Hyperbolicity of the non-linear models of Maxwell's equations

Denis Serre*
ENS Lyon, UMPA (UMR 5669 CNRS),
46, allée d'Italie, F-69364 Lyon Cedex 07.

July 31, 2003

Abstract

We consider the class of nonlinear models of electromagnetism that has been described by Coleman & Dill [7]. A model is completely determined by its energy density $W(B, D)$. Viewing the electromagnetic field (B, D) as a 3×2-matrix, we show that polyconvexity of W implies the local well-posedness of the Cauchy problem within smooth functions of class H^s with $s > 1 + d/2$.

The method follows that designed by C. Dafermos in his book [9] in the context of nonlinear elasticity. We use the fact that $B \times D$ is a (vectorial, non-convex) entropy, and we enlarge the system from 6 to 9 equations. The resulting system admits an entropy (actually the energy) that is convex.

Since the energy conservation law does not derive from the system of conservation laws itself (Faraday's and Ampère's laws), but also need the compatibility relations $\mathrm{div} B = \mathrm{div} D = 0$ (the latter may be relaxed in order to take in account electric charges), the energy density is not an entropy in the classical sense. Thus the system cannot be symmetrized, strictly speaking. However, we show that the structure is close enough to symmetrizability, so that the standard estimates still hold true.

Enlarged systems of conservation laws

Consider a system of conservation laws

$$\partial_t u + \sum_{\alpha=1}^{d} \partial_\alpha f^\alpha(u) = 0, \quad t > 0,\, x \in I\!R^d, \tag{1}$$

*This research was funded in part by the EC as contract HPRN-CT-2002-00282. It was initiated when the author was visiting the Isaac Newton Institute for Mathematical Sciences (Cambridge, UK) within the program "Nonlinear hyperbolic waves in phase dynamics and astrophysics". Part of this research was done when the author was kindly hosted by the Norvegian Academy of Sciences and Letters.

where $u(x,t) \in I\!R^n$ is the unknown and f^α are given smooth fluxes. Recall that the existence of an additional conservation law

$$\partial_t \eta(u) + \sum_\alpha \partial_\alpha q^\alpha(u) = 0, \tag{2}$$

compatible with (1), and where η is a (scalar) function, strictly convex in the sense that $D^2\eta > 0$, ensures that (1) is hyperbolic in the direction of the time t. See [9, 11, 15] for instance. Actually, it allows to symmetrize (1) in the form

$$A^0(z)\partial_t z + \sum_\alpha A^\alpha(z)\partial_\alpha z = 0, \tag{3}$$

where $z := d\eta(u)$ is the "dual variable". Symmetrization means that the matrices $A^0(z)$, $A^\alpha(z)$ are symmetric, the first one being positive definite. As is well-known, the symmetric form (3) has nice consequences for the Cauchy problem (see [9] for instance):

i. Given an initial data of class $H^s(I\!R^d)$ (actually, uniformly locally in H^s is sufficient) with $s > 1 + d/2$ (which ensures that $H^s \subset \mathcal{C}^1$), there exists a positive time T and a unique classical solution in the strip $(0,T) \times I\!R^d$.

ii. The uniqueness holds in the following stronger sense: The classical solution, when it exists, coincide with every weak entropy solutionsThe latter are essentially bounded fields which satisfy (1) in the distributional sense, together with the "entropy inequality"

$$\partial_t \eta(u) + \sum_\alpha \partial_\alpha q^\alpha(u) \leq 0. \tag{4}$$

Obviously, the convexity of the entropy (which turns out to be the mechanical energy in isothermal models) is only a sufficient condition for hyperbolicity, but not a necessary one. It has been well-known for a long time that the mechanical energy of a hyperelastic material cannot be convex (see [6].) This observation led C. Dafermos [9] to the following procedure (see also [10].)

Assume that the system (1) is compatible with some special conservation laws, where the conserved quantities will be denoted by the vector P:

$$\partial_t P(u) + \sum_\alpha \partial_\alpha \pi^\alpha(u) = 0. \tag{5}$$

Although the components of P play a role very similar to that of η, we are reluctant to call them entropies. It turns out that the conservation laws (5) do not always depend on the equation of state of the underlying medium. On the contrary, it is common in thermodynamics that the knowledge of the entropy η determines completely the equation of state.

Assume now that the, *a priori* non convex, entropy can be rewritten as a strictly convex function of $(u_1, \ldots, u_n, P_1, \ldots, P_r)$:

$$\eta(u) = \phi(u_1, \ldots, u_n, P_1, \ldots, P_r), \quad D^2\phi > 0.$$

Then we are tempted to increase the number of unknowns as well as of equations by writing a system

$$\partial_t u + \sum_\alpha \partial_\alpha f^\alpha(u) = 0, \tag{6}$$

$$\partial_t P + \sum_\alpha \partial_\alpha \pi^\alpha(u) = 0, \tag{7}$$

expecting that ϕ is an entropy of the resulting system. If it were the case, then the enlarged system would be symmetrizable and the existence and uniqueness properties mentionned above would apply. Furthermore, the new system contains (1) in the sense that if an initial data (u^0, P^0) satisfies $P^0 \equiv P(u^0)$, then the classical solution will satisfy identically $P \equiv P(u)$, and therefore u will be a classical solution of (1) with data u^0. Hence the existence and uniqueness properties hold also for (1).

The situation is not so simple however, because ϕ is not in general an entropy of the enlarged system ! It is clear that there is not a unique way (if any) to write η as a convex function of (u, P), and one needs to find one such function that is an entropy. Also, the fluxes f^α and π^α may need to be rewritten as functions of (u, P) instead of u only, and they do in general. For this reason, there does not exist yet a satisfactory theory of enlargement of systems of conservation laws. We content ourselves to treat each system of physical interest on a case-by-case basis. This is what Dafermos did for the system governing the motion of a hyperelastic material. The purpose of the present note it to give a convenient treatment of the non-linear models of electromagnetism. Our work has been partly influenced by that of Y. Brenier [5], who treated the special case of the Born–Infeld model. We emphasize that, thanks to the very special structure of the Born–Infeld model, Brenier could extend it to a rather simple system of ten equations/unknowns, with pretty accurate information such as the knowledge of the wave velocities, while in the general case, our extension consists in nine equations with a pretty involved nonlinear structure. In particular, our work does not contain that of Brenier.

Before entering into details, we mention the last difficulty that we have to overcome. Recall that the Maxwell's equations contain, besides evolutionary conservation laws (Faraday's and Ampère's laws), constraints like $\mathrm{div}B = 0$ and $\mathrm{div}D = 0$ (say, in the absence of charges.) Although the conservation of energy relies only upon the former, and thus the energy is an entropy in the classical sense, this is not any more true at the level of the enlarged system. There, the conservation of energy explicitly involves the constraint. That means that the enlarged system may not be symmetrized in the usual way, a fact that could spoil the high order estimates. We show however that the structure of such constrained systems with a convex entropy allows for convenient estimates. Therefore the well-known existence and stability theorems still apply.

The Coleman–Dill model for electromagnetism

We place ourselves in the context of an electromagnetic field (E, B) obeying to a non-linear system of conservation laws. The ambiant space is $I\!R^3$ ($d = 3$.) The law of Faraday

$\partial_t B + \mathrm{curl} E = 0$ must be completed with Ampère's law $\partial_t D - \mathrm{curl} H = 0$ (assuming for simplicity that there are neither charges nor currents.) In standard Maxwell's equations, the equations of state are linear: $D = \epsilon E$ and $H = \mu B$, where ϵ and μ are constant symmetric tensors (often scalars.)

There are several reasons for dropping the standard, linear Maxwell's equations. On the one hand, there exist media in which the equations of state become non-linear. On the other hand, the electric field in the vacuum grows like r^{-2} near a punctual charge, and this growth is responsible for a infinite total energy ! Several corrections of the electromagnetic theory have been made to resolve this contradiction, and one of them was to postulate a non-linear energy density which forces the electromagnetic field to remain finite, though behaving like r^{-2} in the far field, a harmless fact. The most famous model in this respect is certainly that designed by M. Born & L. Infeld [4].

A general theory of non-linear electromagnetic models is due to B. D. Coleman & E. H. Dill. It assumes that the model be compatible with some energy conservation, where the stored energy has the form of a smooth function $W(B, D)$. Then, taking the conserved quantities B and D as primary variables, the equations of state read

$$E_i := \frac{\partial W}{\partial D_i}, \quad H_i := \frac{\partial W}{\partial B_i}.$$

Whence the system

$$\partial_t B + \mathrm{curl}\, \frac{\partial W}{\partial D} = 0, \quad \partial_t D - \mathrm{curl}\, \frac{\partial W}{\partial B} = 0 \quad \mathrm{div} B = \mathrm{div} D = 0. \tag{8}$$

The last two equation are constraints that are compatible with Faraday's and Ampère's laws, in the sense that they remain true forever if they were at initial time. This system is compatible with the energy conservation law

$$\partial_t W + \mathrm{div}(E \times H) = 0. \tag{9}$$

The energy flux $E \times H$ is often called the "Poynting vector". Notice that (9) relies only upon Faraday's and Ampère's laws, it does not need that the constraints be satisfied. When W is a convex function of $u := (B, D)$, we conclude as usual that the system (8) is symmetrizable hyperbolic, and therefore that local existence and uniqueness properties hold.

However, it is not always the case that W is convex. For instance, in the Born–Infeld model, one has

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

which fails to be convex far away from the origin, though (8) remains hyperbolic as is it well-known. When W is not convex, the theorems of Chapter 5 in [9] do not apply. Even Theorems 5.3.1 and 5.3.2, which deal with systems with linear differential constraints (called "involutions"), do not apply, since they require that the entropy (here the energy) of the system be strictly convex in the directions of the "involution cone" $\mathcal{C}$. In our context, the involutions are the constraints $\mathrm{div} B = 0$ and $\mathrm{divD} = 0$, hence the cone is

$I\!R^6$, meaning that the convexity along the cone is the usual convexity. Hence the local existence and the stability of classical solutions remain open questions when W fails to be convex.

Remark: In the relativistic formalism, the Ampère's law is viewed as the Euler–Lagrange equation of a Lagrangian $\mathcal{L}(B,E) = L(\|E\|^2 - \|B\|^2, E\cdot B)$, constrained by Faraday's law. This gives the relations

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

More precisely,

$$W(B,D) = \sup_{e\in I\!R^3}\{D\cdot e - \mathcal{L}(B,e)\}. \tag{10}$$

The fact that $\mathcal{L}$ depends only on two scalar quantities $\|E\|^2 - \|B\|^2$ and $E\cdot B$ is due to the invariance under Lorentz transformations: The electro-magnetic field must be viewed as the 2-differential form $\Omega_{EM} := dt\times(E\cdot dx) + dx\times(B\times dx)$. The invariants of differential forms of degree two under the action of the Lorentz group $\mathbf{O}(1,3)$ turn out to be the above quantities. In this formulation, it is desirable to express the convexity of W in terms of a property of L.

The extra dependent variable

The key observation is that, for physically relevant solutions, that is those satisfying the natural constraints divB $=$ div$D = 0$, the vector $P := B\times D$ obeys to some conservation law

$$\partial_t P_i + \mathrm{div}(E_i D + H_i B) + \partial_i(W - E\cdot D - H\cdot B) = 0. \tag{11}$$

It is amazing that equation (11) is not any more in conservation form when B or D fails to be solenoidal. This fact ressembles much the case of a hyperelastic material, where extra conservation laws hold only when the tensor part of the unknown is a deformation tensor. In both situations, the constraints have the form $Lu = 0$ where L is a linear differential operator in the space variable, which are compatible with the evolution in the sense that, if they are satisfied at initial time, then they persist when time increases.

We point out an important difference however, in that (11) does involve W itself. Therefore, its Rankine–Hugoniot conditions are usually not compatible with those of (8) once shock waves develop. An exception to this flaw is the Born–Infeld model, at least when shocks have moderate amplitude, since its characteristic fields are linearly degenerate.

The advantage of enlarging the system becomes clear in the Born–Infeld case. Then the energy density W becomes a convex function of (B,D,P), when written in the form

$$\sqrt{1+\|B\|^2+\|D\|^2+\|P\|^2}.$$

There remains however to find a new way to write the fluxes in (8, 11), in such a way that the above function be an entropy of the enlarged system. More precisely, what we need is the following. Given a convex function $\phi(B, D, P)$ that coincides with W on the "equilibrium" submanifold

$$\Sigma := \left\{(B, D, B \times D)\,;\; B, D \in \mathbb{R}^3\right\},$$

find a system

$$\partial_t B + \operatorname{curl} E = 0, \quad \operatorname{div} B = 0, \tag{12}$$

$$\partial_t D - \operatorname{curl} H = 0, \quad \operatorname{div} D = 0, \tag{13}$$

$$\partial_t P + \operatorname{Div} T = 0, \tag{14}$$

where

i. $E = E(B, D, P)$, $H = H(B, D, P)$ and $T = T(B, D, P)$ coincide, on Σ, with $\partial W/\partial D$, $\partial W/\partial B$ and $E \otimes D + H \otimes B + (W - E \cdot D - H \cdot B)I_3$ respectively,

ii. ϕ is an entropy of the resulting system.

Of course, the second point is the difficult one. Once this program is achieved, we may apply the local existence and uniqueness properties to (12, 13, 14) and, whenever $P \equiv B \times D$ holds at initial time, this remains true for every time. In the latter situation, (B, D) is a classical solution to (8).

The enlarged system

As mentionned above, there is not yet a systematic method for solving the above program. Thus we give the system that fits the above requirements, without convincing explanations. To begin with, the chain rule suggests natural equations of state for E and H:

$$E = \frac{\partial \phi}{\partial D} - B \times \frac{\partial \phi}{\partial P}, \quad H = \frac{\partial \phi}{\partial B} + D \times \frac{\partial \phi}{\partial P}. \tag{15}$$

There remains to choose $T(B, D, P)$ in an appropriate way and this is the less clear point. The following choice works:

$$\begin{aligned} T(B, D, P) \; := \; & \frac{\partial \phi}{\partial B} \otimes B + \frac{\partial \phi}{\partial D} \otimes D - P \otimes \frac{\partial \phi}{\partial P} \\ & + \left(\phi - B \cdot \frac{\partial \phi}{\partial B} - D \cdot \frac{\partial \phi}{\partial D} - P \cdot \frac{\partial \phi}{\partial P}\right) I_3. \end{aligned} \tag{16}$$

The fact that T coincides with

$$T_0 := E \otimes D + H \otimes B + (W - E \cdot D - H \cdot B)I_3$$

on the equilibrium manifold $P = B \times D$ is tricky. It involves the following crucial identity for vectors $X, Y, Z \in I\!R^3$:

$$(X \times Y) \otimes Z + (Y \times Z) \otimes X + (Z \times X) \otimes Y = \det(X, Y, Z)\, I_3.$$

Last but not least, one obtains the following identity:

$$\partial_t \phi(B, D, P) + \mathrm{div}(E \times H) = \mathrm{div}\left(\left((P - B \times D) \cdot \frac{\partial \phi}{\partial P}\right) \frac{\partial \phi}{\partial P}\right) - \frac{\partial \phi}{\partial P} \cdot ((\mathrm{div} B) H - (\mathrm{div} D) E). \tag{17}$$

Notice that the right-hand side in (17) vanishes identically when the solution comes from a solution of (8). Then (17) reduces to (9) as expected. In particular, the last two terms vanish for solutions of (12, 13, 14), because of the solenoidal constraints.

To summarize, we have built a system (12, 13, 14) of nine conservation laws in nine unknowns, where the equations of state are (15, 16). We call it the "enlarged system". It is endowed with the entropy ϕ, meaning that it is formally compatible with (17). Of course, because of the presence of non-conservative terms in the right-hand side of (17), we may not apply Theorem 5.1.1 of [9]. However, we easily adapt its proof (see next section) and we obtain:

Theorem 1 *Assume that the function* $U := (B, D, P) \mapsto \phi$ *is strictly convex, that is* $D^2\phi > 0_9$, *and smooth enough. Assume a* $\mathcal{C}^1(I\!R^3)$*-initial data* U^0 *that takes values in some compact subset* $\mathcal{O}$ *of* $I\!R^9$, *and such that* $\nabla U^0 \in H^s$ *for some* $s > 3/2$. *Assume also that* $\mathrm{div} B_0$ *and* $\mathrm{div} D_0$ *vanish identically.*

Then there exists $\tau > 0$ *and a unique* $\mathcal{C}^1$*-solution* U *of the initial-value problem of the enlarged system for* $0 \le t < \tau$. *Furthermore,*

$$\nabla_{x,t} U \in \mathcal{C}^0([0, \tau); H^s(I\!R^3)).$$

Since the equation of state coincide with that of the Maxwell's equation on the equilibrium manifold, we have the following corollary, which we prove by choosing $P^0 := B^0 \times D^0$.

Theorem 2 *Assume that the function* $(B, D, P) \mapsto \phi$ *is strictly convex, that is* $D^2\phi > 0_9$, *and smooth enough. Assume a* $\mathcal{C}^1(I\!R^3)$*-initial data* $V^0 = (B^0, D^0)$ *that takes values in some compact subset* $\mathcal{O}$ *of* $I\!R^6$, *and such that* $\nabla V^0 \in H^s$ *for some* $s > 3/2$. *Assume also that* $\mathrm{div} B_0$ *and* $\mathrm{div} D_0$ *vanish identically.*

Then there exists $\tau > 0$ *and a unique* $\mathcal{C}^1$*-solution* V *of the initial-value problem of the Maxwell's equations (8) for* $0 \le t < \tau$. *Furthermore,*

$$\nabla_{x,t} V \in \mathcal{C}^0([0, \tau); H^s(I\!R^3)).$$

We warn the reader that weak entropy solutions of (8) do not solve (12, 13, 14) in general, because the Rankine–Hugoniot relations of (11) are not compatible with those of (8). This phenomenon is studied in greater details below. For the moment, let us say that it prevents to transfer the weak-strong uniqueness property (Theorem 5.2.1 in [9]) from (12, 13, 14) to (8). Hence the enlargement of Maxwell's system resolves the local existence question and the uniqueness within classical solutions, but not the weak-strong uniqueness. For classical solutions, we have:

Theorem 3 *Assume that the function $(B, D, P) \mapsto \phi$ is strictly convex, that is $D^2\phi > 0_9$, and smooth enough.*

Suppose V and $\bar{V}$ are classical solutions of Maxwell's equations (8) on $[0, \tau)$, taking values in a compact subset $\mathcal{O}$ of $\mathbb{R}^6$, with initial data V^0 and $\bar{V}^0$ that satisfy the constraints. Then

$$\int_{|x|<R} \|V(x,t) - \bar{V}(x,t)\|^2 dx \le a e^{bt} \int_{|x|<R+Mt} \|V^0(x) - \bar{V}^0(x)\|^2 dx \tag{18}$$

holds for any $R > 0$ and $t \in [0, \tau)$, with positive constants a, b and M that depend only on $\mathcal{O}$, except for b, which depends also on the Lipschitz constants of the solutions.

Evolution of $P - B \times D$. One checks easily that $\delta := P - B \times D$ satifies the evolution equation

$$\partial_t \delta = \left(\frac{\partial \phi}{\partial P} \cdot \nabla \right) \delta + \left(\operatorname{div} \frac{\partial \phi}{\partial P} \right) \delta + \left(\nabla \frac{\partial \phi}{\partial P} \right) \delta, \tag{19}$$

where in the last term $(\nabla X)\delta$ stands for the vector of components $(\partial_\alpha X) \cdot \delta$. Equation (19) confirms that the enlarged system is compatible with the nonlinear Maxwell's system, in the following sense:

i. Given a classical solution (B, D) of (8), then $(B, D, B \times D)$ is a classical solution of the enlarged system,

ii. Given a classical solution (B, D, P) of the enlarged system that satisfies $P = B \times D$ at initial time (Maxwell-type initial data), then $P \equiv B \times D$ remains true for positive time and (B, D) is a solution of (8).

Stability an existence for constrained systems of conservation laws

Our system (12, 13, 14) belongs to the general class of systems of conservation laws (see (1)), constrained by linear, constant coefficient differential operators in the space variables, say

$$\sum_{\alpha=1}^{d} M^\alpha \partial_\alpha u = 0. \tag{20}$$

The matrices M^α are $m \times n$. More precisely, we focus on those systems whose fluxes satisfy the algebraic relations (see [9], Section 5.3)

$$M^\alpha f^\beta + M^\beta f^\alpha \equiv 0, \quad 1 \le \alpha, \beta \le d, \tag{21}$$

ensuring the compatibility of the constraints. For instance, in Maxwell's equations, we have $d = 3$, $n = 6$, $m = 2$.

We assume that a given system (1,20) is compatible with an entropy balance law, in the sense that there exists a smooth function $\eta(u)$, strictly convex in the sense that $D^2\eta > 0_n$, a smooth flux $\vec{q}(u)$ and a field (a differential form of order one) $Z(u)$, such that every smooth solution of the unconstrained conservation laws (1) satisfies

$$\partial_t \eta(u) + \mathrm{div}\vec{q}(u) = Z(u) \cdot \sum_{\alpha=1}^{d} M^\alpha \partial_\alpha u. \tag{22}$$

Algebraically, this means that $\eta, f, \vec{q}$ and Z satisfy the relations

$$\frac{\partial q^\alpha}{\partial u_i} = \frac{\partial \eta}{\partial u_k} \frac{\partial f_k^\alpha}{\partial u_i} + \sum_{p=1}^{m} Z_p M_{pi}^\alpha. \tag{23}$$

Of course, a smooth solution of (1,20) satisfies the more usual identity (2).

Due to the presence of the sum in the right-hand side of (23), one cannot in general symmetrize the system in the form

$$\partial_t \frac{\partial L^0}{\partial z_k} + \sum_\alpha \partial_t \frac{\partial L^\alpha}{\partial z_k} = 0,$$

with $z := \nabla_u \eta$ the "dual variable". Although the Russian school had a lot of successes in designed more elaborated symmetrization (see for instance the review text by S. Godunov [11]), we wish to present estimates that are valid in our abstract formalism.

To begin with, we consider two solutions u and v of the full system, the former being smooth. The case where v is smooth too is used in the proof of the contraction property in the existence procedure. When v is only bounded, as it occurs in weak-strong uniqueness results, we also assume that it satisfies the entropy inequality (4). Defining

$$\Delta(x, t) := \eta(v) - \eta(u) - d\eta(u) \cdot (v - u)$$

and

$$Q^\alpha(x, t) := q^\alpha(v) - q^\alpha(u) - d\eta(u) \cdot (f^\alpha(v) - f^\alpha(u)),$$

we have the well-known inequality

$$\partial_t \Delta + \mathrm{div} Q \leq J, \tag{24}$$

where, using Einstein's convention

$$J := -D^2\eta(u)(u_t, v - u) - D^2\eta(u)(\partial_\alpha u, f^\alpha(v) - f^\alpha(u)).$$

Due to (1), the first term in J equals $D^2\eta(u)(df^\alpha(u)\partial_\alpha u, v - u)$. In the unconstrained theory, $df^\alpha(u)$ would be $D^2\eta(u)$-symmetric ; thus this term could be rewritten in the form $D^2\eta(u)(\partial_\alpha u, df^\alpha(u)(v - u))$, and we should end up with

$$J = -D^2\eta(u)(\partial_\alpha u, f^\alpha(v) - f^\alpha(u) - df^\alpha(u)(v - u)) = \mathcal{O}(\|v - u\|^2) = \mathcal{O}(\Delta).$$

Whence the inequality

$$\frac{d}{dt}\int_{\|x\|<R-Mt}\Delta(x,t)\,dt \le C\int_{\|x\|<R-Mt}\Delta(x,t)\,dt, \tag{25}$$

for some convenient $M>0$. With the help of Gronwall's inequality, (25) yields an $L^\infty_t L^2_x$-estimate that includes a finite velocity propagation result.

As far as constrained systems are concerned, $df^\alpha(u)$ is not any more $D^2\eta(u)$-symmetric. However, it is not that much far. Since $D^2q^\alpha(u)$ and $d\eta(u)D^2f^\alpha(u)$ are symmetric, the defect of symmetry in $D^2\eta(u)(df^\alpha(u)\cdot,\cdot)$ is the matrix (recall that M has constant coefficients)

$$\left(\frac{\partial Z_p}{\partial u_j}M^\alpha_{pi}-\frac{\partial Z_p}{\partial u_i}M^\alpha_{pj}\right)_{1\le i,j\le n}.$$

When applied to $\partial_\alpha u\otimes(v-u)$, this matrix results into (use the constraint satisfied by u)

$$(\partial_\alpha Z)\cdot M^\alpha(u-v).$$

Since both v and u satisfy the constraint, we obtain

$$J=\mathcal{O}(\Delta)+\partial_\alpha\left(Z\cdot M^\alpha(u-v)\right).$$

We warn the reader that an integration does not immediately yield the same inequality as (25), for the last term in J, though in conservative form, is not quadratic in $v-u$. Therefore, the boundary term on $\partial B_{\|x\|<R-Mt}$ cannot be absorbed by that of the quadratic flux Q. Similarly, the obvious bound by

$$\|\partial_\alpha Z\|_\infty\|v-u\|_{L^2(B_{R-Mt})}$$

cannot be handled by Gronwall's inequality.

We overcome this difficulty as follows. First, we apply (24) to the case where v is a constant state. Using again the constraint, we obtain

$$\partial_t\Delta+\mathrm{div}Q\le\mathcal{O}(\Delta)+\partial_\alpha\left((Z(u)-Z(v))\cdot M^\alpha(u-v)\right). \tag{26}$$

Integrating over a domain B_{R-Mt} with a large enough[1] $M>0$, we obtain (25) ; then we may conclude with the help of Gronwall's estimate. This proves that a smooth solution remains compactly supported if its initial data was so. Going back to the case where v is a weak entropy solution, we assume that u_0 was compactly supported, up to a constant $\bar{u}$. As we just proved it, $u(t)\equiv\bar{u}$ for $\|x\|>R+Mt$ for some R (depending on u_0) and $M>0$. Integrating (26), the conservative term in J drops out[2], and we obtain (25), where the integrals carry over the whole domain $\mathbb{R}^d$. Gronwall's inequality gives the conclusion.

We now sketch the procedure that yields the existence result. As usual, we smooth out the initial datum, through convolution. Since a convolution preserves the constraint, we

[1] The constant M depends on the sup-norm of the matrices M^α and of $\nabla_u Z$.

[2] Here, we replace $Z(u)$ by $Z(u)-Z(\bar{u})$, thanks to the constraint.

are free to assume that the initial data is of class $\mathcal{C}^\infty$. We define inductively approximate solutions, by solving linear Cauchy problems with variable coefficients: If the k-th iterate u is known, then the $(k+1)$-th, say v, is defined by

$$\partial_t v + \partial_\alpha(f^\alpha(u) + df^\alpha(u)(v-u)) = 0, \quad v(0) = u_0. \tag{27}$$

The advantage of this procedure is that (27) is compatible with the constraints, since differentiating (21) yields

$$M^\alpha df^\beta(u)w + M^\beta df^\alpha(u)w = 0, \quad u, w \in I\!R^n.$$

Then, thanks to the existence of the entropy, the linear system, together with the constraints, are governed by a hyperbolic operator, ensuring well-posedness in Sobolev spaces H^s.

There remains to prove that the iteration is stable and convergent on some time interval $(0,T)$ with $T>0$. For the sake of simplicity, we only establish *a priori* estimates for the smooth solutions u of the non-linear Cauchy problem. Given a multi-index r of length $l \geq 1$, we denote by v_r the space derivative $\partial^r u$. It satisfies an identity

$$\partial_t v_r + df^\alpha(u)\partial_\alpha v_r = P_r.$$

Hereabove, P_r is a universal polynomial in the quantities $D^k f(u)$ (up to order $l+1$) and v_s (up to order l). We wish to estimate an $L^\infty_t L^2_x$-norm of v_r. To do so, we compute the time derivative of $D^2\eta(u)(v_r, v_r)$. Using the formula above, we have

$$\partial_t D^2\eta(u)(v_r,v_r) = D^3\eta(u)(-\partial_\alpha f^\alpha(u), v_r, v_r) + 2D^2\eta(u)(P_r - df^\alpha(u)\partial_\alpha v_r, v_r).$$

In the sequel, we denote by Q universal polynomials as above, depending also on some derivatives of η. They could differ from one line to the other, but their order with respect to u remains equal to l. For instance, we have

$$\partial_t D^2\eta(u)(v_r,v_r) = Q - 2D^2\eta(u)(df^\alpha(u)\partial_\alpha v_r, v_r).$$

Differentiating (23), we may replace $D^2\eta df^\alpha$ by

$$D^2q^\alpha - d\eta D^2 f^\alpha - (dZ)M^\alpha,$$

where the first two terms are symmetric matrices. Hence we have

$$\partial_t D^2\eta(u)(v_r,v_r) + \partial_\alpha\left(D^2q^\alpha(u)(v_r,v_r) + d\eta D^2 f^\alpha(u)(v_r,v_r)\right) = Q - 2((v_r\cdot\nabla_u)Z)M^\alpha\partial_\alpha v_r.$$

Since v_r satisfies the same constraint as u, we end up with

$$\partial_t D^2\eta(u)(v_r,v_r) + \partial_\alpha\left(D^2q^\alpha(u)(v_r,v_r) + d\eta D^2 f^\alpha(u)(v_r,v_r)\right) = Q.$$

The rest of the computation is the same as in the unconstrained case. We integrate over $I\!R^d$ and sum over the multi-indices of length less than or equal to ℓ, where ℓ is larger than $1 + d/2$. The right-hand side may be estimated thanks to Moser's type estimates, and one concludes with the help of Gronwall's inequality.

We summarize our results in the following two statements.

Theorem 4 *Assume that the fluxes f^α satisfy the identities (21), so that (1) is compatible with the constraints (20). Assume also that the unconstrained system (1) is compatible with a balance law (22) where $D^2\eta > 0_n$.*

Let u_0 be bounded, constant outside of a ball B_R, with $\nabla_x u_0 \in H^s(\mathbb{R}^d)$ $(s > 1 + d/2)$. Assume that it satisfies the constraint (20).

Then there exists a $T > 0$ and a unique classical solution of (1,20), such that

$$\nabla_{x,t} u \in \bigcap_{k=0}^{[s]} \mathcal{C}^k([0,T); H^{s-k}(\mathbb{R}^d)),$$

satisfying $u(0) = u_0$. The solution is constant outside of the ball B_{R+Mt}, where M depends only on the system itself and the L^∞-norm of u.

Theorem 5 *Assume the same structural hypotheses as in Theorem 4. Let u be a classical solution as above (in particular, u is constant outside a compact set), and let v be a bounded weak entropy solution ; both are defined on a compact interval $[0,T]$. Then there exists constants c_0, c_1 that do not depend on v, such that*

$$\int_{\mathbb{R}^d} \|v(x,t) - u(x,t)\|^2 dx \le c_0 e^{c_1 t} \int_{\mathbb{R}^d} \|v(x,0) - u_0(x)\|^2 dx.$$

It is an open question to prove the local stability, that is an inequality of the form (18).

Polyconvexity and hyperbolicity

Let us consider $M := (B, D)$ as a 3×2 matrix, instead of a 6-vector. Then $B \times D$ is the set of non-trivial minors of M. Hence the convexity of the map $(B, D, P) \mapsto \phi$ is nothing but the *polyconvexity* of the energy density W, as defined by Ball [1]. Hence Theorems 2 and 3 state that, under strict polyconvexity, the Cauchy problem is locally well-posed within H^s for $s > 1 + d/2$. In particular, strict polyconvexity implies hyperbolicity of Maxwell's equations (8). It is thus natural to ask whether hyperbolicity implies polyconvexity, at least in its weak form (that is, $D^2\phi \ge 0_9$).

The polyconvex functions have been the object of numerous papers because of their role in elasticity theory. There does not seem to be any kind of explicit description of polyconvexity, besides its crude definition. An obviously weaker property is the *rank-one convexity*, meaning that

$$s \mapsto W(U + sV)$$

is convex whenever V has rank one. Notice that, writing $V = (B', D')$, this means the equality $B' \times D' = 0$; in particular, the map $s \mapsto (B + sB') \times (D + sD')$ is affine.

It turns out that, in the context of matrices of size $m \times q$ with $\min(m, q) = 2$, the quadratic polyconvex functions are precisely those that are rank-one convex. In other words, they are characterized (in our context) by $W(B, D) \ge 0$ whenever $B \times D = 0$.

This is a consequence of old results by Terpstra [17] or by the author [13]. Since the proofs in these papers are far from direct, we feel free to give here a more explicit one. We actually need the, even weaker, following statement.

Lemma 1 *Let Q be a quadratic form on $\mathbf{M}_2(I\!R)$ that takes (strcitly) positive values on rank-1 matrices. Then there exists some real constant μ, such that $Q + \mu \det$ is positive definite.*

Proof.

Let $\alpha > 0$ denote the minimum value of Q on the set Γ of rank-1 matrices that have unit norm, say for the norm $\sqrt{\mathrm{Tr}(M^T M)}$. Denote by S the symmetric matrix of the form Q and J that of the form det. We study the function

$$\mu \mapsto m(\mu) := \min_{\|M\|=1} (Q(M) + \mu \det M).$$

This is a continuous function, and we want to prove that it takes at least one positive value. Hence assume the contrary $m(\mu) \leq 0$ for every $\mu \in I\!R$.

Let M_μ be a minimum of $Q + \mu \det$ over Γ. By compactness, such points exist, and they admit at least one cluster point M_+ as $\mu \to +\infty$. From $Q(M_\mu) + \mu \det M_\mu \leq 0$, we find that $\det M_+ \leq 0$. If $\det M_+ = 0$, then $Q(M_+) \geq \alpha$. Therefore

$$\frac{1}{\mu} Q(M_\mu) + \det M_\mu \leq 0,$$

where the first term is positive for μ large enough, implies that $\det M_\mu < 0$. Finally, we have proven that there exists large positive μ's such that $\det M_\mu < 0$. Making $\mu \to -\infty$ instead, we find that there exist large enough negative μ's such that $\det M_\mu > 0$.

We now remark that the set

$$\bigcup_{\mu \in I\!R} \{\mu\} \times \mathrm{argmin}\{Q(M) + \mu \det M \,;\, \|M\| = 1\}$$

is a connected set. This is a rather unsual property, which is a consequence of continuity, compactness, and the crucial fact that each section $\mathrm{argmin}\{Q(M) + \mu \det M \,;\, \|M\| = 1\}$ is itself connected. As a matter of fact, it is the intersection of the unit sphere with the eigenspace of $S + \mu J$, a symmetric matrix, associated to its least eigenvalue.

From above, we deduce that there exists a number μ and a minimal matrix M_μ, such that $\det M_\mu = 0$. But then $m(\mu) = Q(M_\mu) \geq \alpha$, since $M_\mu \in \Gamma$.

QED

We warn the reader that the proof, proceeding *ad absurdum*, does not tell that the maximum of $m(\mu)$ as μ runs over $I\!R$, equals α. We only know that $0 < \max m(\mu) \leq \alpha$, the second inequality being the classical one, $\max\min \leq \min\max$.

It is now tempting to compare the hyperbolicity of Maxwell's system with rank-one convexity. Let us write its evolution part under the form

$$\partial_t U + \sum_{\alpha=1}^{3} A^\alpha(U) \partial_\alpha U = 0.$$

Given a unit vector ξ, define

$$A(\xi;U) := \sum_{\alpha=1}^{3} \xi_\alpha A^\alpha(U).$$

Theorem 6 *Assume that W is stricly rank-one convex, in the sense that*

$$\frac{d^2}{ds^2} W(U + sV) > 0$$

whenever V has rank one. Then the Maxwell's system (8) is hyperbolic.

We remark that we do not need to prove that $A(\xi;U)$ has a real spectrum, a fact that could even fail. Since (8) contains the constraints $\mathrm{div} B = \mathrm{div D} = 0$ and the latter are compatible with the system, $A(\xi;U)$ admits a four-dimensional invariant subspace $N(\xi) = \xi^\perp \times \xi^\perp$. What we have to prove is that the restriction of $A(\xi;U)$ to $N(\xi)$ is diagonalisable.

Proof.

Because of rotational invariance, we content ourselves with $\xi = \vec{e}^1$. Since the constraints reduce to $\partial_1 B_1 = \partial_1 D_1 = 0$ and the evolution part gives $\partial_t B_1 = \partial_t D_1 = 0$, we may restrict to the system describing the evolution of $u = (B_2, B_3, D_2, D_3)^T$:

$$\partial_t u + A(U)\partial_1 u = 0.$$

It is therefore enough to prove that $A(U)$ is diagonalizable.

One easily checks that $A(U) = -JS$, where (see above)

$$J := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

and S is the Hessian matrix of W, with respect to (B_2, B_3, D_2, D_3). The assumption tells precisely that S defines a quadratic form that is positive on rank-one matrices. From Lemma 1, there exists a constant μ such that $S + \mu J$ is positive definite. Therefore,

$$A(U) - \mu I_4 = -J(S + \mu J)$$

is diagonalizable with real eigenvalues, as the product of two symmetric matrices, one of them being positive definite, and so is $A(U)$.

QED

Though being a necessary property for local well-posedness, it is not clear whether the hyperbolicity is sufficient. According to Métivier [12] (see also Taylor [16]), the local well-posedness has been proved in the case where there exists a symbolic symmetrizer $A^0(\xi;U)$, that is a symmetric positive definite matrix depending smoothly on its variables, such that $A^0(\xi;U)A(\xi;U)$ be symmetric. This contains the following situations:

- Friedrichs symmetrizability. This is the case that we exploit here.
- Hyperbolic systems whose wave velocities are simple in every directions[3]. Such systems are called *constantly hyperbolic* systems by Benzoni & all. [2].

It is not clear that our nonlinear Maxwell models have wave velocities of constant multiplicities. The multiplicities in the standard linear model, or in Born–Infeld, equal two (see [5]), but the speeds split for general nonlinear models, and the multiplicity two could persist only on a submanifold. Clearly, a more detailed analysis of the wave speeds is needed. For the moment, well-posedness, say *a priori* estimates, remain open under the assumption of strict rank-one convexity.

We also point out that rank-one convexity is not required for having hyperbolicity. For instance, the diagonal matrix $S := \mathrm{diag}\{-1,1,1,-1\}$ is not rank-one convex, although the product

$$A = -JS = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

is a diagonalisable matrix with a real spectrum ± 1. There seems, however that the set of symmetric matrices S such that $-JS$ is diagonalizable with real spectrum be disconnected and that the rank-one convex is one of the connected components, up to some boundary points. For instance, we have the following statement.

Proposition 1 *Let S be a symmetric matrix, defining a quadratic form Q that satisfies*

$$\min_{M\in\Gamma} Q(M) = 0.$$

Then there exists a real number θ such that $S-\theta J$ is positive semi-definite, but not *positive definite. In particular, $-\theta$ is an eigenvalue of A.*

If moreover $\dim\ker(S-\theta J) = 1$, *then $-\theta$ is not semi-simple.*

This proposition shows that in generic examples where W has a transition from strictly rank-one convexity to rank-one non-convexity (akward terminology), system (8) fails to be hyperbolic.

Proof.

Let M be an element of Γ where Q vanishes. In particular, the minimum of Q over $I\!R\Gamma$ is achieved at M. Extremality tells that there exists a numnber θ such that $SM = \theta JM$ (JM is the normal to $I\!R\Gamma$ at M.) Hence $-\theta$ is an eigenvalue of A.

Up to the shift $S \mapsto S - \theta J$, we may assume $\theta = 0$. Also, up to an equivalence within $\mathbf{M}_2(I\!R)$, we may assume that

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

[3]Without extra complexity, we may weaken the assumption by asking that the multiplicities of the wave velocities do not depend on the direction.

Since $SM = 0$, we see that $Q(N)$ depends only on n_{12}, n_{21} and n_{22}. Thus let n_{12}, n_{21} and n_{22} be given, with $n_{22} \neq 0$; we may choose n_{11} in such a way that $\det N = 0$, whence $Q(N) \geq 0$. Therefore the condition $n_{22} \neq 0$ implies $Q(N) \geq 0$. By continuity, we obtain that S is positive semi-definite. It is not positive definite since $SM = 0$.

At last, we remark that not only $AM = 0$, but also

$$A^T JM = -SJJM = -SM = 0.$$

Hence the subspace $(JM)^\perp$ is invariant under A, and the spectrum of A consists in $\lambda = 0$ (the eigenvalue of A^T associated to JM), together with the spectrum of the restriction of A to $(JM)^\perp$. But the latter contains again $\lambda = 0$, because $M \in (JM)^\perp \cap \ker A$. Hence $\lambda = 0$ is a multiple eigenvalue of A, at least algebraically. However, the rank of A equals that of S. Therefore, assuming that $\dim \ker S = 1$, that is $\mathrm{rk} S = 3$, the geometric multiplicity of zero as an eigenvalue of A is one. This proves that the eigenvalue is not semi-simple.

QED

Compatibility of the Rankine–Hugoniot relations

We prove here what we claimed in the previous sections.

Theorem 7 *Let (B, D) be a piecewise smooth solution of the Maxwell's system (8). Hence $(B, D, P := B \times D)$ is a solution of the enlarged system, except perhaps accross discontinuities.*

Assume moreover that $(B, D, P = B \times D)$ satisfies the jump relation for (14) (this means that it is a weak solution of the enlarged system.) Then (B, D) also satisfies the Poynting equation (9)

In other words, a field that satisfies the enlarged system and that keeps $P \equiv B \times D$ does not have dissipative shocks. We expect that its discontinuities are contacts.

Proof.

Let $(B, D, B \times D)$ be a discontinuous solution of the enlarged system. Obviously, (B, D) is a weak solution of the Maxwell system. Consider a discontinuity accross a smooth hypersurface. We denote by ν the unit normal to the surface, and σ its normal velocity. The Rankine–Hugoniot relations for (12, 13, 14) are

$$\begin{aligned} \sigma[B] &= -[E \times \nu], \\ \sigma[D] &= [H \times \nu], \\ \sigma[P] &= [(D \cdot \nu)\phi_D + (B \cdot \nu)\phi_B - (\phi_P \cdot \nu)P] \\ &\quad +[\phi - B \cdot \phi_B - D \cdot \phi_D - P \cdot \phi_P]\nu. \end{aligned}$$

Starting from these identities, plus the fact that $P \equiv B \times D$, we have to show that the jump condition $[E \times H] \cdot \nu = \sigma[\phi]$ associated to (9) holds true.

As usual, $[g] := g^+ - g^-$ is the jump of a quantity g. We shall also use the notation

$$\langle g\rangle := \frac{1}{2}(g^+ + g^-).$$

We point out that, for every bilinear map Q, there holds

$$[Q(g,h)] = Q([g],\langle h\rangle) + Q(\langle g\rangle,[h]).$$

To begin with, we eliminate the derivatives ϕ_B and ϕ_D by using the equations of state, and the vector P by using the assumption. This yields to the following form of the third jump relation:

$$\begin{aligned}\sigma[B\times D] &= [(D\cdot\nu)E + (B\cdot\nu)H] + [\phi - B\cdot H - D\cdot E + (B\times D)\cdot\phi_P]\nu\\ &\quad +[(D\cdot\nu)B\times\phi_P + (B\cdot\nu)\phi_P\times D + (\phi_P\cdot\nu)D\times B].\end{aligned}$$

Because of circular symmetry, the brackets in the last line equals $[\det(D,B,\phi_P)]\nu$. Hence there remains

$$\sigma[B\times D] = [(D\cdot\nu)E + (B\cdot\nu)H] + [\phi - B\cdot H - D\cdot E]\nu.$$

Let us develop the bilinear terms. First of all:

$$\sigma[B\times D] = \sigma([B]\times\langle D\rangle + \langle B\rangle\times[D]).$$

Together with the Rankine–Hugoniot relations, that gives

$$\sigma[B\times D] = \langle B\rangle\times[H\times\nu] + \langle D\rangle\times[E\times\nu].$$

Next,

$$[(D\cdot\nu)E] = \langle D\cdot\nu\rangle[E] + [D\cdot\nu]\langle E\rangle,$$

and similarly

$$[(B\cdot\nu)H] = \langle B\cdot\nu\rangle[H] + [B\cdot\nu]\langle H\rangle.$$

Using then the formula

$$X\times(Y\times Z) = (X\cdot Z)Y - (X\cdot Y)Z, \tag{28}$$

there comes

$$[\phi - B\cdot H - D\cdot E]\nu = -(\langle B\rangle\cdot[H])\nu - (\langle D\rangle\cdot[E])\nu - [B\cdot\nu]\langle H\rangle - [D\cdot\nu]\langle E\rangle.$$

Developing again, we obtain

$$[\phi]\nu = ([B]\cdot\langle H\rangle)\nu + ([D]\cdot\langle E\rangle)\nu - [B\cdot\nu]\langle H\rangle - [D\cdot\nu]\langle E\rangle.$$

Then using again Formula (28), we have

$$[\phi]\nu = [B]\times(\nu\times\langle H\rangle) + [D]\times(\nu\times\langle E\rangle).$$

We multiply by σ and use again the Rankine–Hugoniot relation, to end with the equivalent relation

$$\sigma[\phi]\nu = [E \times \nu] \times \langle H \times \nu\rangle + \langle E \times \nu\rangle \times [H \times \nu].$$

This exactly means that

$$\sigma[\phi]\nu = [(E \times \nu) \times (H \times \nu)],$$

or in other words

$$(\sigma[\phi] + [H \times E] \cdot \nu)\nu = 0,$$

which implies the desired identity.

QED

Examining these calculations, we see that we have proved the following. For every discontinuous field that satisfies

$$\sigma[B] = -[E \times \nu], \quad \sigma[D] = [H \times \nu] \tag{29}$$

plus the equations of state, there holds

$$\begin{aligned}\sigma^2[B \times D] &= \sigma[(D \cdot \nu)\phi_D + (B \cdot \nu)\phi_B - (\phi_P \cdot \nu)B \times D] \\ &\quad -\sigma[B \cdot \phi_B + D \cdot \phi_D + (B \times D) \cdot \phi_P]\nu + ([E \times H] \cdot \nu)\nu.\end{aligned}$$

Assume now that (B, D) is an *admissible* weak solution of Maxwell's system, meaning that it satisfies (29), together with the "entropy" inequality

$$\epsilon := \sigma[\phi] - [E \times H] \cdot \nu \geq 0.$$

Then we derive, denoting $P := B \times D$,

$$\begin{aligned}\sigma^2[P] &= \sigma[(D \cdot \nu)\phi_D + (B \cdot \nu)\phi_B - (\phi_P \cdot \nu)P] \\ &\quad +\sigma[\phi - B \cdot \phi_B - D \cdot \phi_D - P \cdot \phi_P]\nu - \epsilon\nu.\end{aligned} \tag{30}$$

This identity may be converted into an integral formula. Assume that (B, D) is smooth away from a smooth hypersurface $\Sigma \subset (0, \tau) \times \Omega$. Then let $\theta \in \mathcal{D}((0, \tau) \times \Omega)^3$ be a test field. Then there holds

$$\int_0^\tau dt \int_\Omega (P \cdot \partial_t\theta + T : \nabla_x\theta)\, dx = \int_0^\tau dt \int_{\Sigma(t)} \frac{\epsilon}{\sigma}\theta \cdot \nu\, dS(x). \tag{31}$$

In particular, we have a new kind of entropy inequality:

$$\left(\theta \cdot \frac{\nu}{\sigma} \geq 0 \text{ on } \Sigma\right) \Longrightarrow \left(\int_0^\tau dt \int_\Omega (P\partial_t\theta + T : \nabla_x\theta)\, dx \geq 0\right). \tag{32}$$

Open questions

We list now a few open questions that seem of some mathematical interest.

i. What is the physical meaning of (32) ? Does it always make sense, or can the normal velocity σ vanish ?

ii. What are the wave speeds in either the Maxwell's equations or the enlarged system ? So long as we restrict to the equilibrium manifold, the Maxwell's velocities are part of the "enlarged" velocities. The three extra velocities can be computed by linearizing (19) around a constant solution that is at equilibrium. We obtain

$$\partial_t \delta' = \left(\frac{\partial \phi}{\partial P} \cdot \nabla \right) \delta',$$

where δ' stands for the infinitesimal perturbation of δ, namely $\delta' = P' - B' \times D - B \times D'$, with obvious notations. Therefore the extra velocities merge into a unique one with multiplicity 3,

$$\lambda(U; \xi) := -\frac{\partial \phi}{\partial P} \cdot \cdot \xi.$$

Does this multiplicity persist away from the equilibrium manifold ? If it did, the corresponding characteristic field would be linearly degenerate and "integrable", according to a theorem of G. Boillat [3] (see also [15] vol I, page 81.) This does not seem to be the case.

iii. Identify, among the energies W that come from an invariant Lagrangian $L(\|E\|^2 - \|B\|^2, E \cdot B)$, those which can be written as convex functions of $(B, D, B \times D)$. We know that the Born–Infeld energy works. Presumably, a small and localized disturbance of L_{BI} yields an admissible energy. The difficulty here is that, given an energy, there is a lot of freedom when writing it as a function of $(B, D, B \times D)$, since we are completely free outside the equilibrium manifold $P = B \times D$, a non-convex set.

Given the Lagrangian $L(\gamma, \delta)$, with $\gamma := (\|E\|^2 - \|B\|^2)/2$ and $\delta := E \cdot B$, there is however a "natural" (although non unique) way to define $\phi(B, D, P)$ such that $W(B, D) \equiv \phi(B, D, B \times D)$. We shall assume that L is even with respect to δ, which means that it is Lorentz- and orientation-invariant. We start form Definition (10). Given B and D, we write

$$\sup_{e \in \mathbb{R}^3} = \sup_{\gamma, \delta} \sup_{e \backslash \gamma, \delta},$$

where $\sup_{e \backslash \gamma, \delta}$ is a supremum over e, constrained by $(\|e\|^2 - \|B\|^2)/2 = \gamma$ and $e \cdot B = \delta$. To begin with, we solve this sub-problem, where L remains constant. The maximum of $D \cdot e$ is achieved at some point e that belong to the plane spanned by

B and D. With the two constraints, the possible points are the intersections of a sphere with a line,

$$e = \frac{\delta}{\|B\|^2} B + a(B \times D) \times B,$$

where a obeys to

$$\frac{\delta^2}{\|B\|^2} + a^2 \|B\|^2 \|B \times D\|^2 = \|B\|^2 + 2\gamma.$$

The supremum is achieved when a is positive. We obtain

$$\sup_{e \backslash \gamma,\delta} \{D \cdot e - L(\gamma,\delta)\} = \frac{\delta\, B \cdot D + \|B \times D\| \sqrt{\|B\|^4 + 2\gamma \|B\|^2 - \delta^2}}{\|B\|^2} - L(\gamma,\delta).$$

Since L is even with respect to δ, we may replace $B \cdot D$ by its absolute value. Then the expression $|B \cdot D|$ equals $(\|B\|^2 \|D\|^2 - \|B \times D\|^2)^{1/2}$. Finally, we may write

$$W(B,D) = h(\|B\|, \|D\|, \|B \times D\|), \tag{33}$$

with

$$h(b,d,p) := \sup_{\gamma,\delta} \left\{ \frac{\delta \sqrt{b^2 d^2 - p^2} + p \sqrt{b^4 + 2\gamma b^2 - \delta^2}}{b^2} - L(\gamma,\delta) \right\}. \tag{34}$$

The convexity of $\phi(B,D,P) := h(\|B\|, \|D\|, \|P\|)$ is equivalent to that of h. Hence we obtained a sufficient condition (a rather obscure one, indeed) in order that an enlarged system with a convex energy exist. Can we make this condition more explicit ? Is this condition necessary ? We leave these questions open. Remark that formula (34) can be used to find H in the Born–Infeld model:

$$L_{BI} = -\sqrt{1 + \|B\|^2 - \|E\|^2 - (E \cdot B)^2}, \quad W_{BI} = \sqrt{1 + \|B\|^2 + \|D\|^2 + \|P\|^2}.$$

Notice that Formula (34) also reads

$$h(b,d,p) := \frac{1}{b} \sup \left\{ \rho \sqrt{b^2 d^2 - p^2}\, \cos\theta + p\rho \sin\theta - bL\left(\frac{\rho^2 - b^2}{2}, \rho b \cos\theta\right) \right\},$$

where the supremum is taken over $\rho \geq 0$ and $\theta \in [0, \pi/2]$.

Remark: The fact that an energy W depends only on $\|B\|$, $\|D\|$ and $B \cdot D$ has been shown above under the assumption that W derives from a Lorenzt-invariant Lagrangian. Actually, this is true in a much more general context, whenever the physics is isotropic, for instance in an isotropic non-linear medium. For an orthogonal change of coordinates must preserve W, while (B,D) are changed into (QB, QD) for some $Q \in \mathbf{O}_3$.

In is interesting to see that, provided $W(B,D)$ has the form $H(\|B\|, \|D\|, B \cdot D)$, the reduction of the energy to a level set of $B \cdot \xi$ and $D \cdot \xi$ (ξ a given unit vector) admits a representation of similar form. For instance, taking $\xi = \vec{e}_1$, thus freezing B_1 and D_1 and writing $B = (B_1, \bar{B})$, $= (D_1, \bar{D})$, we have

$$\|B\| = \sqrt{B_1^2 + \|\bar{B}\|^2}, \quad \|D\| = \sqrt{D_1^2 + \|\bar{D}\|^2}, \quad B \cdot D = B_1 D_1 + \bar{B} \cdot \bar{D}.$$

Hence

$$W = \bar{H}(B_1, D_1; \|\bar{B}\|, \|\bar{D}\|, \bar{B} \cdot \bar{D}),$$

where B_1, D_1 play the role of parameters.

Planar waves

We consider now solutions of Maxwell's equations that depend only on time and a single space variable, say x_1. Then B_1 and D_1 depend only on time and may be considered as prescribed data. To avoid complications due to inhomogeneity, we assume that B_1 and D_1 are constant initially, hence constant forever. Without loss of generality, we may assume that $B_1 \equiv D_1 \equiv 0$, thanks to the previous remark. Denoting $x := x_1$, Maxwell's equations reduce to

$$\begin{aligned} &\partial_t B_2 - \partial_x(\partial W/\partial D_3) = 0, \qquad &\partial_t B_3 + \partial_x(\partial W/\partial D_2) = 0, \\ &\partial_t D_2 + \partial_x(\partial W/\partial B_3) = 0, \qquad &\partial_t D_3 - \partial_x(\partial W/\partial B_2) = 0. \end{aligned}$$

We assume a form (33) for the energy, which is equivalent to the form $H(\|B\|, \|D\|, |B \cdot D|)$. Then the above system rewrites, in complex variables $w := B_2 + iB_3$ and $z := D_3 - iD_2$,

$$\begin{aligned} \partial_t w - \partial_x(h_p w) - \partial_x\left(\frac{1}{d} h_d z\right) &= 0, \\ \partial_t z - \partial_x(h_p z) - \partial_x\left(\frac{1}{b} h_b w\right) &= 0. \end{aligned}$$

In general, this system of four equations does not decouple into closed proper sub-systems. We remind the reader that the analysis done in [14] concluded to the existence of a weak entropy solution for any bounded initial data, provided W has the form of a function of $(\|B\|^2 + \|D\|^2)^{1/2}$, with suitable convexity properties. But such an assumption fits hardly with the requirement that W comes from an invariant Lagrangian. Therefore, we wish to relax it. The main property that we wish to preserve is the decoupling of the system. It turns out that an energy of the form

$$W(B, D) = h(r, p), \quad r = \sqrt{b^2 + d^2}$$

permits this simplification. For then the system rewrites

$$\begin{aligned} \partial_t(w + z) - \partial_x(h_p(w + z)) - \partial_x\left(\frac{1}{r} h_r(w + z)\right) &= 0, \\ \partial_t(w - z) - \partial_x(h_p(w - z)) + \partial_x\left(\frac{1}{r} h_r(w - z)\right) &= 0. \end{aligned}$$

Hence, writing $w+z =: \rho\exp(i\theta)$ and $w-z =: \sigma\exp(i\alpha)$ (polar decomposition of complex numbers), we obtain a 2×2 system in (ρ,σ):

$$\partial_t\rho - \partial_x(h_p\rho) - \partial_x\left(\frac{1}{r}h_r\rho\right) = 0, \tag{35}$$

$$\partial_t\sigma - \partial_x(h_p\sigma) + \partial_x\left(\frac{1}{r}h_r\sigma\right) = 0. \tag{36}$$

The fact that the above system is closed follows from the identities

$$2r^2 = \rho^2+\sigma^2, \quad 4p = \rho^2-\sigma^2.$$

We emphasize that the energy (in)equality (9) reads in terms of (ρ,σ) only and hence can be used as an entropy criterion for the system (35, 36):

$$\partial_t h(r,p) + \partial_x\left((r^{-2}h_r^2+h_p^2)p - 2rh_rh_p\right) \le 0.$$

In particular, our analysis above gives us the non-trivial fact that the strict convexity of h implies the hyperbolicity of the sub-system.

We postpone the study of the Cauchy problem for this 2×2 system to a future work. For the moment, let us just say that, given a weak entropy solution of (35, 36) that is non-negative ($\rho\ge 0$, $\sigma\ge 0$), we may build a weak entropy solution of the plane wave system by solving the following transport equations

$$\left(\partial_t - h_p\partial_x - \frac{1}{r}h_r\partial_x\right)\theta = 0, \quad \left(\partial_t - h_p\partial_x + \frac{1}{r}h_r\partial_x\right)\alpha = 0. \tag{37}$$

We recall that, following the procedure in [14], we actually may solve the conservation laws

$$\partial_t(\rho f(\theta)) - \partial_x(h_p\rho f(\theta)) - \partial_x\left(\frac{1}{r}h_r\rho f(\theta)\right) = 0,$$

$$\partial_t(\sigma g(\alpha)) - \partial_x(h_p\sigma g(\alpha)) + \partial_x\left(\frac{1}{r}h_r\sigma g(\alpha)\right) = 0,$$

for every smooth functions f and g simultaneously. The choices of the sine and cosine functions give exactly the Maxwell's equations for planar waves.

Remarks and questions: - Under the assumption that the reduced energy h depends only on (r,p), the system of planar waves is equivalent, at a formal level, to (35, 36, 37). In particular, the wave velocities $-h_p \pm r^{-1}h_r$ are linearly degenerate. The fact that the system also contains the transport equations $\partial_t B_1 = 0$ and $\partial_t D_1 = 0$ indicates that there are actually four "linear" velocities in this one-dimensional model. Is it true that reasonable models, say with $W = H(\|B\|,\|D\|,B\cdot D)$ always admit four linearly degenerate velocities in each direction ? - Which energies of the form $h(r,p)$ derive from an invariant Lagrangian $L(\gamma,\delta)$.

Acknowledgements. The author thanks warmly C. Dafermos and Y. Brenier for fruitful discussions, and the later for providing a preliminary version of [5].

References

[1] J. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rational Mech. Anal.*, **63** (1977), pp 337–403.

[2] S. Benzoni–Gavage, D. Serre. *First-order hyperbolic systems of partial differential equations. With applications.* In preparation.

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

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

[5] Y. Brenier. Hydrodynamic structure of the augmented Born–Infeld equations. Preprint (2003).

[6] Ph. Ciarlet. *Mathematical elasticity.* North–Holland, Amsterdam (1988).

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

[8] C. Dafermos. Quasilinear hyperbolic systems with involutions. *Arch. Rational Mech. Anal.*, **94** (1986), pp 373–389.

[9] C. Dafermos. *Hyperbolic conservation laws in continuum physics.* Grundlehren der mathematischen Wissenschaften, **325**. Springer-Verlag, Heidelberg (2000).

[10] S. Demoulini, D. Stuart, A. Tzavaras. A variational approximation scheme for three–dimensional elastodynamics with polyconvex energy. *Arch. Rational Mech. Anal.*, **157** (2001), pp 325–344.

[11] S. Godunov. Lois de conservation et intégrales d'énergie. *Lecture Notes in Math.* **1270**, pp 135–149. Springer-Verlag, Heidelberg (1987).

[12] G. Métivier. Problèmes de Cauchy et ondes non linéaires. *Journées EDPs de St-Jean-de-Monts* (1986). Available on http://www.numdam.org .

[13] D. Serre. Formes quadratiques et calcul des variations. *J. Maths. Pures Appl.* **62** (1983), pp 177–196.

[14] D. Serre. Les ondes planes en électromagnétisme non-linéaire. *Physica D***31** (1988), pp 227–251.

[15] D. Serre. *Systems of conservation laws*. Cambridge Univ. Press, Cambridge (vol I: 1999, vol II: 2000).

[16] M. Taylor. *Pseudodifferential operators*. Princeton University Press, Princeton (1981).

[17] F. J. Terpstra. Die Darstellung biquadratischer Formen als Summen von Quadraten mit Anwendung auf die Variationsrechnung. *Math. Ann.***116** (1938), pp 166–180.