, yields classically

$$d\lambda(\xi_0) = \frac{L(\xi_0)A(\theta)R(\xi_0)}{L(\xi_0)R(\xi_0)}.$$

Proving (22) thus amounts to proving that $L(\xi_0)A(\theta)R(\xi_0) = 0$ for every $\theta \perp \xi_0$. For such directions θ, we have

$$A(\theta)R(\xi_0) = J(\theta)(M_0 \otimes I_3)R(\xi_0) = \begin{pmatrix} m_\delta(\theta \cdot v)\xi_0 \\ \theta \times v - m_\gamma(\theta \cdot v)\xi_0 \end{pmatrix},$$

where we have used that $D^2hR(\xi_0) = (M_0 \otimes I_3)R(\xi_0)$. We observe that the right-hand side belongs to $\xi_0 \times \xi_0$ (recall that $v \perp \xi_0$), the kernel of $J(\xi_0)$. Since $L(\xi_0) = R_0^T J(\xi_0)$, we deduce that $L(\xi_0)A(\theta)R(\xi_0) = 0$, as promissed.

$\blacksquare$

Since the first-order condition does not give any information, we have to content ourselves with the second-order one (23). We compute $D^2\lambda$ in two steps. We begin with $dL(\xi_0)\theta$, which is the solution of

$$(dL(\xi_0)\theta)(A(\xi_0) - I_6) + L(\xi_0)(A(\theta) - (\xi_0 \cdot \theta)I_6) = 0. \tag{24}$$

This solution is unique up to the addition of $\kappa L(\xi_0)$. Then we have

$$(LR)(\xi_0)D^2\lambda(\xi_0)\theta \otimes \theta = (dL(\xi_0)\theta)(A(\theta) - (\xi_0 \cdot \theta)I_6)R(\xi_0), \tag{25}$$

where the right-hand side does not depend on the choice of the solution of (24), since we know that $L(\xi_0)(A(\theta) - (\xi_0 \cdot \theta)I_6) \equiv 0$. Since LR is positive, our second-order condition (23) can be written

$$(dL(\xi_0)\theta)(A(\theta) - (\xi_0 \cdot \theta)I_6)R(\xi_0) \le (LR)(\xi_0)|\theta|^2 - (\xi_0 \cdot \theta)^2,$$

for every $\theta \in \mathbb{R}^3$. Because both factors $dL(\xi_0)\theta$ and $(A(\theta) - (\xi_0 \cdot \theta)I_6)R(\xi_0)$ vanish when θ is colinear with ξ_0, this is equivalent to

$$(\theta \perp \xi_0) \Longrightarrow \left((dL(\xi_0)\theta)(A(\theta) - (\xi_0 \cdot \theta)I_6)R(\xi_0) \le (LR)(\xi_0)|\theta|^2\right). \tag{26}$$

This property is presumably a restriction about D^2m, because $dL(\xi_0)$ must depend on it. Curiously, the dependence is almost certainly nonlinear.

4.4 Waves propagating along $D \times B$

Let us assume that $D \times B \neq 0$. We consider the propagation in the direction of $D \times B = P$, namely we choose $\xi = \pm P/|P|$. We recall that the relevant eigenvalues are those of the restriction of $A(\xi)$ to the subspace $\Pi(\xi) = \xi^\perp \times \xi^\perp$. In the present case, this space is what we denoted $E_\parallel$ in Paragraph 3.1. We thus can use the matrix $M_1(\alpha)$. In the basis $\{f^1, \dots, f^4\}$, the restriction of $A(\xi)$ is given by the 4×4 matrix

$$\tilde{A} = \tilde{J} M_1(\alpha),$$

where $\tilde{J}$ is the matrix of $(X, X') \mapsto (\xi \times X', -\xi \times X)$ in the same basis. A straightforward calculation gives

$$\tilde{J} = J_1(\alpha) = \frac{1}{\sqrt{4\beta\delta - \gamma^2}} \begin{pmatrix} 0 & 0 & \gamma & 2\delta \\ 0 & 0 & -2\beta & -\gamma \\ -\gamma & -2\delta & 0 & 0 \\ 2\beta & \gamma & 0 & 0 \end{pmatrix}.$$

We deduce that $\tilde{A}$ depends only upon α: $\tilde{A} = A_1(\alpha)$.

We now move $\alpha = \alpha(s)$ along a characteristic curve passing through α_0. We recall that when $s > s_+(\alpha_0)$, then $P \neq 0$. Thus the calculations made above apply. When $s \to +\infty$, $M_1(\alpha(s))$ admits a limit $M_\infty(\alpha_0)$. Since $\alpha(s) = \alpha_0 + sV$, the matrix $J_1(\alpha(s))$ admits a limit too, equal to

$$J_\infty(\alpha_0) = \begin{pmatrix} 0 & 0 & -m_\gamma & m_\beta \\ 0 & 0 & -m_\delta & m_\gamma \\ m_\gamma & -m_\beta & 0 & 0 \\ m_\delta & -m_\gamma & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0_2 & -K^{-1} \\ K^{-1} & 0_2 \end{pmatrix}.$$

Therefore the matrix $\tilde{A}$ admits a limit $A_\infty(\alpha_0) = J_\infty(\alpha_0) M_\infty(\alpha_0)$.

When $\mathrm{rk} D^2 m(\alpha_0) = 2$**,** we have

$$A_\infty(\alpha_0) = I_4.$$

It reveals the remarkable fact that for large fields, all the wave velocities in the direction of P tend to $+1$. This was obvious for the Born–Infeld model, according to Brenier's formula for the wave speeds.

If the model is causal, then the spectrum of $A_1(\alpha(s))$ must belong to $[-1, 1]$. Expanding this matrix in terms of s^{-1}, we find the necessary condition that the spectrum of the coefficient of s^{-1} must be real and non-positive.

When $\mathrm{rk} D^2 m(\alpha_0) = 1$**,** we have

$$A_\infty(\alpha_0) = J_\infty(\alpha_0) \left\{ M_0 \otimes I_2 + \frac{1}{w_\beta w_\delta - w_\gamma^2} \hat{w}\hat{w}^T \begin{pmatrix} m_\delta & -m_\gamma & 0 & 0 \\ -m_\gamma & m_\beta & 0 & 0 \\ 0 & 0 & m_\delta & -m_\gamma \\ 0 & 0 & -m_\gamma & m_\beta \end{pmatrix} \right\}.$$

This is a rank-one perturbation of the matrix

$$A_0 := J_\infty(\alpha_0) M_0 \otimes I_2,$$

which satisfies $A_0^2 = I_2$. The eigenvalues of A_0 are ± 1, which are semi-simple with multiplicity two. We deduce that $A_\infty(\alpha_0)$ still has the eigenvalues ± 1, although their multiplicities might drop to one. Its spectrum has the form $\{-1, \lambda, \mu, 1\}$. In order to find λ and μ, it is enough to compute the trace and the determinant of $A_\infty(\alpha_0)$. One has on the one hand

$$\det A_\infty(\alpha_0) = \det J_\infty(\alpha_0) \,\det M_\infty(\alpha_0) = \det M_\infty(\alpha_0) = -1,$$

according to (16). On the other hand, there holds

$$\mathrm{Tr} A_\infty(\alpha_0) = \mathrm{Tr} A_0 + \frac{1}{w_\beta w_\delta - w_\gamma^2} \hat{w}^T \begin{pmatrix} m_\delta & -m_\gamma & 0 & 0 \\ -m_\gamma & m_\beta & 0 & 0 \\ 0 & 0 & m_\delta & -m_\gamma \\ 0 & 0 & -m_\gamma & m_\beta \end{pmatrix} J_\infty(\alpha_0) \hat{w}.$$

Thanks to (10), and because A_0 is traceless, this gives

$$\mathrm{Tr} A_\infty(\alpha_0) = \frac{1}{w_\beta w_\delta - w_\gamma^2} \hat{w}^T \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \hat{w} = 2.$$

Finally, we have $\lambda + \mu = 2$ and $\lambda\mu = 1$, thus $\lambda = \mu = 1$. The matrix $A_\infty(\alpha_0)$ has the eigenvalues -1 (simple) and $+1$ (triple).

We summarize these calculations in the following statement.

Theorem 4.2 *Assume that $P := D \times B$ is non-zero. Then the wave speeds $c_1, \dots, c_4$ in the direction $\xi = P/|P|$ depend only on α.*

Along a characteristic curve $s \mapsto \alpha(s)$ (one has $P \neq 0$ for $s > s_+(\alpha_0)$), these velocities have limits when $s \to +\infty$, described as follows:

Case $\mathrm{rk} D^2 m(\alpha_0) = 2$**.** *Then*

$$\lim_{s \to +\infty} c_j = 1, \qquad \forall j = 1, \dots, 4.$$

Case $\mathrm{rk} D^2 m(\alpha_0) = 1$**.** *Then*

$$\lim_{s \to +\infty} c_1 = -1 \quad \text{and} \quad \lim_{s \to +\infty} c_j = 1, \quad \forall j = 2, \dots, 4.$$

Case $D^2 m(\alpha_0) = 0_3$**.** *Then*

$$c_1 = c_2 \equiv -1, \qquad c_3 = c_4 \equiv 1.$$

In other words, the velocities tend to ± 1. The number of those tending to $+1$ is $2 + \mathrm{rk} D^2 m(\alpha_0)$.

Remark. In the mildly nonlinear case $\mathrm{rk}D^2m(\alpha_0) = 1$, we deduce that for large s, the symbol $A(\xi)$ admits a simple eigenvalue for $\xi = P/|P|$. This situation is thus not suitable for the application of Boillat's Theorem in [3], which tells us that systems of conservation laws with multiple characteristics are linearly degenerate. If we wish to have double eigenvalues for every state and every direction, the rank of D^2m must be everywhere equal to either 0 or 2. The Lorentz invariant models with double eigenvalues (meaning that birefringence does not occur) have been characterized by Boillat [2] and independently by Plebanski [14]: the Born–Infeld and the linear systems are the only non-birefringent models, besides a singular one, given by the action density $L = -\sigma/\pi$.

4.5 Asymptotic transport and vanishing dispersion

As metionned above, the relevant eigenvalues of the symbol $A(\xi)$ are those associated to its restriction to the invariant subpace $\Pi(\xi) = \xi^\perp \times \xi^\perp$. We denote $Q_{B,D}(X;\xi)$ the characteristic polynomial of this restriction. Since $X^2 Q_{B,D}(X;\xi) = \det(XI_6 - A(\xi))$, where the right-hand side is divisible by X^2, we see that $Q_{B,D}$ is a polynomial in all its arguments X and ξ, homogeneous of degree 4.

If the model is hyperbolic and causal, $Q_{B,D}(X,\xi)$ has the property that for every ξ, the roots $X_j(\xi)$ belong to $[-|\xi|, |\xi|]$. Let us decompose

$$Q_{B,D}(X,\xi) = X^4 - \sigma_1(\xi)X^3 + \sigma_2(\xi)X^2 - \sigma_3(\xi)X + \sigma_4(\xi).$$

Each polynomial σ_k is homogeneous of degree k. It is nothing but the elementary symmetric polynomial of degree k in the roots X_j. Therefore our assumption implies

$$\sigma_k(\xi) \leq \binom{4}{k} |\xi|^k, \qquad \forall \xi \in \mathbb{R}^3, \tag{27}$$

This shows that $Q_{B,D}$ remains in a compact set as (B,D) varies in $\mathbb{R}^6$. In particular, this polynomial has cluster point(s) Q_∞ as $s \to +\infty$ along an α-characteristic curve. Such cluster points are still homogeneous of degree four and hyperbolic in the direction of X. They still ave the property that

$$X_j(\xi) \in [-|\xi|, |\xi|], \qquad \forall j = 1, \cdots, 4, \quad \forall \xi \in \mathbb{R}^3. \tag{28}$$

We focus now on the case $\mathrm{rk}D^2m(\alpha_0) = 2$. According to Theorem 4.2, Q_∞ has the property that for $\xi_0 := P/|P|$,

$$X_j(\xi_0) = |\xi_0| \ (= 1), \qquad \forall j = 1, \ldots, 4. \tag{29}$$

Lemma 4.3 *Assume* $\mathrm{rk}D^2m(\alpha_0) = 2$ *and set* $\xi_0 := P/|P|$, *where* $P = D \times B$. *Then*

$$Q_\infty(X,\xi) = (X - \xi_0 \cdot \xi)^4. \tag{30}$$

In particular, $Q_{B,D}$ *has a unique limit when* $s \to +\infty$ *along an* α*-characteristic curve.*

For the proof, we received help from J.-Y. Welschinger.

Proof

Let $\xi \in \mathbb{R}^3$ be given with $\xi \neq 0$ and $\xi \perp \xi_0$. We know from [12] that we can label the roots X_j in such a way that the one-parameter functions $\tau \mapsto X_j(\xi_0 + \tau\xi) =: g_j(\tau)$ be analytic (in general, this is not the natural ordering). By assumption, we have

$$|g_j(\tau)| \leq \sqrt{1 + \tau^2|\xi|^2}.$$

Since we also have $g_j(0) = 1$, according to Theorem 4.2, we deduce that $g_j'(0) = 0$. In other words, $g_j(\tau) - 1 = O(\tau^2)$.

Finally, we write

$$Q_\infty(1; \xi_0 + \tau\xi) = \prod_{j=1}^{4}(1 - g_j(\tau)).$$

The left-hand side is a polynomial in τ of degree at most $n = 4$, while the right-hand side is an $O(\tau^{2n})$. We conclude that

$$Q_\infty(1; \xi_0 + \tau\xi) \equiv 0, \qquad \forall \xi \perp \xi_0.$$

By homogeneity, this means

$$Q_\infty(\xi_0 \cdot \xi; \xi) \equiv 0.$$

In other words, Q_∞ is divisible by $X - \xi_0 \cdot \xi$. Applying the same argument to the quotient $Q_\infty(X;\xi)/(X - \xi_0 \cdot \xi)$, and proceeding by induction, we conclude that

$$Q_\infty(X, \xi) = (X - \xi_0 \cdot \xi)^4.$$

■

The meaning of (30) is that at very high fields, the system is essentially a transport equation at velocity $\xi_0 = P/|P|$. Since the velocities X_j are linear in ξ ($X_j(\xi) = \xi_0 \cdot \xi$), wave dispersion disappears in the high-field limit.

4.6 A potential application.

Theorem 4.2 and Paragraph 4.5 are interesting from a practical point of view. Non-linear models for electrodynamics are considered because the linear classical Maxwell model is inconsistent at a charged particle, say an electron. One could imagine a high-field experiment that computes the wave velocities.

In the direction normal to B and D, the velocities will separate into positive and negative ones, each one approximately equal to ± 1. Let $n_\pm$ be their respective numbers ; one has $n_- + n_+ = 4$. Then $|n_+ - n_-|/2$ equals the rank of D^2m. For instance, the model has to be linear if, and only if, $n_- = n_+ = 2$. It will be as nonlinear as possible when $n_- = 0$ and $n_+ = 4$. This case is illiustrated by the Born–Infeld model for instance.

In this rank-2 non-linear case, the wave propagation ressembles a pure transport in the direction of $D \times B$; at the leading order, there is neither dispersion nor polarization because all the waves are transported with the same velocity.

In conclusion, the qualitative behaviour of the high-field wave propagation would be a way to detect non-linearity in electrodynamics.

Errata

We point out that Proposition 1, page 15 of [16] is incorrect. The system (1) is not always hyperbolic under the assumptions that h_{DD} is positive definite and h is 'Lorentz invariant'.

References

[1] I. Białynicki-Birula. Nonlinear electrodynamics: variations on a theme by Born and Infeld. *Quantum theory of particles and fields*, World Sci. Publishing, Singapore, (1983), pp 31–48.

[2] G. Boillat. Nonlinear electrodynamics: Lagrangians and equations of motion. *J. Math. Phys.*, **11** (1970), pp 941–951.

[3] G. Boillat. Chocs caractéristiques. *C. R. Acad. Sci. Paris, Série I*, **274** (1972), pp 1018–21.

[4] G. Boillat. Convexité et hyperbolicité en 'electrodynamique non linéaire. *C. R. Acad. Sci. Paris, Série I*, **290** (1980), pp 259–261.

[5] G. Boillat, A. Strumia. Limitation des vitesses d'ondes et de choc quand la densité relativiste d'entropie (ou d'énergie) est convexe. *C. R. Acad. Sci. Paris, Série I*, **274** (1988), pp 1018–21.

[6] M. Born. Théorie non-linéaire du champ électromagnétique. *Annales de l'Inst. Henri Poincaré*, **7** (1937), pp 155–265.

[7] M. Born, L. Infeld. Foundations of the new field theory. *Proc. Roy. Soc. London A*, **144** (1934), pp 425–451.

[8] Y. Brenier. Hydrodynamic structure of the augmented Born–Infeld equations. *Arch. Rat. mech. Anal*, **172** (2004), pp 65–91.

[9] D. Christodoulou. *The action principle and partial differential equations.* Annals of Math. Studies, **116**. Princeton Univ. Press (2000).

[10] B. D. Coleman, E. H. Dill. Thermodynamic restrictions on the constitutive equations of electromagnetic theory. *Z. Angew. Math. Phys.*, **22** (1971), pp 691–702.

[11] G. W. Gibbons, D. A. Rasheed. Electric-magnetic duality rotations in non-linear electrodynamics. *Nucl. Phys.*, **B454** (1995), pp 185–206.

[12] T. Kato. *Perturbation theory for linear operators.* Grundlehren der mathematischen Wissenschaften **132**. Springer-Verlag, Heidelberg (1976).

[13] G. Mie. Grundlagen einer Theorie der Materie. *Annalen der Physik*, **37** (1912), pp 511–ff, **39** (1912), pp 1–ff, **40** (1913), pp 1–ff.

[14] J. Plebanski. *Lecture on nonlinear electrodynamics.* Nordita (1970).

[15] D. Serre. `http://www.umpa.ens-lyon.fr/ serre/DPF/exobis.pdf` : Additional list of exercises to *Matrices; theory and applications*. Grad. Text in Math. **216**, Springer-Verlag (2002).

[16] D. Serre. Hyperbolicity of the non-linear models of Maxwells equations. *Arch. Rat. Mech. Anal*, **172** (2004), pp 309–331