On an equation from the theory of field dislocation mechanics

Amit ACHARYA & Luc TARTAR*

to the memory of Professor Giovanni PRODI

Abstract.

Global existence and uniqueness results for a quasilinear system of partial differential equations in one space dimension and time representing the transport of dislocation density are obtained. Stationary solutions of the system are also studied, and an infinite dimensional class of equilibria is derived. These time (in)dependent solutions include both periodic and aperiodic spatial distributions of smooth fronts of plastic distortion representing dislocation twist boundary microstructure. Dominated by hyperbolic transport-like features and at the same time containing a large class of equilibria, our system differs qualitatively from regularized systems of hyperbolic conservation laws and neither does it fit into a gradient flow structure.

Introduction.

Since this article is written for a special issue of the Bolletino dell'Unione Matematica Italiana, in memory of Giovanni PRODI, the introduction is written in the first person, by the second author.

During my second year as a graduate student, in Paris in 1968–1969, my advisor, Jacques-Louis LIONS, gave a course on nonlinear partial differential equations, and he wrote a book in this way [Li]. After a secretary had typed his handwritten notes, he asked me to proof-read one chapter at a time, and it was then that I first encountered the name of Giovanni PRODI, who was quoted for questions about Navier–Stokes equation. I had already taken the habit of reading very little, and I could not have read an article in Italian at the time, but I did not even try to read the joint work of Giovanni PRODI and my advisor, a note to the Comptes Rendus de l'Académie des Sciences, written in French, of course.

I am not sure when I first met Giovanni PRODI, and I suppose that it was in Pisa when I first visited the Scuola Normale Superiore at the invitation of Ennio DE GIORGI, and I remember well a visit in February 1982, just after having visited my good friend Carlo SBORDONE in Napoli at the time of carnevale, and while in Pisa some new friends took me to carnevale in Viareggio, which I could not appreciate so well since I did not know the politicians who were caricatured, but I enjoyed the fireworks afterward, the best I ever saw. I could well have been in Pisa for a short visit before that occasion.

Since I first associated the name of Giovanni PRODI to Navier–Stokes equation, I thought of writing an article in his memory on this subject, but I have no important remark to make that I have not already written in my four books published by Springer in the lecture notes series of Unione Matematica Italiana, the first one on Navier–Stokes equation [Ta3], the second one on Sobolev spaces and interpolation spaces [Ta4], the third one on kinetic theory [Ta5], and the fourth one on homogenization [Ta6].

However, I find useful to emphasize again that it gives a bad impression about mathematicians that after convincing a philanthropist like L. T. CLAY to give one million dollars away for a prize on the Navier–Stokes equation, those in charge of writing on that problem could not find a mathematician competent enough to tell them what the Navier–Stokes equation is about. I certainly received an excellent training in Paris, but I thought that everyone learned that the basic laws of continuum mechanics are about conservation of mass, conservation of linear momentum, conservation of angular momentum (usually taken care of by the symmetry of the Cauchy stress tensor), and conservation of energy, together with some constitutive equations forming equations of state, although I was not told in a clear way that the validity of equations of state is questionable out of equilibrium, which was the reason why kinetic theory was invented, by MAXWELL, and by BOLTZMANN.

Then, since conservation of mass gives an hyperbolic equation $\rho_t + div(\rho\, u) = 0$ for the density (of mass) ρ, one avoided this "difficulty" by assuming incompressibility, $\rho = \rho_0$, which has the unfortunate

* Amit ACHARYA, Department of Civil and Environmental Engineering, Luc TARTAR, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A. Research supported by ONR Grant N000140910301.

effect of making unphysical the "pressure" p appearing in the equation, since after that it is only a Lagrange multiplier defined up to addition of a "constant", depending upon t, and although the speed of sound at usual atmospheric pressure (and reasonable temperature) is about 300 meters per second for air and about 1500 meters per second for water, the incompressibility hypothesis makes it infinite. Then, one forgot to tell why the equation of conservation of energy was not considered, which is that one assumed the dynamic viscosity μ to be constant, without saying it and without observing that having it independent of temperature is quite contrary to evidence, so that one finally only considers an equation for the velocity field u, which uses a reduced "pressure" $\frac{p}{\rho_0}$, and where the kinematic viscosity $\nu = \frac{\mu}{\rho_0}$ appears. Then, trying to describe the group of transformations which leave the equation invariant, besides translation in space or time, and scaling, one forgot to mention invariance by rotation, since fluids like air or water are isotropic, and one forgot to mention Galilean invariance, which makes the difference between the slightly unphysical Stokes equation and the more physical Navier–Stokes equation (first introduced in 1821 by NAVIER using a molecular approach, and then derived more mathematically using stress, in 1843 by SAINT-VENANT, and in 1845 by STOKES). Then, one chose domains without boundary, like the not so physical $\mathbb{R}^3$ or the unphysical torus $\mathbb{T}^3$, although the mathematical problem is to find good bounds for the vorticity, and that it is observed that vorticity is created at the boundary.

I am not sure why some harmonic analysts are fascinated by the scaling property, since it only says that if one changes the unit of length and the unit of time, then the unit of velocity is automatically changed, and the scaling property is just observing that the unknown u is a velocity. Of course, there could be unexpected consequences of scaling, and Sobolev's embedding theorem is a classical one, about which I have introduced a general method for its extension to Lorentz spaces, published in this journal [Ta1], and my work is in the spirit of what Jaak PEETRE had done,[1] and also extends a result of Haïm BREZIS and Stephen WAINGER, but although it uses a classical scaling argument for the domain $x \in \mathbb{R}^N$, it uses a nonlinear scaling for the target space $\mathbb{R}$ by considering various functions $\varphi(u)$, and adapting the choice of φ to the rearrangement of $|u|$; I think that it would be a generalization of that type of argument and not of the classical scaling that could be useful, so that I expect that a different kind of scaling argument could be discovered.

Giovanni PRODI participated in the 1950s and 1960s in the "rinascita" of Italian mathematics together with a few other colleagues, some of whom I met a few times, like Guido STAMPACCHIA and Ennio DE GIORGI who died before him, or Enrico MAGENES who died more recently, and in some way my decision to write lecture notes for the graduate courses which I taught (after 1999) followed a similar plan to revitalize a way of doing mathematics with a serious interaction with continuum mechanics and physics, in the spirit of POINCARÉ, HADAMARD, and Jean LERAY in France; I do not put Laurent SCHWARTZ in this category although his theory of distributions helped give a more general framework than what Sergei SOBOLEV had done for problems in continuum mechanics or physics, nor do I put my advisor either, since I could not find in them a true will to improve the understanding of continuum mechanics or physics through the development of new mathematical tools, which should permit to go a step further, and correct some of the "mistakes" made by the practitioners. However, since engineers and physicists are not mathematicians, they are allowed to guess results without calling them conjectures, so that a "mistake" may just be what a mathematician calls a conjecture which is proven to be false, but I shall measure the success of my plan by the number of mathematicians who will understand the errors in reasoning which were made for arriving at some popular models, and hopefully have the courage to mention such mistakes, and suggest ways for finding better models.

Since no realistic approach to plasticity in solids can be made without knowing about dislocations, the point of view that we have taken is to start from a reasonable description of physical reality, the theory of field dislocation mechanics, of which the first author is a specialist, and try to solve some of the systems of partial differential equations which appear in this theory, by improving the actual mathematical methods, but since such a goal has not been attained yet, we have looked at simpler systems in one space dimension, for which some variants of more or less classical methods can be used. We first present the general theory, and then explain how to derive the simpler models which will be treated mathematically afterward, and

[1] I realized afterward that Jaak PEETRE could not have proven my generalization $W^{1,1}(\mathbb{R}^N) \subset L^{1^*,1}(\mathbb{R}^N)$, since he followed the method of Sergei SOBOLEV, which gives $W^{1,p}(\mathbb{R}^N) \subset L^{p^*}(\mathbb{R}^N)$ only for $1 < p < N$, and then I did not attribute correctly the case $p = 1$ of "Sobolev's embedding theorem", which was proven independently by Emilio GAGLIARDO and by Louis NIRENBERG.

it would certainly be desirable that some mathematicians pretending to work on realistic problems would follow this scheme instead of specializing in unrealistic questions and then look for inadequate models of the kind that they like which were used in the past, due to a lack of a better alternative, and pretend that it is still important to solve them, without explaining in what way they are now known to be deficient, and which mathematical questions about them are nevertheless useful to be settled.

Although the equation that we consider looks like an unconventional regularization (certainly for those who never questioned the "artificial viscosity method") of an hyperbolic system of conservations laws, transport is the main feature of our system, with a weaker dissipation effect for $\varepsilon > 0$ than in hyperbolic systems of conservations laws, for which an up-to-date reference is the book of my good friend Constantine DAFERMOS [Da], where he describes various types of E-conditions for selecting which discontinuities one wants to accept. One should then be careful in comparing our results to what happens for hyperbolic system of conservations laws, as we show in sections 3 and 4.

1. **Field Dislocation Mechanics.**

Field Dislocation Mechanics, denoted FDM afterwards, was developed by the first author in [Ac1], [Ac2], [Ac3], [Ac4], [Ac5], as a pde-model for understanding plasticity of solids as it arises from the nucleation, the motion, and the interaction of defects, in an otherwise elastic motion of a material. FDM advances the pioneering work of KRÖNER [Kr], MURA [Mu], FOX [Fo], and John WILLIS [Wi] to produce the first pde model for the prediction of coupled dislocation internal stress, dislocation microstructure evolution, and permanent deformation in generic bodies of finite extent. A unique feature of FDM is the treatment of dissipative dislocation transport accounting for nonlinearity due to geometric and crystal elasticity effects. While simplifications of the model to a small deformation setting were developed and will be analyzed in this article, an essential feature of the general theory is to make no restrictions on the material and the geometric nonlinearities. It is perhaps fair to say that there is a general bias that, in the context of modeling of defect dynamics in an elastic material, dislocations should be thought of as discrete objects that are best studied as singularities within an otherwise linear theory, with their nonlinear interactions "put in by hand." The primary goals of FDM are: to understand such discreteness as field localizations without singularities that emerge naturally in nonlinear theory, whose interactions are decided by partial differential equations which encode conservation statements of a geometric and topological nature, coupled with nonlinear elasticity; and to understand the emergence of macroscopic plasticity as a homogenization of this underlying system of non-linear pde. This article is one of the first rigorous studies (cf. [Ac&Ma&Zi]) of the dynamics of such nonsingular localizations arising in perhaps the most simplified, but exact, dynamic problem of FDM, which nevertheless appears to be surprisingly rich. For instance, NABARRO [Na] states "The attempt to build up a dislocation theory while neglecting the non-Hookean forces which hold a dislocation together and prevent its spreading thinly over the glide plane is bound to encounter difficulties similar to the 'purely' electromagnetic theory of the electron.". As we show, in FDM, even with linear elasticity, it is possible to hold a dislocation together and this is because of nonlinearity and non-scalar nature of dislocation transport.

The full 3-dimensional, small deformation theory is presented first (for details, see [Ac4]) and an exact simplification for a time-dependent system in one space dimension is then derived. The physical context that this simplification represents is described, and primary questions of interest related to the system are mentioned.

The notation used is as follows: for a second-order tensor A, a vector v, and a spatially constant vector field c,

$$
\begin{aligned}
(A \times v)^T c &= (A^T c) \times v \text{ for all } c \\
(divA) \cdot c &= div(A^T c) \text{ for all } c \\
(curlA)^T c &= curl(A^T c) \text{ for all } c.
\end{aligned}
$$

In rectangular Cartesian components,

$$
\begin{aligned}
(A \times v)_{im} &= e_{mjk} A_{ij} v_k \\
(divA)_i &= A_{ij,j} \\
(curlA)_{im} &= e_{mjk} A_{ik,j}
\end{aligned}
$$

where one sums over repeated indices, e_{mjk} is a component of the *third-order alternating tensor* X, and $_{,j}$ as an index means ∂_j. Also, the vector $X\,AB$ is defined as

$$(X\,AB)_i = e_{ijk} A_{jr} B_{rk}.$$

The spatial derivative, for the component representation is with respect to rectangular Cartesian coordinates on the body. For all manipulations with components, we shall always use such rectangular Cartesian coordinates. The symbol *div* represents the divergence, *grad* the gradient, and *div grad* the Laplacian. We have occasion to use the identity $curl\,curl(\cdot) = grad\,div(\cdot) - div\,grad(\cdot)$, often for an argument for which $div(\cdot) = 0$. A superposed dot represents a time derivative.

The complete set of equations is

$$\text{on } R \qquad \begin{cases} curl\,\chi = \alpha \\ div\,\chi = 0 \\ div(grad\,\dot{z}) = div(\alpha \times V) \\ U^e := grad(u - z) + \chi \quad ; \quad U^p := grad\,z - \chi \\ div[T(U^e)] = 0 \\ \dot{\alpha} = -curl(\alpha \times V) \end{cases} \tag{1}$$

where R is the body, and the various fields are defined as follows. χ is the incompatible part of the elastic distortion tensor U^e, u is the total displacement field, and $u - z$ is a vector field whose gradient is the compatible part of the elastic distortion tensor. U^p is the plastic distortion tensor. α is the dislocation density tensor, and V is the dislocation velocity vector. $\alpha \times V$ represents the flow of Burgers vector carried by the dislocation density field moving with velocity V relative to the material. The argument of the *div* operator in the fifth equation in (1) is the (symmetric) stress tensor, and the functions V, T are constitutively specified. All the statements in (1) are fundamental statements of kinematics or conservation. In particular, the sixth equation in (1) is a purely geometric statement of conservation of Burgers vector content carried by a density of lines (see [Ac5] for a derivation) and the fifth equation in (1) is the balance of linear momentum in the absence of inertia and body forces.

As for boundary conditions,

$$\text{on } \partial R, \text{ with } n \text{ denoting the normal,} \qquad \begin{cases} \chi\, n = 0 \\ (grad(\dot{z}) - \alpha \times V)\, n = 0 \end{cases} \tag{2}$$

are imposed along with standard conditions on displacement and/or traction. For the dislocation density field, analysis from the linear partial differential equation point of view [Ac2] indicates that it suffices to prescribe $\alpha\,(V \cdot n)$ on inflow parts of the boundary, but we believe that the nonlinear problem admits other physically motivated possibilities [Ac&Ro].

In order to define sensible, minimal constitutive relations, we require that the mechanical power supplied by external agencies always be greater than or equal to the rate at which energy is stored in the body at all instants of time. To ensure this, we define a stored-energy function,

$$\psi = \widehat{\psi}(U^e_{sym}) + \frac{\varepsilon}{2}\,\alpha : \alpha, \text{ where } \alpha : \alpha = \alpha_{ij}\alpha_{ij}, \tag{3}$$

in terms of which the stress is defined as

$$T = \frac{\partial \psi}{\partial U^e_{sym}},$$

and the "driving force" for the dislocation velocity is

$$V \to X\,(T + \varepsilon\,[curl\,\alpha]^T)\,\alpha. \tag{4}$$

In the above, $\varepsilon = \mu\, b^2$ is to be considered a small parameter, where μ is the shear modulus (from linearized elastic response) of the material and b is a typical interatomic spacing. The stored energy models the elastic straining of the material and energy contained in *dislocation cores*. In particular, the function $\widehat{\psi}$ could be non-convex in its argument. We now assume the simplest possible linear kinetic equation for the dislocation velocity in terms of its thermodynamic driving force:

$$V = \frac{X\left(T + \varepsilon\left[curl\,\alpha\right]^T\right)\alpha}{\mathrm{B}}, \tag{5}$$

where $\mathrm{B} > 0$ is a material constant called the "drag." It is in the determination of the response function ψ and V that FDM is designed to draw input from the apparatus of molecular dynamics or quantum mechanics. It can now be shown [Ac4] that solving the sixth equation of (1) is equivalent to

$$\dot{U}^p = curl\,U^p \times \frac{X\left(T + \varepsilon\left[curl\,\alpha\right]^T\right) curl\,U^p}{\mathrm{B}}, \tag{6}$$

and therefore solving (1)-(2) with the given constitutive equations is the same as solving (6) along with the fifth equation in (1) with the same constitutive equations and then solving for the fields $\chi, grad(z)$ from the first three equations in (1) and (2). It should be noted, however, that the sixth equation in (1) is the fundamental statement of defect evolution and (6) a derived construct, given the structure of (1)-(2) that is required in the nonlocal thermodynamic arguments to infer the dissipative driving force (4). This overall model based on the driving force (4) with $\varepsilon = 0$ was first proposed in [Ac2]; it was subsequently rediscovered and analyzed in [Li&Se1].

Our objective now is to solve the fifth equation in (1) and (6).

All tensor indices run from 1 to 3. We make the ansatz that $T_{ij}, i, j = 1, 2$ are the only non-vanishing (symmetric) stress components. Similarly, we assume that $U^p_{ij}, i, j = 1, 2$ are the only non-vanishing plastic distortion components. Additionally, we assume that all fields vary only in the x_3 direction. As a consequence we have $div(U^p) = 0$ in this special case, which implies

$$curl\,\alpha = -curl\,curl\,U^p = div\,grad\,U^p.$$

Thus (6) becomes

$$\dot{U}^p = curl\,U^p \times \frac{X\left(T + \varepsilon\left[div\,grad\,U^p\right]^T\right) curl\,U^p}{\mathrm{B}}. \tag{7}$$

We also note that under the stated assumptions, the fifth equation in (1) written in components as

$$T_{ij,j} = 0$$

is identically satisfied. Thus our task reduces to solving (7) with the assumed constitutive equation for the stress tensor. Since force equilibrium is identically satisfied, any displacement field u with a spatial variation in its gradient being only in the x_3 direction and one that does not produce any T_{3j} component would suffice to generate the elastic distortion U^e; for definiteness we assume $u = 0$ on R for all times.

The physical problem we think of modeling is a bar of uniform rectangular cross section in the (x_1, x_2) plane with axis being the x_3 direction. All lateral surfaces of the cylinder are rigidly constrained and when the bar is of finite length, it has free surfaces at the ends. The bar is assumed to contain a non-uniform initial plastic distortion field possibly containing dislocations, and we would like to understand how this dislocation distribution evolves and the type of spatial equilibria admitted by the model. Based on the assumptions, the only non-vanishing components of the dislocation density tensor,

$$\alpha_{r1} = U^p_{r2,3} \text{ for } r = 1, 2$$
$$\alpha_{r2} = -U^p_{r1,3} \text{ for } r = 1, 2$$

are uniform on any cross-section, possibly varying from one cross-section to another. Contact with discrete dislocations would have been made if the dynamics produces regions of length on the order of b where the dislocation density is non-vanishing with adjoining areas where it is zero or close to it. Thus we look for dislocation walls of finite width, corresponding to smooth fronts in the plastic distortion.

We would now like to write (7) in simpler notation adapted to the exact ansatz we have assumed. We first note that

$$(curl\, U^p)_{ri} = e_{i3k} U^p_{rk,3} = e_{i31} U^p_{r1,3} + e_{i3\,2} U^p_{r2,3},$$

so that

$$(curl\, U^p)_{r1} = -U^p_{r2,3} \text{ for } r = 1, 2, 3$$
$$(curl\, U^p)_{r2} = U^p_{r1,3} \text{ for } r = 1, 2, 3$$

are the only possibly non-zero components of $curl\, U^p$. Denote

$$T + \varepsilon\, [curl\, \alpha]^T = [T + \varepsilon\, div\, grad\, U^p]^T = A^T,$$

since T is symmetric. Next we expand

$$(X\, A^T B)_k = e_{kij} A_{ri} B_{rj},$$

assuming $A_{3i} = A_{i3} = B_{3i} = B_{i3} = 0$ for $i = 1, 2, 3$. With

$$\xi_k := (X\, A^T B)_k = e_{ki1} A_{ri} B_{r1} + e_{ki2} A_{ri} B_{r2},$$

this yields

$$\xi_1 = e_{1i2} A_{ri} B_{r2} = -A_{r3} B_{r2} = 0$$
$$\xi_2 = e_{2i1} A_{ri} B_{r1} = A_{r3} B_{r1} = 0$$
$$\xi_3 = e_{3i1} A_{ri} B_{r1} + e_{3i2} A_{ri} B_{r2} = -A_{r2} B_{r1} + A_{r1} B_{r2}$$

Choosing $B = curl\, U^p$, we have

$$\xi_3 = X\, (T + \varepsilon\, [div\, grad\, U^p]^T)\, curl\, U^p = (T_{r2} + \varepsilon\, U^p_{r2,33})\, U^p_{r2,3} + (T_{r1} + \varepsilon\, U^p_{r1,33})\, U^p_{r1,3}. \tag{8}$$

The only non-trivial statements of (7) now are

$$\begin{aligned} \mathrm{B}\, \dot{U}^p_{rm} &= e_{mnp} B_{rn} \xi_p = e_{mn3} B_{rn} \xi_3 \text{ for } r = 1, 2, \text{ i.e.} \\ \mathrm{B}\, \dot{U}^p_{r1} &= e_{123} B_{r2} \xi_3 = U^p_{r1,3} \xi_3 \text{ for } r = 1, 2, \\ \mathrm{B}\, \dot{U}^p_{r2} &= e_{213} B_{r1} \xi_3 = U^p_{r2,3} \xi_3 \text{ for } r = 1, 2. \end{aligned} \tag{9}$$

Finally, for $\psi = \psi(U^e_{sym}, \alpha)$, a class of function to which (3) belongs, we note that

$$(U^e_{sym})_{ij} = \frac{1}{2}\, [u_{i,j} - U^p_{ij} + u_{j,i} - U^p_{ji}]$$
$$\text{hence } \frac{\partial \psi}{\partial U^p_{rj}} = -\frac{\partial \psi}{\partial U^e_{mn}} \frac{\partial (U^e_{sym})_{mn}}{\partial U^p_{rj}} = -T_{rj}.$$

Denoting the coordinate $x_3 := x$ and introducing the array

$$\varphi = (U^p_{11}, U^p_{21}, U^p_{12}, U^p_{22}),$$

(9) is now expressed as

$$\varphi_t = \left(\varepsilon\, \varphi_{xx} - \frac{\partial \psi}{\partial \varphi}, \varphi_x \right) \varphi_x, \tag{10}$$

where we work in physical units such that $\mathrm{B} = 1$, $(\cdot, \cdot)$ represents the standard inner-product of two vectors in $\mathbb{R}^N$, and $\partial \psi / \partial \varphi$ is a function of (φ, φ_x) in general.

For the simplified problem being discussed here, $u_i = 0$. When linear elasticity (possibly anisotropic) is assumed in (3),

$$\widehat{\psi}(U^e_{sym}) = \frac{1}{2}\, C_{ijk\ell}(-U^p_{ij})\,(-U^p_{k\ell}), \text{ hence } \frac{\partial\widehat{\psi}}{\partial U^p_{rj}} = C_{rjk\ell}U^p_{k\ell},$$

where $C_{ijk\ell}$ has the minor symmetries, and

$$\frac{\partial\psi}{\partial\varphi} = K\,\varphi,$$

where K is a 4×4 matrix. K cannot be positive-definite because of the minor symmetries of C.

In the setting of linear elasticity with $\varepsilon = 0$, [Li&Se1], [Li&Se2] raise the interesting question of whether (10) and its 3-d general form, the fifth equation in (1) or (6), admit discontinuities in the plastic distortion field, with the intent of probing the dynamics of such discontinuities as grain boundaries and dislocation cell walls. They carry out their analysis numerically, utilizing a variety of numerical regularizations and conclude that in the system case, as opposed to the scalar, discontinuities do arise. Unfortunately, discontinuities of φ in (10) are problematic due to the nonlinearity in φ_x. At any rate, questions of whether singularities can arise in a pde-model can hardly be discussed by numerics, and one goal of this article is to answer this question with definiteness for the case with core energy. An analysis of traveling waves for the scalar case in (10) with non-convex ψ and $\varepsilon \neq 0$ has been carried out in [Ac&Ma&Zi], providing interesting physical insight into patterned dislocation wall equilibria.

2. **Why prove existence of solutions for good models in mechanics?**

In a book which Garrett BIRKHOFF edited [Bi], containing some translations into English of passages of articles in analysis from the 19th century (originally written in French, German, or Italian), he wrote (page 403) "Even Poincaré, ... , gave in 1890 a discussion of the heat and wave equations which concluded ignominiously with a physical "proof" of the completeness of the eigenfunctions of the Laplace operator, based on a "molecular hypothesis" which essentially said that solids could be treated as finite sets of particles.", but it is curious that Garrett BIRKHOFF wanted to insult POINCARÉ in this way and write a few lines after "By 1893, Poincaré had also established the completeness of the eigenfunctions of the Laplace equation using the still nascent theory of integral equations." with a footnote referring to [Po2], because if Garrett BIRKHOFF meant [Po1] as the first reference, it is clear from the last paragraph in [Po1] "Je pourrai dire alors que les conclusions des §§2, 3 et 4 sont démontrées d'un façon rigoureuse au point de vue physique. Peut-être même est-il permis d'espérer que, par une sorte de passage à la limite, on pourra fonder sur ces principes une démonstration rigoureuse même au point de vue analytique."[2] that POINCARÉ did not pretend to have proved his results from a mathematical point of view: he said that, with a sequel, he will have shown enough evidence to convince physicists, and he was actually quite cautious in conjecturing that a mathematical proof could be made using the same ideas (and he could have thought of inventing a new approach to the mathematical question, different from the formal argument based on physical intuition).

The questions which POINCARÉ raised in [Po1] about the necessity or the usefulness to give rigorous proofs (from a mathematical point of view) for physics problems which may just be approximations is still valid, but there is something that POINCARÉ was obviously not aware of, even after discovering what one now calls "chaos", which is an effect for ordinary differential equations.

An important message of the second author in his books [Ta3], [Ta5], [Ta6], is that 20th century mechanics (like plasticity or turbulence in continuum mechanics, and atomic physics or phase transitions in physics) seem to force upon us to go "beyond partial differential equations", because some homogenization problems show that the form of equations may change when one goes from one level to another, obviously so from the microscopic level of atoms to a mesoscopic level, but also possibly from one mesoscopic level to another, or from mesoscopic level to macroscopic level, and that some homogenized/effective equations might not be partial differential equations (and definitely not ode's). From the observed diversity of the constructions that nature builds at so many different scales, it seems clear that the mathematical tools

[2] What POINCARÉ wrote means "I shall be able to say that the conclusions of §§2, 3 and 4 are proved in a rigorous fashion from the physical point of view. Perhaps even is it allowed to hope that, by some kind of limit procedure, one will base on these principles a rigorous proof even from the analytical point of view.".

(beyond partial differential equations) which will permit to understand how this diversity comes out of a unified framework have not been developed yet, but one may hope to find a hierarchy of simpler models which permit to understand in a better way some of the questions which puzzled the previous generation of researchers. It is then important that mathematicians develop more efficient tools for proving or disproving the validity of some approximations, for the new models which specialists of continuum mechanics and physics propose, and for proposing better models than those old ones for which some limitations have already been pointed out.

3. **Stationary solutions for the model with $\varepsilon > 0$.**

There is no clear interpretation to what (10) could mean if φ takes its values in $\mathbb{R}^d$ (with dimension $d > 1$) and is allowed to be discontinuous, and a numerical approach (like that of [Li&Se1] and [Li&Se2]) cannot resolve this question, since one may easily confuse a very sharp layer where φ goes continuously from a_- to a_+ with a jump of φ from a_- to a_+.

At the moment, the only reasonable framework for allowing discontinuities in partial differential equations *of continuum mechanics* is to write equations in conservative form, and use the definition of derivatives in the sense of distributions (of Laurent SCHWARTZ), although STOKES was the first to derive a jump condition in 1848, followed by RIEMANN, who did it independently in 1860, but the jump conditions which must be imposed are now named after RANKINE and HUGONIOT; without the mention of any of these names, one should explain how a discontinuous function may satisfy a nonlinear partial differential equation.

One should pay attention to the example of the function $w(x) = sign(x)$ on $\mathbb{R}$, which is discontinuous and satisfies $w_x = 2\delta_0$, and since $w^2 = 1$, one can multiply w^2 by a Dirac mass; however, since $w^3 = w$, one has $(w^3)_x = 2\delta_0$, although $3w^2 w_x = 6\delta_0$. If one changes the nonlinearity w^3 to $f(w)$ with f smooth, then $\big(f(w)\big)_x = \big(f(+1) - f(-1)\big)\,\delta_0$, but $f'(w)\,w_x$ has no meaning unless $f'(+1) = f'(-1)$; if instead, the function w takes the value a on $(-\infty, 0)$ and the value b on $(0, +\infty)$, the term $(w^3)_x$ has a meaning, and is $(b^3 - a^3)\,\delta_0$, while $3w^2 w_x$ has no meaning if $|a| \neq |b|$.

We shall then only accept discontinuous solutions for equations written in conservative form, and since (for $\varepsilon > 0$) the jth component of (10) shows the quantity $\varphi_{jx} \sum_k \varphi_{kx}\varphi_{kxx}$, which cannot be written as $\big(A_j(\varphi, \varphi_x)\big)_x + B_j(\varphi, \varphi_x)$, we shall not allow φ_x to be discontinuous either. It is then natural to look for a solution such that φ_x is continuous, and such that the derivative of $\frac{|\varphi_x|^2}{2}$ in the sense of distributions, which is the meaning given to the quantity $(\varphi_{xx}, \varphi_x)$, is locally integrable.[3]

In this section, we look for stationary solutions on an interval $(-L, +L)$, which are constant on $(-\infty, -L)$ and on $(+L, +\infty)$, and satisfy Neumann conditions at $\pm L$, so that solving (10) in $(-L, +L)$ creates a solution for $x \in \mathbb{R}$. We make the restrictive assumption that ψ only depends upon φ_x in the separated form (3), i.e. one rewrites (10) as

$$\varphi_t + \Big(\frac{\partial\widehat{\psi}}{\partial\varphi} - \varepsilon\,\varphi_{xx}, \varphi_x\Big)\,\varphi_x = 0, \tag{11}$$

for a smooth enough function $\widehat{\psi}$ defined on $\mathbb{R}^d$, and one observes that

$$\Big(\frac{\partial\widehat{\psi}}{\partial\varphi} - \varepsilon\,\varphi_{xx}, \varphi_x\Big) = \Big(\widehat{\psi}(\varphi) - \varepsilon\,\frac{|\varphi_x|^2}{2}\Big)_x,$$

so that, since φ_x is assumed to be continuous, in any open interval where $\varphi_x \neq 0$, one obtains a stationary solution of (11) by solving

$$\widehat{\psi}(\varphi) - \varepsilon\,\frac{|\varphi_x|^2}{2} = constant.$$

Theorem 1: Assume that $\widehat{\psi}$ is a function of class C^1 in $\mathbb{R}^d$, that $a_-, a_+ \in \mathbb{R}^d$ are given such that

$$\widehat{\psi}(a_-) = \widehat{\psi}(a_+),\ \text{denoted}\ \widehat{\psi}(a_\pm), \tag{12}$$

[3] In one dimension, in order to have v, w continuous and $v\,w_x$ defined, it is not necessary that w be absolutely continuous (i.e. $w \in W^{1,1}_{loc}(\mathbb{R})$): for example, $H^s(\mathbb{R})$ is a (Banach) algebra if $\frac{1}{2} < s < 1$, so that by interpolation any $v \in H^s(\mathbb{R})$ is a multiplier for $H^t(\mathbb{R})$ for $0 \leq t \leq s$, hence for $v, w \in H^s(\mathbb{R})$ one has $v\,w_x \in H^{s-1}(\mathbb{R})$, which may be multiplied by elements from $H^s(\mathbb{R})$.

and that one can join a_- to a_+ by a curve $M(s)$ of class C^2, parametrized by arc-length s on an interval $[s_-, s_+]$, and such that

$$\begin{aligned} &\widehat{\psi}\big(M(s)\big) > \widehat{\psi}(a_\pm) \text{ for } s \in (s_-, s_+), \\ &\Big(\frac{\partial \widehat{\psi}}{\partial \varphi}\big(M(s)\big), M'(s)\Big) \neq 0 \text{ for } s = s_- \text{ and } s = s_+, \end{aligned} \tag{13}$$

and

$$2L \geq \ell(s_-, s_+) = \sqrt{\frac{\varepsilon}{2}} \int_{s_-}^{s_+} \big[\widehat{\psi}\big(M(s)\big) - \widehat{\psi}(a_\pm)\big]^{-1/2} ds. \tag{14}$$

Then, there exists a stationary solution φ of (11) of the form $\varphi(x) = M\big(s(x)\big), x \in \Omega = (-L, +L)$, with $s(\pm L) = s_\pm$, $s_x(\pm L) = 0$, and $s_{xx} \in L^\infty(\Omega)$, so that $\varphi(\pm L) = a_\pm$, $\varphi_x(\pm L) = 0$, and $\varphi_{xx} \in L^\infty(\Omega; \mathbb{R}^d)$.
Proof: The first part of (13) puts a geometric restriction, that a_- and a_+ must be on the boundary of the same connected component of $\{a \in \mathbb{R}^d \mid \widehat{\psi}(a) > \widehat{\psi}(a_\pm)\}$. The second part of (13) means that $\frac{\partial \widehat{\psi}}{\partial \varphi}\big(M(s_\pm)\big) \neq 0$ and that $M'(s_\pm)$ is not tangent to the equipotential $\widehat{\psi}(a) = \widehat{\psi}(a_\pm)$ at $a_\pm$, and the equipotential is a C^1 hyper-surface near a_- and a_+ by the implicit function theorem, so that $\widehat{\psi}\big(M(s)\big) - \widehat{\psi}(a_\pm)$ behaves as $|s - s_\pm|$ near $s_\pm$, which makes the integral in (14) converge. One "defines" $s(x)$ by

$$s_x = \sqrt{\frac{2}{\varepsilon}} \big[\widehat{\psi}\big(M(s)\big) - \widehat{\psi}(a_\pm)\big]^{1/2}, \text{ and } s(-L) = s_-, \tag{15}$$

but since it is a classical case of non-uniqueness (and we do not want the constant solution s_-), one means the solution corresponding to

$$x = -L + \sqrt{\frac{\varepsilon}{2}} \int_{s_-}^{s} \big[\widehat{\psi}\big(M(\sigma)\big) - \widehat{\psi}(a_\pm)\big]^{-1/2} d\sigma, \text{ for } s_- < s < s_+ \tag{16}$$

and condition (14) asserts that s reaches s_+ for a value $x_* \in (-L, +L]$, and one defines

$$\varphi(x) = M\big(s(x)\big), \text{ with } s(x) = s_- \text{ for } x < -L \text{ and } s(x) = s_+ \text{ for } x > x_*. \tag{17}$$

Since (15), which follows from (16), implies that s is of class C^2, one deduces by deriving (15) that

$$s_{xx} = \sqrt{\frac{2}{\varepsilon}} \frac{1}{2} \big[\widehat{\psi}\big(M(s)\big) - \widehat{\psi}(a_\pm)\big]^{-1/2} \Big(\frac{\partial \widehat{\psi}}{\partial \varphi}\big(M(s)\big), M'(s)\Big) s_x = \frac{1}{\varepsilon} \Big(\frac{\partial \widehat{\psi}}{\partial \varphi}\big(M(s)\big), M'(s)\Big), \tag{18}$$

which is then bounded for $x \in (-L, x_*)$. One deduces from (17) that

$$\varphi_x = M'(s)\, s_x, \text{ and } \varphi_{xx} = M'(s)\, s_{xx} + M''(s)\, s_x^2, \tag{19}$$

so that, because $|M'(s)| = 1$, which implies $\big(M''(s), M'(s)\big) = 0$, one has

$$\varepsilon\, (\varphi_{xx}, \varphi_x) = \varepsilon\, s_{xx} s_x = \Big(\frac{\partial \widehat{\psi}}{\partial \varphi}, M'(s)\Big) s_x = \Big(\frac{\partial \widehat{\psi}}{\partial \varphi}, \varphi_x\Big), \tag{20}$$

which implies that one has a stationary solution of (11).

Remark 2: It is enough that the curve $M(s)$ be of class C^1 with an integrable curvature, between s_- and s_+, more precisely

$$\int_{s_-}^{s_+} |M''(s)|\, ds < \infty \text{ implies } \varphi_{xx} \in L^1(\Omega; \mathbb{R}^d). \tag{21}$$

Indeed, s_x and s_{xx} are bounded by (15) and (18), so that (19) implies

$$\int_{-L}^{+L} |\varphi_{xx}|\, dx \leq \int_{-L}^{+L} |s_{xx}|\, dx + \int_{-L}^{+L} |M''(s)|\, s_x^2\, dx \leq 2L\, |s_{xx}|_\infty + |s_x|_\infty \int_{s_-}^{s_+} |M''(s)|\, ds.$$

Remark 3: Assuming that a curve M is chosen satisfying (13), i.e. one only considers the case $\varphi(x) = M\big(s(x)\big)$ for a function s, and one imposes $s(\pm L) = s_\pm$ (which is stronger than the condition $\varphi(\pm L) = a_\pm$ in the case $a_+ = a_-$), what can be said about uniqueness?

There is uniqueness only if $2L = \ell(s_-, s_+)$: there must exist a value $x_0 \in (-L, +L)$ with $s_- < s(x_0) < s_+$, and uniqueness for the differential equation (15) holds around $s(x_0)$ as long as $s_- < s(x) < s_+$ since $\big[\widehat{\psi}\big(M(s)\big) - \widehat{\psi}(a_\pm)\big]^{1/2}$ is C^1 on (s_-, s_+), and this solution tends to s_- for a value $x_- < x_0$ and tends to s_+ for a value $x_+ > x_0$, and $x_+ - x_- = \ell(s_-, s_+)$; then, there is no choice other than having $x_- = -L$ and $x_+ = +L$.

A first form of non-uniqueness holds if $2L > \ell(s_-, s_+)$, in that one may translate the solution in x of any amount from 0 to $2L - \ell(s_-, s_+)$.

A second form of non-uniqueness holds if $2L \geq 3\ell(s_-, s_+)$: one observes that our equation is invariant by changing x into $-x$, i.e. choosing the opposite sign in (15), so that going from s_+ to s_- gives a solution connecting a_+ to a_-, hence one obtains a solution by gluing together the constructed piece going from s_- to s_+, the reversed piece going from s_+ to s_-, and again the first piece going from s_- to s_+. If $2L > 3\ell(s_-, s_+)$, there is even room for waiting for a while at s_- or at s_+, of amounts adding up to $2L - 3\ell(s_-, s_+)$.

Of course, another family of solutions is possible when $2L \geq 5\ell(s_-, s_+)$, $2L \geq 7\ell(s_-, s_+)$, and so on.

Remark 4: Once a_- and a_+ can be connected by a curve M, there are plenty of curves doing the job, so that the coefficient of $\sqrt{\varepsilon}$ in $\ell(s_-, s_+)$, defined at (14), is not a definite number depending only upon a_-, a_+, and $\widehat{\psi}$, but may take all values from a smallest positive one to $+\infty$. For a_+ near a_-, one estimates the order of this smallest value by replacing $\widehat{\psi}$ by its linearization at a_-, and besides a coefficient $\big|\frac{\partial\widehat{\psi}}{\partial\varphi}(a_-)\big|^{-1/2}$ one finds a universal constant times $|a_+ - a_-|^{1/2}$ by a scaling argument.[4]

Remark 5: It may be useful to point out that in the case of linear elasticity, i.e. $\widehat{\psi}(\varphi) = (K\,\varphi, \varphi)$ for a matrix K, the existence result of a stationary solution connecting different values that we obtained holds for a system but has no analogue for a scalar equation: indeed, if $d = 1$ and $\widehat{\psi}(\varphi) = \kappa\,\varphi^2$ with $\kappa > 0$, the region $\widehat{\psi} > c$ is disconnected if $c > 0$, and for $c = 0$ one cannot take $a_- = a_+ = 0$ since one has $\frac{\partial\widehat{\psi}}{\partial\varphi}(0) = 0$, hence the second part of (13) does not hold.

4. **Limits of stationary solutions as ε tends to 0.**

When ε tends to 0, the solutions constructed are rescaled versions of the solution corresponding to $\varepsilon = 1$, and they converge to the discontinuous function taking the value a_- for $x < -L$ and the value a_+ for $x > -L$. Of course, there is then no way to deduce from the sole knowledge of a_- and a_+ what *internal structure* the discontinuity shows. Also, besides the constraint (12), which says that one must have $\widehat{\psi}(a_-) = \widehat{\psi}(a_+)$, (13) requires more, i.e. the existence of the curve M, which implies that a_- and a_+ are on the boundary of the same connected component of $\{a \in \mathbb{R}^d \mid \widehat{\psi}(a) > \widehat{\psi}(a_\pm)\}$.

One should be careful in comparing our result to what happens for hyperbolic systems of conservation laws [Da], where there are various types of E-conditions for selecting which discontinuities one wants to accept, and one usually does not accept both the jump from a_- to a_+ and the jump from a_+ to a_-, except in the case of contact discontinuities, because our system (11) with $\varepsilon = 0$ is not an hyperbolic system of conservation laws.

However, one may perform some manipulations for transforming (11) into a regularized version of an hyperbolic system of conservation laws by using the variables

$$\chi = \Big(\frac{\partial\widehat{\psi}}{\partial\varphi}, \varphi_x\Big) = \big(\widehat{\psi}(\varphi)\big)_x, \quad \Phi = \varphi_x, \tag{22}$$

[4] If one stays near a_-, the equipotential $\widehat{\psi}(M) = \widehat{\psi}(a_-)$ looks like the tangent hyperplane at a_- and $\widehat{\psi}(M) - \widehat{\psi}(a_-)$ looks like $\big|\frac{\partial\widehat{\psi}}{\partial\varphi}(a_-)\big|$ times the signed distance to the hyperplane, so that the crucial question is like for $d = 2$ if one considers a curve $y = g(x)$ with $g(x_\pm) = 0$ and $0 < g(x)$ for $x_- < x < x_+$ and one looks for g minimizing $\int_{x_-}^{x_+} \frac{\sqrt{1+(g')^2}}{\sqrt{g}}\,dx$, and rescaling means taking $x = x_- + \xi\,(x_+ - x_-)$ and $g(x) = (x_+ - x_-)\,\gamma(\xi)$, which makes a coefficient $\sqrt{x_+ - x_-}$ appear, time the minimum value corresponding to the interval $(0, 1)$.

so that (11) becomes

$$\varphi_t + \big(\chi - \varepsilon\,(\Phi_x, \Phi)\big)\,\Phi = 0, \tag{23}$$

and taking the derivative of (23) in x gives the first equation of (24), while taking the scalar product of (23) with $\frac{\partial\widehat{\psi}}{\partial\varphi}$ and then taking the derivative in x gives the second equation of (24):

$$\begin{aligned} &\Phi_t + \big[\big(\chi - \varepsilon\,(\Phi_x, \Phi)\big)\,\Phi\big]_x = 0,\\ &\chi_t + \big[\big(\chi - \varepsilon\,(\Phi_x, \Phi)\big)\,\chi\big]_x = 0, \end{aligned} \tag{24}$$

which is an unconventional regularization of the case $\varepsilon = 0$, i.e. (25), which is a system of $d+1$ equations in conservative form, the last one being Burgers's equation with a different coefficient than usual

$$\begin{aligned} &\Phi_t + (\chi\,\Phi)_x = 0,\\ &\chi_t + (\chi^2)_x = 0. \end{aligned} \tag{25}$$

Since the precise function $\widehat{\psi}$ used does not appear in (24), one must show how a solution of (24) permits to define φ satisfying (11), and for doing this one assumes that the initial data for (24) and (11) are linked by the relations

$$\Phi_0 = \varphi_{0x}; \quad \chi_0 = \Big(\frac{\partial\widehat{\psi}}{\partial\varphi}(\varphi_0), \varphi_{0x}\Big) = \Big(\frac{\partial\widehat{\psi}}{\partial\varphi}(\varphi_0), \Phi_0\Big). \tag{26}$$

One defines $a_\varepsilon(x,t) = \chi - \varepsilon\,(\Phi_x, \Phi)$, assuming that the solution of (24) is smooth enough, so that the characteristic curves defined for any basis point $y \in \mathbb{R}$ by

$$\frac{d}{dt}x(t;y) = a_\varepsilon(x(t;y),t), \text{ with } x(0;y) = y,$$

are well defined for $t \in [0,T]$, and that uniqueness holds for equations of the type (27) and (29), so that for any initial data v_0 which is smooth enough, the solution of

$$v_t + a_\varepsilon v_x = 0 \text{ in } \mathbb{R}\times(0,T), \text{ with } v\,|_{t=0} = v_0 \tag{27}$$

is given by

$$v(x(t;y),t) = v_0(y) \text{ for } y \in \mathbb{R}, \tag{28}$$

and the solution of

$$w_t + (a_\varepsilon w)_x = 0 \text{ in } \mathbb{R}\times(0,T), \text{ with } w\,|_{t=0} = v_{0x} \tag{29}$$

is given by

$$w(x,t) = v_x(x,t) \text{ in } \mathbb{R}\times(0,T), \tag{30}$$

since v_x is a solution of (29).

One then defines φ not as a solution of (23) but as the solution of

$$\varphi_t + a_\varepsilon\,\varphi_x = 0 \text{ in } \mathbb{R}\times(0,\infty), \text{ with } \varphi\,|_{t=0} = \varphi_0, \tag{31}$$

and one applies the preceding remark to v_0 being a component of $U_0 = \varphi_{0x}$, and one deduces from the first equation in (24) and (31) that $\Phi = \varphi_x$. Then, one observes that (31) implies

$$\big(g(\varphi)\big)_t + a_\varepsilon\,\big(g(\varphi)\big)_x = 0 \text{ in } \mathbb{R}\times(0,\infty), \tag{32}$$

for every smooth function g, and choosing $g = \widehat{\psi}$, and $v_0 = \big(\widehat{\psi}(\varphi_0)\big)_x$, one deduces from the second equation in (24) and (32) that $\chi = \big(\widehat{\psi}(\varphi)\big)_x$, so that $a_\varepsilon(x,t) = \chi - \varepsilon\,(\Phi_x, \Phi) = \big(\frac{\partial\widehat{\psi}}{\partial\varphi} - \varepsilon\,\varphi_{xx}, \varphi_x\big)$, hence the solution of (31) satisfies (11).

5. **The system of conservation laws obtained for $\varepsilon = 0$.**

In order to check if (25) is an hyperbolic system, one rewrites it for smooth solutions as

$$\begin{pmatrix}\Phi\\ \chi\end{pmatrix}_t + A(\Phi,\chi)\begin{pmatrix}\Phi\\ \chi\end{pmatrix}_x = 0, \text{ with } A(\Phi,\chi) = \begin{pmatrix}\chi\, I & \Phi\\ 0 & 2\chi\end{pmatrix}, \tag{33}$$

so that A is hyperbolic (i.e. diagonalizable with real eigenvalues) if $\chi \neq 0$, but also at the origin, and the general theory of Peter LAX (for the Riemann problem) applies for the upper half space $\chi > 0$ or the lower half space $\chi < 0$, but mixing the two creates some difficulties because for $\chi = 0$ and $\Phi \neq 0$, the matrix $A(\Phi,\chi)$ still has real eigenvalues but is non-diagonalizable.

If $\chi \neq 0$, then $A(\Phi,\chi)$ has an eigenvalue χ of multiplicity d, with eigen-space $\mathbb{R}^d \times \{0\}$, and this field is linearly degenerate, i.e. the derivative of the eigenvalue in the direction of an eigenvector is everywhere 0 (since the eigenvalue χ is independent of Φ). Also, $A(\Phi,\chi)$ has a simple eigenvalue 2χ, with eigenvector $\begin{pmatrix}\Phi\\ \chi\end{pmatrix}$, and this field is genuinely nonlinear, i.e. the derivative of the eigenvalue in the direction of an eigenvector is everywhere $\neq 0$ (since it is $+2$).

The linear degeneracy of system (25) leads to accept contact discontinuities, because for a constant χ, say $\chi = v$, and Φ_0 smooth, then $\Phi(x,t) = \Phi_0(x - v\,t)$ and $\chi(x,t) = v$ is a smooth solution of (25), and one is led to accept limits of such solutions corresponding to Φ_0 not necessarily smooth. However, the initial data should be of the form (26), so that these Riemann data for (25) are usually not "physical", i.e. they do not correspond to a problem for φ.

Even for the case $\chi = 0$, if one imposes condition (26) to initial data for (25), then φ_0 must be such that $\widehat{\psi}(\varphi_0)$ is constant, so that if one takes a sequence of smooth initial data φ_0 equal to a_- for $x < 0$, and satisfying $\widehat{\psi}(\varphi) = \widehat{\psi}(a_-)$ for $x > 0$, one can only expect to obtain in the limit a function φ jumping from a_- to a point a_+ on the same connected component of the equipotential $\widehat{\psi}(\varphi) = \widehat{\psi}(a_-)$, while we have constructed such limiting solutions under the weaker condition that a_+ and a_- are on the boundary of the same component of $\widehat{\psi}(\varphi) > \widehat{\psi}(a_-)$.[5]

It seems then questionable to draw conclusions for the behaviour of our equation (11) when ε tends to 0 from information corresponding to the solution of the Riemann problem for (25).

6. **The evolution problem for $\varepsilon > 0$.**

We now consider the evolution equation (11) using the same approach as for stationary solutions, in that one seeks φ taking its values on a smooth curve, parametrized using arc-length s, i.e.

$$\varphi(x,t) = M\big(s(x,t)\big), s_- \leq s(x,t) \leq s_+, x \in \Omega = (-L,+L), t \in [0,T], \tag{34}$$

but we now impose Neumann conditions

$$\varphi_x(\pm L, t) = 0, t \in (0,T), \tag{35}$$

so that extending φ by $\varphi(-L,t)$ on $(-\infty,-L)$ and $\varphi(+L,t)$ on $(+L,+\infty)$ does not usually give a solution of (11) for $x \in \mathbb{R}$ (since the "constants" in $(-\infty,-L)$ and $(+L,+\infty)$ may depend upon t). From (34), one deduces that $\varphi_t = s_t M'(s)$ and $\varphi_x = s_x M'(s)$, and since $|M'(s)| = 1$, one deduces that $|\varphi_x|^2 = s_x^2$, which implies $(\varphi_{xx},\varphi_x) = s_{xx}s_x$,[6] and the equation (11) complemented with (34)-(35) consists in seeking a function s satisfying

$$s_t + F(s)\, s_x^2 - \varepsilon\, s_x^2 s_{xx} = 0, x \in \Omega = (-L,+L), t \in (0,T), s_x(\pm L,t) = 0, t \in (0,T), s\big|_{t=0} = s_0, x \in \Omega, \tag{36}$$

[5] If $d \geq 2$ and $\widehat{\psi}(\varphi) = |\varphi|^2(|\varphi|^2 - 1)^2$, then if $0 < c < \frac{4}{27}$ the equation $r^2(r^2-1)^2 = c$ has three positive roots $0 < r_1 < \frac{1}{\sqrt{3}} < r_2 < 1 < r_3$, and the region $\widehat{\psi}(\varphi) > c$ has two connected components, one being $r_1 < |\varphi| < r_2$ and the other being $|\varphi| > r_3$, and the boundary of the first region has two connected components, one being $|\varphi| = r_1$ and the other being $|\varphi| = r_2$.

[6] One has $\varphi_x = s_x M'(s)$ and $\varphi_{xx} = s_{xx}M'(s) + s_x^2 M''(s)$, so that imposing M of class C^2 is natural (but too strong a condition) for talking about φ_{xx}; however, since M'' is perpendicular to M' (because M' stays a unit vector) the term in s_x^2 does not appear in the result for (φ_{xx},φ_x).

where one uses

$$E(s) = \widehat{\psi}\big(M(s)\big); \quad F(s) = E'(s) = \Big(\frac{\partial \widehat{\psi}}{\partial \varphi}\big(M(s)\big), M'(s)\Big), s_- < s < s_+. \tag{37}$$

Theorem 6: Assume that $\widehat{\psi}$ is a function of class C^1 in $\mathbb{R}^d$, that a curve $M(s)$ of class C^2 is parametrized by arc-length s on an interval $[s_-, s_+]$, and that

$$\varphi_0(x) = M\big(s_0(x)\big), x \in \Omega, \text{ with } s_0 \in W^{1,4}(\Omega), s_- \le s_0(x) \le s_+, x \in \Omega, \tag{38}$$

then there exists a unique solution s of (36), equivalent to (11) with (34)-(35), satisfying

$$\begin{aligned} &s_- \le s(x,t) \le s_+, x \in \Omega, t \in (0,T); \quad s_x \in L^4\big((0,T) \times \Omega\big); \quad s_t \in L^2\big((0,T) \times \Omega\big) \\ &(s_x)^3 \in L^2\big((0,T); H^1(\Omega)\big). \end{aligned} \tag{39}$$

Most of the techniques used in the various steps of the proof are not new, and were already taught by Jacques-Louis LIONS in the late 1960s [Li],[7] so that they should not be attributed to recent authors, as it is unfortunately done sometimes. For some technical results on Sobolev spaces, one may consult the course of Shmuel AGMON [Ag], or [Ta4].

One extends F outside the interval $[s_-, s_+]$ so that it is bounded and (globally) Lipschitz continuous, although the solution does not depend upon which extension one takes, but one will show the existence of a solution by a method of approximation which does not rely on the maximum principle, so that one needs the equation to be defined for $s \in \mathbb{R}$; one will then show that the solution is unique and satisfies some regularity properties, one of them being that $a \le s_0 \le b$ in Ω implies $a \le s \le b$ in $\Omega \times (0,T)$, so that the solution takes its values in the interval $[s_-, s_+]$.

The first step is to use a Faedo–Galerkin approximation for a simpler problem, where the non-linear term $F(s)\, s_x^2$ is replaced by a given function f, and passing to the limit requires the use of the monotonicity method, which was introduced independently by Eduardo ZARANTONELLO for a problem in continuum mechanics, and by George MINTY for a problem on electrical circuits.[8]

The second step is to use a fixed point argument for a truncated equation, in order to find f such that the solution s satisfies $F(s)\, \min\{s_x^2, h^2\} = f$ for a constant h, and for verifying the hypotheses of the Schauder–Tychonoff fixed point theorem, for a space endowed with a weak topology, one needs an argument of compactness: one uses what is sometimes called the Aubin–Lions lemma, since it seems due to Jean-Pierre AUBIN and it uses a lemma of Jacques-Louis LIONS.[9]

The third step is to let h tend to $+\infty$ in order to obtain a solution.

The fourth step is to prove some regularity properties, for deducing that the solution is unique.

The fifth step is to prove L^∞ estimates by the maximum principle, in the spirit of the work of Guido STAMPACCHIA [St].

First step: For $s_0 \in L^2(\Omega)$ and $f \in L^2\big(0,T; L^2(\Omega)\big)$, and with $\Omega = (-L, +L)$, one wants to solve the equation

$$s_t + f - \varepsilon\, s_x^2 s_{xx} = 0, x \in \Omega, t \in (0,T); \quad s\big|_{t=0} = s_0 \in L^2(\Omega); \quad s_x\big|_{x=\pm L} = 0, t \in (0,T), \tag{40}$$

where $s_x^2 s_{xx}$ means $\big(\frac{s_x^3}{3}\big)_x$, of course. Also, since the Neumann boundary conditions do not have a meaning for functions in $W^{1,4}(\Omega)$, they are imposed in a variational way; they can also be given a meaning for

[7] In particular, one does not use the generalization called the compensated compactness method, which the second author developed in the late 1970s, based on his joint work with François MURAT, but a good reference for the history of various related questions is [Ta2], written for a volume in memory of Luigi AMERIO.

[8] After them, monotonicity was given a general framework in functional analysis by the works of Haïm BREZIS, Felix BROWDER, Jacques-Louis LIONS, and Terry ROCKAFELLAR (in alphabetic order).

[9] It is a variant of the Fréchet–Kolmogorov compactness argument, and like almost all the compactness theorems in functional analysis, it uses the ideas of ARZELÁ and ASCOLI on equi-continuity.

solutions of some partial differential equations, by using the normal trace of functions in the space $H(div)$ introduced by Jacques-Louis LIONS.

Besides $H = L^2(\Omega)$, one considers $V = W^{1,4}(\Omega)$, which is $\subset C^{0,3/4}(\overline{\Omega})$. The solution s is found in the space $C([0,T];H) \cap L^4(0,T;V)$, and since $|u|_\infty \le C\,(|u_x|_4^{2/5}|u|_2^{3/5} + L^{-1/2}|u|_2)$ for all $u \in V$,[10] where $|\cdot|_p$ denotes the norm in $L^p(\Omega)$, one deduces that $|s|_\infty \in L^{10}(0,T)$.

The existence part uses the fact that s_x^3 is monotone in s_x, i.e. $(a^3-b^3)\,(a-b) \ge 0$ for all $a,b \in \mathbb{R}$, and even $\ge \frac{1}{4}\,|a-b|^4$, which will be crucial for proving some strong convergences.

One uses a Faedo–Galerkin basis of V, i.e. linearly independent functions $w_1,\ldots,w_n,\ldots \in V$, whose linear combinations are dense in V, and one approaches the initial data s_0 by a sequence s_{0n} converging to s_0 strongly in H, where s_{0n} belongs to $V_n = span\{w_1,\ldots,w_n\} \subset V$, and one solves

$$\begin{aligned}&\int_\Omega \frac{ds_n}{dt}\,w_j\,dx + \int_\Omega f\,w_j\,dx + \frac{\varepsilon}{3}\int_\Omega s_{nx}^3 w_{jx}\,dx = 0, j = 1,\ldots,n, \text{ and } t \in (0,T),\\ &\text{with } s_n = \sum_{k=1}^n c_k(t)\,w_k, \text{ and } s_n\big|_{t=0} = s_{0n}.\end{aligned} \tag{41}$$

This is a differential system for the unknown coefficients $c_1,\ldots,c_n$,[11] because the matrix with entries $\int_\Omega w_j w_k\,dx$ is invertible, and since the non-linear term is algebraic in $c_1,\ldots,c_n$, hence locally Lipschitz continuous, there is a unique solution on a maximal interval $(0,T_n)$ with $0 < T_n \le T$, and for showing that $T_n = T$ one needs bounds: multiplying the jth equation by c_j and summing in j gives

$$\frac{d}{dt}\Big(\int_\Omega \frac{s_n^2}{2}\,dx\Big) + \frac{\varepsilon}{3}\int_\Omega s_{nx}^4\,dx = -\int_\Omega f\,s_n\,dx \le |f(\cdot,t)|_2|s_n(\cdot,t)|_2, \tag{42}$$

and one then uses a variant of Gronwall's inequality,[12] to deduce that

$$|s_n(\cdot,t)|_2 \le |s_{0n}|_2 + \int_0^t |f(\cdot,\tau)|_2\,d\tau, t \in (0,T_n), \text{ so that } T_n = T, \tag{43}$$

and then

$$\frac{\varepsilon}{3}\int_0^T |s_{nx}(\cdot,t)|_4^4\,dt \le \frac{|s_{0n}|_2^2}{2} + \int_0^T |f(\cdot,t)|_2|s_n(\cdot,t)|_2\,dt \le \frac{1}{2}\Big(|s_{0n}|_2 + \int_0^T |f(\cdot,t)|_2\,dt\Big)^2. \tag{44}$$

From the nature of the bounds, one could have taken $f \in L^1\big(0,T;L^2(\Omega)\big)$, but L^2 in time is enough for our purpose, and it is better for using weak convergence arguments later. For existence of s_n, the bounds may depend upon n, but for passing to the limit it is important to have the bounds independent of n.

One extracts a subsequence s_m such that

$$\begin{aligned}&s_m \rightharpoonup s_\infty \text{ in } L^2\big((0,T)\times\Omega\big) \text{ weak},\\ &s_{mx} \rightharpoonup s_{\infty x} \text{ in } L^4\big((0,T)\times\Omega\big) \text{ weak},\\ &s_{mx}^3 \rightharpoonup \xi_\infty \text{ in } L^{4/3}\big((0,T)\times\Omega\big) \text{ weak},\\ &s_m(\cdot,T) \rightharpoonup \sigma_\infty \text{ in } L^2(\Omega) \text{ weak},\end{aligned} \tag{45}$$

[10] One may choose $|u_x|_4 + L^{-5/4}|u|_2$ as a norm on $W^{1,4}(\Omega)$, the power of L being chosen so that the two terms are measured in the same unit. Then, there is a linear extension to functions on $\mathbb{R}$ (by mirror symmetry at the ends of the interval, followed by a truncation) whose norm is bounded by a constant independent of L, for the norm of H or that of V. On $\mathbb{R}$ a scaling argument shows that $|v|_\infty \le C\,|v_x|_4^{2/5}|v|_2^{3/5}$.

[11] There should be another index to represent that they are the coefficients of s_n.

[12] One has $|s_n(\cdot,t)|_2^2 \le |s_{0n}|_2^2 + 2\int_0^t |f(\cdot,\tau)|_2|s_n(\cdot,\tau)|_2\,d\tau = g_n(t)$, so that $g_n'(t) = 2|f(\cdot,t)|_2|s_n(\cdot,t)|_2 \le 2|f(\cdot,t)|_2\sqrt{g_n(t)}$, hence $(\sqrt{g_n})'(t) \le |f(\cdot,t)|_2$, which gives by integration $\sqrt{g_n}(t) \le |s_{0n}|_2 + \int_0^t |f(\cdot,\tau)|_2\,d\tau$.

although one has better information, that $s_n, s_\infty \in L^\infty\big(0,T;L^2(\Omega)\big)$. One multiplies by a smooth function of t and one integrates by parts in t, and then one passes to the limit; after using the fact that linear combinations of the w_j are dense, one finds (in a weak formulation) that

$$
\begin{aligned}
&s_{\infty t} + f - \frac{\varepsilon}{3}\,\xi_{\infty x} = 0, \text{ in } (0,T)\times\Omega,\\
&s_\infty(\cdot,0) = s_0; \quad s_\infty(\cdot,T) = \sigma_\infty; \quad \xi_\infty\,|_{\pm L} = 0 \text{ in } (0,T),
\end{aligned} \tag{46}
$$

and there are some technical details, for showing that s_∞ is continuous in t with values in $L^2(\Omega)$,[13] so that the initial data and final data make sense, and for the Neumann condition one uses an argument of Jacques-Louis LIONS concerning the space $H(div;(0,T)\times\Omega)$, but there is an important observation, which is that one has $\xi_\infty = s_{\infty x}^3$, and it is where the monotonicity property is crucial. Multiplying by s_m and integrating in (x,t), and comparing to what one obtains by multiplying the limiting equation by s_∞ and integrating in (x,t), one obtains[14]

$$
\limsup_{m\to\infty} \int_0^T \int_\Omega s_{mx}^4\,dx\,dt \le \int_0^T \int_\Omega \xi_\infty s_{\infty x}\,dx\,dt, \tag{47}
$$

and one can deduce from (47) that

$$
\begin{aligned}
&\xi_\infty = s_{\infty x}^3 \text{ a.e. in } (0,T)\times\Omega,\\
&s_{mx} \to s_{\infty x} \text{ in } L^4\big(0,T;L^4(\Omega)\big) \text{ strong.}
\end{aligned} \tag{48}
$$

Indeed, for the first part of (48), one deduces from (47) and from the definition of ξ_∞ that

$$
\limsup_{m\to\infty} \int_0^T \int_\Omega (s_{mx}^3 - v^3)\,(s_{mx} - v)\,dx\,dt \le \int_0^T \int_\Omega (\xi_\infty - v^3)\,(s_{\infty x} - v)\,dx\,dt, \text{ for all } v \in L^4\big(0,T;L^4(\Omega)\big), \tag{49}
$$

and the monotonicity implies

$$
0 \le \int_0^T \int_\Omega (\xi_\infty - v^3)\,(s_{\infty x} - v)\,dx\,dt, \text{ for all } v \in L^4\big(0,T;L^4(\Omega)\big); \tag{50}
$$

one concludes with a trick used by George MINTY, which is to take $v = s_{\infty x} + \eta\,w$ with $w \in L^4\big(0,T;L^4(\Omega)\big)$ and, after dividing by η (and paying attention to the change of inequality for $\eta < 0$) let η tend to 0 either from the positive side or the negative side, so that

$$
\int_0^T \int_\Omega (\xi_\infty - s_{\infty x}^3)\,w\,dx\,dt = 0 \text{ for all } w \in L^4\big(0,T;L^4(\Omega)\big), \text{ hence } \xi_\infty = s_{\infty x}^3 \text{ a.e. in } (0,T)\times\Omega. \tag{51}
$$

For the second part of (48), one uses $v = s_{\infty x}$ in (49), and one deduces that

$$
\limsup_{m\to\infty} \int_0^T \int_\Omega |s_{mx} - s_{\infty x}|^4\,dx\,dt \le 0, \tag{52}
$$

which gives the strong convergence of s_{mx} to $s_{\infty x}$ in $L^4\big(0,T;L^4(\Omega)\big)$, hence a way to identify $\xi_\infty = s_{\infty x}^3$ without using Minty's trick.

[13] The function space to use is $s_\infty \in L^\infty\big(0,T;L^2(\Omega)\big) \cap L^4\big(0,T;W^{1,4}(\Omega)\big)$, but the derivative $s_{\infty t}$ is the sum of two terms, one in $L^2\big(0,T;L^2(\Omega)\big)$, or more generally in $L^1(0,T;L^2(\Omega))$, and another which is the x derivative (in the sense of distributions) of a function in $L^{4/3}\big((0,T);L^{4/3}(\Omega)\big)$, and there are technical steps (see [Li] or [Ta4]) of truncation, regularization, for showing the density of smooth enough functions, and deduce that the formula of integration by parts is valid, and that one actually has $s_\infty \in C\big([0,T];L^2(\Omega)\big)$.

[14] For obtaining (47), one uses the fact that $\liminf_{m\to\infty} |s_m(\cdot,T)|_2^2 \ge |\sigma_\infty|_2^2$.

Once one has proved all the technical details about the validity of the integration by parts, the uniqueness of a solution of (40) is easy, since two solutions s and $\overline{s}$ imply after subtraction and multiplication by $s-\overline{s}$ that

$$\frac{d}{dt}\,\frac{|s-\overline{s}|_2^2}{2}+\int_\Omega (s_x^3-\overline{s}_x^3)\,(s_x-\overline{s}_x)\,dx=0, \text{ which implies } \frac{d}{dt}\,\frac{|s-\overline{s}|_2^2}{2}\le 0, \text{ hence } s=\overline{s}, \tag{53}$$

since at time 0 one has $|s-\overline{s}|_2=0$.

Second step: For the fixed point argument, one uses the version by TYCHONOFF, that if one has a (nonempty) compact convex set K of a locally convex space, and a continuous mapping S from K into K, then there exists $k\in K$ with $S(k)=k$. One uses the weak topology on a closed bounded convex set of $Z_2=L^2\big(0,T;L^2(\Omega)\big)$, which is (sequentially) compact for the weak topology (which is metrizable on bounded sets like K), and the continuity of the mapping will result from a compactness argument, often called the Aubin–Lions lemma.[15]

Using the fact that $F(s)$ is a continuous bounded function, and choosing $h>0$ (which will later tend to $+\infty$), one defines $S_h(f)$ by

$$S_h(f)=F(s)\,\min\{s_x^2,h^2\}, \text{ where } s \text{ satisfies (40).} \tag{54}$$

Since S_h sends all the space Z_2 into a bounded subset of Z_2, because of the introduction of the parameter h, one takes for K a closed ball of Z_2 which contains the image of S_h, and one needs to check the continuity of S_h from Z_2 to Z_2, i.e. show that if $f_n\rightharpoonup f_\infty$ in Z_2 weak, then $S_h(f_n)\rightharpoonup S_h(f_\infty)$ in Z_2 weak.

Since the injection of $W^{1,4}(\Omega)$ into $L^4(\Omega)$ is compact, the Aubin–Lions lemma implies (for $p>1$) that if a sequence s_n is bounded in $L^p\big(0,T;W^{1,4}(\Omega)\big)$, and the sequence s_{nt} is bounded in $L^p(0,T;E)$ for a Banach space E in which $L^4(\Omega)$ is continuously embedded, then the sequence belongs to a compact of $L^p\big(0,T;L^4(\Omega)\big)$ (for the strong topology).

If $f_n\rightharpoonup f_\infty$ in Z_2 weak, then the corresponding solutions s_n of (40) are bounded in $L^4\big(0,T;W^{1,4}(\Omega)\big)$, and the derivatives s_{nt} have a term bounded in $L^2\big(0,T;L^2(\Omega)\big)$ and a term bounded in $L^{4/3}(0,T;E)$ where E is the dual of $W^{1,4}(\Omega)$, so that Aubin–Lions lemma applies with $p=\frac{4}{3}$. A *subsequence* s_m converges strongly in $L^{4/3}\big(0,T;L^4(\Omega)\big)$, but since s_m is bounded in $L^{10}\big(0,T;L^\infty(\Omega)\big)$, one can deduce from Hölder inequality that the same subsequence converges strongly in other spaces of the form $L^q\big(0,T;L^r(\Omega)\big)$ with related values q,r, for example in $L^a\big(0,T;L^a(\Omega)\big)$ for $a<10$. However, it is important to show that *the whole sequence converges*, and this is done by showing that the limit is unique, because it satisfies (40) for f_∞: for showing that, one extracts another subsequence such that (45) holds, then the monotonicity argument applies because

$$\int_0^T\int_\Omega f_m s_m\,dx\,dt\to\int_0^T\int_\Omega f_\infty s_\infty\,dx\,dt, \tag{55}$$

since s_m converges to s_∞ strongly in $L^2\big(0,T;L^2(\Omega)\big)$. Once it is known that the whole sequence converges to s_∞, the monotonicity argument implies that s_{nx} converges to $s_{\infty x}$ in $L^4\big(0,T;L^4(\Omega)\big)$ strong, so that $\min\big\{s_{nx}^2,h^2\big\}$ converges to $\min\big\{s_{\infty x}^2,h^2\big\}$ in $L^4\big(0,T;L^4(\Omega)\big)$ strong, because $\min\{z^2,h^2\}$ is a bounded Lipschitz continuous function of $z\in\mathbb{R}$. Then, using the fact that F is continuous and bounded, $F(s_n)$ converges to $F(s_\infty)$ in $L^\infty\big(0,T;L^\infty(\Omega)\big)$ weak $\star$ and a subsequence converges almost everywhere, hence in $L^b\big(0,T;L^b(\Omega)\big)$ strong for any $b<\infty$ by Lebesgue's dominated convergence theorem. This shows the continuity of S_h for the weak topology, hence the existence of a fixed point, which solves

$$s_t+F(s)\min\{s_x^2,h^2\}-\varepsilon\,s_x^2 s_{xx}=0, x\in\Omega, t\in(0,T), s\big|_{t=0}=s_0\in L^2(\Omega), s_x\big|_{x=\pm L}=0. \tag{56}$$

Third step: One wants to let h tend to ∞, and one needs to obtain bounds independent of h. Multiplying by s, and denoting by $||F||_\infty$ the norm of F in $L^\infty\big(0,T;L^\infty(\Omega)\big)$, one has

$$\frac{d}{dt}\Big(\int_\Omega\frac{s^2}{2}\,dx\Big)+\frac{\varepsilon}{3}\int_\Omega s_x^4\,dx=-\int_\Omega F(s)\,s_x^2\,s\,dx\le||F||_\infty|s_x|_4^2|s|_2\le\frac{\eta}{2}\,|s_x|_4^4+\frac{||F||_\infty^2}{2\eta}\,|s|_2^2 \text{ for } \eta>0, \tag{57}$$

[15] Jacques-Louis LIONS attributed this variant to Jean-Pierre AUBIN.

and choosing $\eta < \frac{2\varepsilon}{3}$ one applies Gronwall's inequality for deducing that

$$|s(\cdot,t)|_2 \le |s_0|_2 e^{C_1 t}, t \in [0,T], \int_0^T |s_x|_4^4 \, dt \le C_2, \tag{58}$$

with $C_1 = \frac{||F||_\infty^2}{2\eta}$, and these bounds are valid for all values of h. For a sequence h_n tending to $+\infty$, one has a sequence $f_n = F(s_n) \min\{s_{nx}^2, h_n^2\}$ which is bounded in $L^2\big(0,T;L^2(\Omega)\big)$ by (58), so that a subsequence f_m converges weakly to f_∞ in $L^2\big(0,T;L^2(\Omega)\big)$, and by the preceding analysis s_m converges to s_∞ strongly in $L^2\big(0,T;L^2(\Omega)\big)$, and s_{mx} converges to $s_{\infty x}$ strongly in $L^4\big(0,T;L^4(\Omega)\big)$, so that s_∞ is the solution corresponding to f_∞.

Then, for $h > 0$, let T_h be the Lipschitz continuous function equal to $-h$ for $u \in (-\infty,-h]$, equal to u for $u \in [-h,+h]$ and equal to $+h$ for $u \in [+h,+\infty)$, i.e. $u \mapsto T_h(u)$ is the truncation operator at $\pm h$, so that $\min\{s_{mx}^2, h_m^2\} = T_{h_m}(s_{mx}^2)$; one has $|T_{h_m}(s_{mx}) - s_{\infty x}|_4 \le |T_{h_m}(s_{mx}) - T_{h_m}(s_{\infty x})|_4 + |T_{h_m}(s_{\infty x}) - s_{\infty x}|_4$, which is $\le |s_{mx} - s_{\infty x}|_4 + |T_{h_m}(s_{\infty x}) - s_{\infty x}|_4$ since T_{h_m} is a contraction. Because $T_{h_m}(v)$ converges to v strongly in $L^4\big(0,T;L^4(\Omega)\big)$ for any $v \in L^4\big(0,T;L^4(\Omega)\big)$ by Lebesgue's dominated convergence theorem, one deduces that $T_{h_m}(s_{mx})$ converges to $s_{\infty x}$ strongly in $L^4\big(0,T;L^4(\Omega)\big)$, hence $\min\{s_{mx}^2, h_m^2\}$ converges to $s_{\infty x}^2$ strongly in $L^2\big(0,T;L^2(\Omega)\big)$.

Then, since s_m converges to s_∞ strongly in $L^2\big(0,T;L^2(\Omega)\big)$, one deduces that $F(s_m)$ converges to $F(s_\infty)$ strongly in $L^2\big(0,T;L^2(\Omega)\big)$ because F is continuous and bounded, so that $F(s_m)$ converges to $F(s_\infty)$ weakly $\star$ in $L^\infty\big(0,T;L^\infty(\Omega)\big)$. Hence $F(s_m)\min\{s_{mx}^2, h_m^2\}$ converges to $F(s_\infty)\, s_{\infty x}^2$ weakly in $L^2\big(0,T;L^2(\Omega)\big)$, and s_∞ satisfies (36).

Fourth step: The natural method for proving uniqueness of a solution is to write the equation for a solution s and for a solution $\overline{s}$, to subtract the two equations and to multiply by $s - \overline{s}$, and the problem is then to bound the integral

$$\int_{-L}^{+L} (F(s)\, s_x^2 - F(\overline{s})\, \overline{s}_x^2)\,(s-\overline{s})\,dx = \int_{-L}^{+L} F(s)\,(s_x^2 - \overline{s}_x^2)\,(s-\overline{s})\,dx + \int_{-L}^{+L} \big(F(s) - F(\overline{s})\big)\,\overline{s}_x^2\,(s-\overline{s})\,dx, \tag{59}$$

and the first term in the right hand side of (59) poses no difficulty, since

$$\Big|\int_{-L}^{+L} F(s)\,(s_x^2 - \overline{s}_x^2)\,(s-\overline{s})\,dx\Big| \le ||F||_\infty |s_x^2 - \overline{s}_x^2|_2 |s-\overline{s}|_2 \le \eta\,|s_x^2 - \overline{s}_x^2|_2^2 + \frac{||F||_\infty^2}{4\eta}\,|s-\overline{s}|_2^2 \tag{60}$$

and with $\eta > 0$ small the first term on the right hand side is controlled, because

$$\frac{\varepsilon}{3}\int_{-L}^{+L} (s_x^3 - \overline{s}_x^3)\,(s_x - \overline{s}_x)\,dx \ge \frac{\varepsilon}{4}\,|s_x^2 - \overline{s}_x^2|_2^2, \tag{61}$$

consequence of the inequality $(a^3-b^3)\,(a-b) \ge \frac{3}{4}\,|a^2-b^2|^2$ for all $a,b \in \mathbb{R}$. The second term in the right hand side of (59) presents a difficulty: F has been extended to be Lipschitz continuous, i.e. there is a Lipschitz constant L_F such that $|F(b) - F(a)| \le L_F\,|b-a|$ for all $a,b \in \mathbb{R}$, and this implies the pointwise estimate

$$\big|\big(F(s) - F(\overline{s})\big)\,\overline{s}_x^2\,(s-\overline{s})\big| \le L_F \overline{s}_x^2\,|s-\overline{s}|^2, \tag{62}$$

from which one deduces

$$\int_{-L}^{+L} \big|\big(F(s) - F(\overline{s})\big)\,\overline{s}_x^2\,(s-\overline{s})\big|\,dx \le L_F |s_x|_\infty^2 |s-\overline{s}|_2^2, \tag{63}$$

and choosing $\eta = \frac{\varepsilon}{4}$ one obtains

$$\frac{1}{2}\,(|s-\overline{s}|_2^2)_t \le \Big(\frac{||F||_\infty^2}{\varepsilon} + L_F |s_x|_\infty^2\Big)\,|s-\overline{s}|_2^2 = \lambda(t)\,|s-\overline{s}|_2^2, \tag{64}$$

and Gronwall inequality applies if $\lambda \in L^1(0,T)$, and in this case it implies $|s-\overline{s}|_2^2 = 0$ since s and $\overline{s}$ have the same initial data s_0; it means that uniqueness holds if $\overline{s}_x \in L^2\big(0,T;L^\infty(\Omega)\big)$, or if $s_x \in L^2\big(0,T;L^\infty(\Omega)\big)$, since the role of s and $\overline{s}$ are equivalent.

Using the Faedo–Galerkin approximation (41), the estimate shown before was that corresponding to $f \in L^1\big(0,T;L^2(\Omega)\big)$, although one only used it for $f \in L^2\big(0,T;L^2(\Omega)\big)$. However, there is a better smoothness estimate in this case, enough for proving uniqueness, and actually the following analysis only uses the hypothesis $\sqrt{t}\,f \in L^2\big(0,T;L^2(\Omega)\big)$.[16] The estimate is obtained by multiplying the equation by $t\,s_{nt}$ (i.e. by multiplying the jth equation with w_j by $t\,c_{jt}$ and summing in j), which gives

$$\int_\Omega t\,|s_{nt}|^2\,dx + \int_\Omega t\,f\,s_{nt}\,dx + \frac{\varepsilon}{3}\int_\Omega t\,s_{nx}^3 s_{nxt}\,dx = 0, \tag{65}$$

and since $|t\,f\,s_{nt}| \le \frac{1}{2}\,t\,|s_{nt}|^2 + \frac{1}{2}\,t\,|f|^2$, one deduces that

$$\frac{\varepsilon}{12}\Big(\int_\Omega t\,s_{nx}^4\,dx\Big)_t + \frac{1}{2}\int_\Omega t\,|s_{nt}|^2\,dx \le \frac{1}{2}\int_\Omega t\,|f|^2\,dx + \frac{\varepsilon}{12}\int_\Omega |s_{nx}|^4\,dx, \tag{66}$$

and since there is a bound independent of n for $\int_\Omega |s_{nx}|^4\,dx$, one deduces bounds independent of n for $\int_\Omega t\,s_{nx}^4\,dx$ in $L^\infty(0,T)$, and for $\int_\Omega t\,|s_{nt}|^2\,dx$ in $L^1(0,T)$, and these bounds are inherited by the solution which is the limit of a subsequence s_m, so that

$$\begin{aligned}&s_0 \in L^2(\Omega), f \in L^2\big(0,T;L^2(\Omega)\big) \text{ imply that the solution of (40) satisfies}\\ &t^{1/4}s_x \in L^\infty\big(0,T;L^4(\Omega)\big), t^{1/2}s_t \in L^2\big(0,T;L^2(\Omega)\big),\end{aligned} \tag{67}$$

hence by the equation (40) it satisfies

$$t^{1/2}(s_x^3)_x \in L^2\big(0,T;L^2(\Omega)\big), \text{ i.e. } t^{1/2}s_x^3 \in L^2\big(0,T;H_0^1(\Omega)\big), \tag{68}$$

since s_x^3 satisfies a Dirichlet condition, as a consequence of the Neumann condition for s. Then, $\sigma = s_x^3$ satisfies $t^{1/2}\sigma \in L^2\big(0,T;H_0^1(\Omega)\big)$, and since $t^{1/4}s_x \in L^\infty\big(0,T;L^4(\Omega)\big)$ is equivalent to $t^{3/4}\sigma \in L^\infty\big(0,T;L^{4/3}(\Omega)\big)$, one deduces an estimate of σ in $L^\infty(\Omega)$ by using an inequality of the type $|\sigma|_\infty \le C\,|\sigma_x|_2^{3/5}|\sigma|_{4/3}^{2/5}$ for all $\sigma \in H_0^1(\Omega)$;[17] then $|\sigma_x|_2 = t^{-1/2}\lambda(t)$ for $\lambda \in L^2(0,T)$ and $|\sigma|_{4/3} \le C\,t^{-3/4}$ give $|\sigma|_\infty = t^{-3/5}\mu(t)$ with $\mu \in L^{10/3}(0,T)$, and since $t^{-3/5}\mu \in L^p(0,T)$ for $1 \le p < \frac{10}{9}$, one has $s_x \in L^q\big(0,T;L^\infty(\Omega)\big)$ for $1 \le q < \frac{10}{3}$, and because $q = 2$ is the condition needed for uniqueness, one finds that uniqueness holds for $s_0 \in L^2(\Omega)$. As a consequence, all the sequence s_n converges to the (unique) solution.

Fifth step: We shall need another regularity result, still for $f \in L^2\big(0,T;L^2(\Omega)\big)$ but with $s_0 \in W^{1,4}(\Omega)$, and the estimates result from multiplying the equation by s_t: more precisely, one works on the Faedo–Galerkin approximation (41), choosing the initial data s_{0n} converging strongly to s_0 in $W^{1,4}(\Omega)$, and one multiplies by s_{nt}, so that instead of (65) one obtains

$$\int_\Omega |s_{nt}|^2\,dx + \int_\Omega f\,s_{nt}\,dx + \frac{\varepsilon}{3}\int_\Omega s_{nx}^3 s_{nxt}\,dx = 0, \tag{69}$$

and instead of (66) one obtains

$$\frac{\varepsilon}{12}\Big(\int_\Omega s_{nx}^4\,dx\Big)_t + \frac{1}{2}\int_\Omega |s_{nt}|^2\,dx \le \frac{1}{2}\int_\Omega |f|^2\,dx, \tag{70}$$

[16] The second author devised this method for linear equations (of the abstract form $u' + A\,u = f, u(0) = u_0$ with A elliptic), and used it in his thesis for interpolating regularity, but it may not be widely known.

[17] From $(\sigma^{5/3})_x = \frac{5}{3}\,\sigma^{2/3}\sigma_x$, one deduces that $|\sigma^{5/3}|_\infty \le \frac{1}{2}\,|(\sigma^{5/3})_x|_1 \le \frac{5}{6}\,|\sigma^{2/3}|_2\,|\sigma_x|_2$, hence $|\sigma|_\infty^{5/3} \le \frac{5}{6}\,|\sigma|_{4/3}^{2/3}\,|\sigma_x|_2$.

and since there is a bound independent of n for $\int_\Omega |s_{0nx}|^4\,dx$, one deduces bounds independent of n for $\int_\Omega s_{nx}^4\,dx$ in $L^\infty(0,T)$, and for $\int_\Omega |s_{nt}|^2\,dx$ in $L^1(0,T)$, and these bounds are inherited by the solution which is the limit of a subsequence s_m, hence

$$\begin{aligned}&s_0 \in W^{1,4}(\Omega), f \in L^2\big(0,T;L^2(\Omega)\big) \text{ imply that the solution of (40) satisfies}\\ &s_x \in L^\infty\big(0,T;L^4(\Omega)\big), s_t \in L^2\big(0,T;L^2(\Omega)\big), \text{ hence } s_x^3 \in L^2\big(0,T;H_0^1(\Omega)\big), s_x \in L^{10}\big(0,T;L^\infty(\Omega)\big).\end{aligned} \tag{71}$$

This regularity enables us to apply the technique of truncation used by Guido STAMPACCHIA [St] for proving a weak form of the "maximum principle", and a consequence is that the way one has extended F outside the interval $[s_-,s_+]$ has no effect on the solution: Guido STAMPACCHIA proved that if $u \in W^{1,1}_{loc}(\mathcal{O})$ for an open set $\mathcal{O} \subset \mathbb{R}^N$, and if g is a Lipschitz continuous function on $\mathbb{R}$ with at most a countable number of discontinuities of the derivative, then $g(u) \in W^{1,1}_{loc}(\mathcal{O})$ with $\frac{\partial g(u)}{\partial x_j} = g'(u)\,\frac{\partial u}{\partial x_j}$ for $j = 1,\ldots,N$,[18] and if one applies it to $g(u) = u_+$ it does not matter what value $g'(0)$ one takes.[19]

One considers two solutions s,σ of (36) with initial data s_0,σ_0, and one wants to shows that

$$s_0 \le \sigma_0 \text{ a.e. in } \Omega \text{ implies } s \le \sigma \text{ a.e. in } \Omega\times(0,T), \tag{72}$$

and for proving this one wants to subtract the two equations and multiply by $(s-\sigma)_+$, so that one assumes that $s_0,\sigma_0 \in W^{1,4}(\Omega)$ in order to use the estimates (71) for both s and σ, which imply $s,\sigma \in H^1\big(\Omega\times(0,T)\big)$, so that one may apply the formula for the partial derivatives of $(s-\sigma)_+$, and observe that $(s-\sigma)_+ \in H^1\big(\Omega\times(0,T)\big)$ since

$$\begin{aligned}&\big((s-\sigma)_+\big)_t = (s-\sigma)_t\chi_+ \text{ and } \big((s-\sigma)_+\big)_x = (s-\sigma)_x\chi_+ \text{ a.e. in } \Omega\times(0,T),\\ &\text{where } \chi_+ \text{ is the characteristic function of } E_+ = \{(x,t) \mid s-\sigma>0\},\end{aligned} \tag{73}$$

and because $\chi_+^2 = \chi_+$ one deduces that

$$(s-\sigma)_t(s-\sigma)_+ = \frac{1}{2}\big(|(s-\sigma)_+|^2\big)_t. \tag{74}$$

Moreover, since one also has $s_x,\sigma_x \in L^\infty\big(0,T;L^4(\Omega)\big) \cap L^2\big(0,T;L^\infty(\Omega)\big)$, one deduces that

$$\big((s-\sigma)_+\big)_x \in L^\infty\big(0,T;L^4(\Omega)\big) \cap L^2\big(0,T;L^\infty(\Omega)\big), \tag{75}$$

and one also has

$$|s-\sigma|\,(s-\sigma)_+ = \chi_+\,|(s-\sigma)_+|^2 = |(s-\sigma)_+|^2 \text{ a.e. in } \Omega\times(0,T). \tag{76}$$

One starts from the variational formulation

$$\int_\Omega \Big[(s-\sigma)_t w + \big(F(s)\,s_x^2 - F(\sigma)\,\sigma_x^2\big)\,w + \frac{\varepsilon}{3}\,(s_x^3-\sigma_x^3)\,w_x\Big]\,dx = 0 \text{ a.e. } t\in(0,T), \text{ for all } w \in W^{1,4}(\Omega), \tag{77}$$

from which one deduces that (79) is also true if $w \in L^\infty\big(0,T;W^{1,4}(\Omega)\big)$, and one then uses $w = (s-\sigma)_+$. One has

$$\frac{\varepsilon}{3}\int_\Omega (s_x^3-\sigma_x^3)\,\big((s-\sigma)_+\big)_x\,dx \ge \frac{\varepsilon}{4}\int_\Omega \chi_+\,(s_x^2-\sigma_x^2)^2\,dx \text{ a.e. } t\in(0,T), \tag{78}$$

[18] For g of class C^1 over $\mathbb{R}$ and u smooth, one has $\frac{\partial g(u)}{\partial x_j} = g'(u)\,\frac{\partial u}{\partial x_j}$, and approaching $u \in W^{1,1}_{loc}(\mathcal{O})$ by a sequence of smooth functions u_n converging strongly to u in $W^{1,1}(\omega)$ for an open set ω with $\overline{\omega} \subset \mathcal{O}$, one deduces that the formula is true for $u \in W^{1,1}_{loc}(\mathcal{O})$; then one takes a sequence g_n of C^1 functions, which converges uniformly on $\mathbb{R}$ to a Lipschitz function g, and such that g_n' converges *everywhere* on $\mathbb{R}$ to a function which one denotes g', and passing to the limit by using Lebesgue's dominated convergence theorem, one deduces the result.

[19] This is the way Guido STAMPACCHIA showed that on the set where $u(x) = 0$ one has $\frac{\partial u}{\partial x_j} = 0$ almost everywhere, for $j = 1,\ldots,N$.

and

$$\int_\Omega \big|F(s)\,s_x^2 - F(\sigma)\,\sigma_x^2\big|\,(s-\sigma)_+\,dx \le \int_\Omega |F(s)-F(\sigma)|\,s_x^2\,(s-\sigma)_+\,dx + \int_\Omega |F(\sigma)|\,|s_x^2-\sigma_x^2|\,(s-\sigma)_+\,dx$$
$$\le L_F||s_x||_\infty^2 \int_\Omega |(s-\sigma)_+|^2\,dx + \frac{\varepsilon}{4}\int_\Omega \chi_+\,(s_x^2-\sigma_x^2)^2\,dx + \frac{||F||_{L^\infty}^2}{\varepsilon}\int_\Omega |(s-\sigma)_+|^2\,dx \text{ a.e. } t\in(0,T), \tag{79}$$

so that

$$\frac{1}{2}\Big(\int_\Omega |(s-\sigma)_+|^2\,dx\Big)_t \le \Big(L_F||s_x||_\infty^2 + \frac{||F||_{L^\infty}^2}{\varepsilon}\Big)\int_\Omega |(s-\sigma)_+|^2\,dx \text{ a.e. } t\in(0,T), \tag{80}$$

and by Gronwall's inequality

$$\int_\Omega |(s-\sigma)_+|^2\,dx \le C\int_\Omega |(s_0-\sigma_0)_+|^2\,dx = 0, \tag{81}$$

since $(s_0-\sigma_0)_+ = 0$ from the hypothesis $s_0\le\sigma_0$ a.e. in Ω; this shows that $(s-\sigma)_+=0$, i.e. $s\le\sigma$.

A consequence is that

$$a\le s_0\le b \text{ a.e. in } \Omega \text{ for constants } a,b\in\mathbb{R} \text{ implies } a\le s\le b \text{ a.e. in } \Omega\times(0,T), \tag{82}$$

since if the initial data σ_0 is a constant c (either a or b) then the solution is $\sigma=c$.

7. Another regularity property of the solution (for $\varepsilon>0$).

There is a formal estimate obtained by multiplying the equation by $-s_{xx}$ and integrating by parts, which supposes that one has enough regularity for writing $(s_x^3)_x$ as a pointwise product $3s_x^2s_{xx}$, and in this way one obtains

$$\Big(\int_\Omega \frac{s_x^2}{2}\,dx\Big)_t + \varepsilon\int_\Omega (s_xs_{xx})^2\,dx = \int_\Omega F(s)\,s_x(s_xs_{xx})\,dx \le ||F||_\infty|s_x|_2|s_xs_{xx}|_2 \le \eta\,|s_xs_{xx}|_2^2 + \frac{||F||_\infty^2|s_x|_2^2}{4\eta}, \tag{83}$$

so that, by choosing $\eta<\varepsilon$ and using Gronwall's inequality, one deduces that[20]

$$\begin{aligned}&s_0\in H^1(\Omega) \text{ implies that the solution of (36) satisfies}\\ &s_x\in L^\infty\big(0,T;L^2(\Omega)\big), s_x^2\in L^2\big(0,T;H_0^1(\Omega)\big), \text{ so that } s_x\in L^6\big(0,T;L^\infty(\Omega)\big),\end{aligned} \tag{84}$$

hence $F(s)\,s_x^2\in L^3\big(0,T;L^\infty(\Omega)\big)\cap L^\infty\big(0,T;L^1(\Omega)\big)$, and $(s_x^3)_x$ (which can be written as $\frac{3}{2}s_x(s_x^2)_x$) belongs to $L^{3/2}\big(0,T;L^2(\Omega)\big)\cap L^2\big(0,T;L^1(\Omega)\big)$, which then permits to find a bound for s_t in the latter intersection, hence $s_t\in L^p\big(0,T;L^q(\Omega)\big)$ with $1\le q\le 2$ and $\frac{3}{p}=\frac{5}{2}-\frac{1}{q}$.

One may prove that (84) is valid by one of two methods already taught by Jacques-Louis Lions in the late 1960s [Li]: the first one is to use a special Faedo–Galerkin basis, made of eigenfunctions of the operator $-\frac{d^2}{dx^2}$, i.e. $w_j(x)=\cos\frac{(j-1)\,\pi\,(x+L)}{2L}$ for $j=1,\ldots,$ in which case a linear combination of $w_1,\ldots,w_n$ gives $-s_{mxx}$ for the solution s_m of (41), and the above computations hold for s_m, and the bounds obtained are inherited by the limit s; the second one is to start from the solution s of (36) and prove that it is more regular by using the method of translations of Louis Nirenberg. For applying this method, one denotes by τ_h the operator of translation by h, i.e. $(\tau_hv)(x)=v(x-h)$ for a function v, but since one works on a bounded interval $\Omega=(-L,+L)$, it is useful to extend u by symmetry around $-L$ and $+L$ (i.e. $u(x)=u(-2L-x)$ for $x\in(-3L,-L)$, or $u(x)=u(2L-x)$ for $x\in(+L,+3L)$) which gives a periodic function of period $4L$, so that all translations are then well defined, and the norms $|\cdot|_p$ considered are computed on a period.[21] One

[20] The last bound for s_x in (84) follows from taking $v=s_x^2$ and using the classical estimates $|v|_\infty\le|v_x|_2^{1/2}|v|_2^{1/2}$ for $v\in H_0^1(\Omega)$ and $|v|_2\le|v|_1^{1/2}|v|_\infty^{1/2}$ by Hölder inequality, which give $|v|_\infty\le|v_x|_2^{2/3}|v|_1^{1/3}$.

[21] Since the equation only involves s, s_x^2, and s_{xx} which are invariant by changing x into $c-x$, and s satisfies a Neumann condition at $\pm L$ (for which the precise meaning uses the space $H(div)$ studied by Jacques-Louis Lions), this type of extension produces a function which satisfies the same partial differential equation.

subtracts the equation for s and for a translated $\tau_h s$, and one multiplies by $s-\tau_h s$, and using the formula $(a^3-b^3)\,(a-b)\geq \frac{3}{4}\,|a^2-b^2|$ for all $a,b\in\mathbb{R}$, one deduces that

$$\frac{1}{2}\,\frac{d}{dt}|s-\tau_h s|_2^2+\frac{\varepsilon}{4}\,|s_x^2-\tau_h s_x^2|_2^2\leq |f-\tau_h f|_2|s-\tau_h s|_2,\ \text{with } f=F(s)s_x^2, \tag{85}$$

and one has the following estimate for $|f-\tau_h f|_2$

$$|f-\tau_h f|_2=|F(s)s_x^2-F(\tau_h s)\tau_h s_x^2|_2\leq ||F||_\infty|s_x^2-\tau_h s_x^2|_2+L_F|s-\tau_h s|_2|s_x|_\infty^2, \tag{86}$$

so that

$$|f-\tau_h f|_2|s-\tau_h s|_2\leq \frac{\varepsilon}{8}\,|s_x^2-\tau_h s_x^2|_2^2+\Big(\frac{2||F||_\infty^2}{\varepsilon}+L_F|s_x|_\infty^2\Big)\,|s-\tau_h s|_2^2, \tag{87}$$

and since $\frac{2}{\varepsilon\,||F||_\infty^2}+L_F|s_x|_\infty^2$ has a finite norm C in $L^1(0,T)$, Gronwall's inequality implies[22]

$$|s-\tau_h s|_2\leq e^C\,|s_0-\tau_h s_0|_2 \text{ on } (0,T),\qquad \int_0^T |s_x^2-\tau_h s_x^2|_2^2\,dt\leq \frac{4e^{2C}}{\varepsilon}\,|s_0-\tau_h s_0|_2^2. \tag{88}$$

If $s_0\in H^1(\Omega)$, then $|s_0-\tau_h s_0|_2\leq |h|\,|(s_0)_x|_2$ for all h, so that $\frac{|s-\tau_h s|_2}{h}$ is bounded in $L^\infty\big(0,T;L^2(\Omega)\big)$ and $\frac{|s_x^2-\tau_h s_x^2|_2}{h}$ is bounded in $L^2\big(0,T;L^2(\Omega)\big)$ when $h\neq 0$ tends to 0; since $\frac{s-\tau_h s}{h}$ tends to s_x and $\frac{s_x^2-\tau_h s_x^2}{h}$ tends to $(s_x^2)_x$ in the sense of distributions as h tends to 0, one deduces (84).

Actually, using $(a^3-b^3)\,(a-b)\geq \frac{3}{4}\,(|a|\,a-|b|\,b)^2$ for all $a,b\in\mathbb{R}$, one has $|s_x|\,s_x\in L^2\big(0,T;H_0^1(\Omega)\big)$, which implies $s_x^2\in L^2\big(0,T;H_0^1(\Omega)\big)$, but the converse is not true.

Remark 7: If $v=|u|\,u\in H_0^1(\Omega)$, one may wonder if v_x is the product of $2|u|$ by u_x, and it is easy to show (using $|\tau_h u-u|\leq C\,|\tau_h v-v|^{1/2}$) that u and $|u|$ belong to the Besov space $X=B_\infty^{1/2,4}$, so that u_x makes sense in the dual Y' of another Besov space, $Y=B_1^{1/2,4/3}$; it is not so clear if one can multiply elements of X by elements of Y'. However, the derivative of u^3 is the product of $3u^2$ by u_x.

8. The class of initial data used (for $\varepsilon>0$).

Given an initial data φ_0, one needs to check if it satisfies our hypotheses, and one first wonders if one can find a curve $M(s)$ of class C^1 with an integrable curvature, as noticed in Remark 2 for stationary solutions, and then if $\varphi_0(x)=M\big(s_0(x)\big)$ for $x\in(-L,+L)$ with s_0 smooth enough.

Since $\varphi_{0x}=s_{0x}M'$ with $|M'|=1$, a sufficient condition is to assume that

$$\varphi_0(x)=a_-\text{ in }[-L,x_-],\quad \varphi_{0x}\neq 0\text{ in }(x_-,x_+),\quad \varphi_0(x)=a_+\text{ in }[x_+,+L]. \tag{89}$$

In order to have s_0 increasing when x varies from $-L$ to $+L$, one chooses

$$s_0(x)=s_-\text{ in }[-L,x_-],\quad s_{0x}=|\varphi_{0x}|\text{ and }M'(s)=\frac{\varphi_{0x}}{|\varphi_{0x}|}\text{ in }(x_-,x_+),\quad s_0(x)=s_+\text{ in }[x_+,+L], \tag{90}$$

so that $s_+-s_-=\int_{x_-}^{x_+}|\varphi_{0x}|\,dx$, and one then assumes that $\varphi_{0x}\in L^1(\Omega;\mathbb{R}^d)$. Then, one also assumes that

$$\varphi_{0xx}\in L^1\big((x_-,x_+);\mathbb{R}^d\big),\text{ so that } s_{0xx}=\Big(\varphi_{0xx},\frac{\varphi_{0x}}{|\varphi_{0x}|}\Big)\text{ in }(x_-,x_+), \tag{91}$$

and since $\varphi_{0xx}=s_{0xx}M'+s_{0x}^2M''$ and M'' is orthogonal to M' because $|M'|=1$, one deduces that $|\varphi_{0x}|^2M''$ is the projection of φ_{0xx} on the orthogonal of φ_{0x}, i.e.

$$|\varphi_{0x}|^2M''=\varphi_{0xx}-\frac{(\varphi_{0xx},\varphi_{0x})}{|\varphi_{0x}|^2}\,\varphi_{0x}\text{ in }(x_-,x_+), \tag{92}$$

[22] With $a=|s-\tau_h s|_2^2$, $b=\frac{\varepsilon}{8}\,|s_x^2-\tau_h s_x^2|_2^2$, and $c=\frac{2}{\varepsilon\,||F||_\infty^2}+L_F|s_x|_\infty^2$, one has $a+2\int_0^t b\leq a(0)+\int_0^t 2c\,a=d$, which satisfies $d_t=2c\,a\leq 2c\,d$, hence $a+2\int_0^t b\leq d\leq a(0)exp\big(2\int_0^t c\big)$, giving $a\leq a(0)\,e^{2C}$ and $2\int_0^T b\,dt\leq a(0)\,e^{2C}$.

hence

$$\int_{s_-}^{s_+} |M''|\,dx = \int_{x_-}^{x_+} \left|\varphi_{0xx} - \frac{(\varphi_{0xx},\varphi_{0x})}{|\varphi_{0x}|^2}\,\varphi_{0x}\right| \frac{dx}{|\varphi_{0x}|}, \tag{93}$$

which one assumes then to be finite.

Then, the condition $s_0 \in W^{1,p}(\Omega)$ follows from $\varphi_{0x} \in L^p(\Omega;\mathbb{R}^d)$, and one notices that assuming (91) implies $\varphi_{0x} \in L^\infty(\Omega;\mathbb{R}^d)$.

One should pay attention to the fact that strong hypotheses of regularity for the initial data are used for transforming the system (10) into the scalar equation (36), but then the system (36) admits uniquely defined solutions even for some discontinuous initial data, so that such discontinuous initial data should be considered as "non physical" for what concerns the system (10).

9. Conclusion.

We have analyzed a system of partial differential equations (10) in one space dimension, and shown that one can study the stationary solutions as well as the evolutionary solutions if the initial data satisfy a geometrical condition, because the analysis can be transformed into studying a scalar partial differential equation, for which we applied more or less classical methods. However, since (10) came out of a special situation for a system of partial differential equations (1) in three space dimensions, motivated by field dislocation mechanics, this work should be considered as a motivation for developing a more general mathematical approach for studying the complete system (1).

References.

[Ac1] Acharya Amit, "A model of crystal plasticity based on the theory of continuously distributed dislocations." J. Mech. Phys. Solids 49 (2001), 761–784.

[Ac2] Acharya Amit, "Driving forces and boundary conditions in continuum dislocation mechanics," R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. 459 (2003), no. 2034, 1343–1363. Proceedings of the Royal Society. A, 459, (2003): 1343–1363. MR1994263 (2004c:74011).

[Ac3] Acharya Amit, "Constitutive analysis of finite deformation field dislocation mechanics." J. Mech. Phys. Solids 52 (2004), no. 2, 301–316. MR2033975 (2004j:74015).

[Ac4] Acharya Amit, "New inroads in an old subject: Plasticity, from around the atomic to the macroscopic scale." J. Mech. Phys. Solids 58 (2010), no. 5, 766–778. MR2642309.

[Ac5] Acharya Amit, "Microcanonical Entropy and mesoscale dislocation mechanics and plasticity." to appear in Journal of Elasticity.

[Ac&Ma&Zi] Acharya Amit & Matthies Karsten & Zimmer Johannes, "Traveling wave solutions for a quasilinear model of field dislocation mechanics." J. Mech. Phys. Solids 58 (2010), 2043–2053.

[Ac&Ro] Acharya Amit & Roy Anish, "Size effects and idealized dislocation microstructure at small scales: predictions of a phenomenological model of Mesoscopic Field Dislocation Mechanics: Part I." J. Mech. Phys. Solids 54 (2006), 1687–1710.

[Ag] Agmon Shmuel, *Lectures on Elliptic Boundary Value Problems*, Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr. Revised edition of the 1965 original. AMS Chelsea Publishing, Providence, RI, 2010. x+216 pp. ISBN: 978-0-8218-4910-1. MR2589244. Van Nostrand Mathematical Studies, No. 2 D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London 1965 v+291 pp. MR0178246 (31 #2504).

[Bi] Birkhoff Garrett, *A Source Book in Classical Analysis*, Edited by Garrett Birkhoff. With the assistance of Uta Merzbach. Harvard University Press, Cambridge, Mass., 1973. xii+470 pp. MR0469612 (57 #9395).

[Da] Dafermos Constantine, *Hyperbolic Conservation Laws in Continuum Physics*. Third edition. Grundlehren der Mathematischen Wissenschaften, 325. Springer-Verlag, Berlin, 2010. xxxvi+708 pp. ISBN: 978-3-642-04047-4 MR2574377. [Second edition. 2005. xx+626 pp. ISBN: 978-3-540-25452-2; 3-540-25452-8. MR2169977 (2006d:35159). First edition. 2000. xvi+443 pp. ISBN: 3-540-64914-X. MR1763936 (2001m:35212)]

[Fo] Fox Norman, "A continuum theory of dislocations for single crystals." IMA Journal of Applied Mathematics 2 (1966): 285–298.

[Kr] Kröner Ekkehart, "Continuum theory of defects." *Physics of Defects*, Les Houches Summer School. North-Holland, 1981. 217–315.
[Li&Se1] Limkumnerd Surachate & Sethna James P., "Mesoscale theory of grains and cells: crystal plasticity and coarsening." Phys. Rev. Lett., (2006): 96, 095503.
[Li&Se2] Limkumnerd Surachate & Sethna James P., "Shocks and slip systems: predictions from a mesoscale theory of continuum dislocation dynamics." J. Mech. Phys. Solids 56 (2008), no. 4, 1450–1459. MR2404020 (2009c:74015).
[Mu] Mura Toshio, "Continuous distribution of moving dislocations." Phil. Mag. 8 (1963): 843–857.
[Na] Nabarro Frank R.N., *Theory of Crystal Dislocations*, Dover Books on Physics and Chemistry, 1987, 821pp. ISBN: 978-0-486-65488-1.
[Po1] Poincaré Henri, "Sur les équations aux dérivées partielles de la physique mathématique." Amer. J. Math. 12 (1890), no. 3, 211–294. MR1505534
[Po2] Poincaré Henri, "Sur les équations de la physique mathématique." Rend. Circ. Mat. Palermo 8 (1894), 57–156.
[St] Stampacchia Guido, *Équations Elliptiques du Second Ordre à Coefficients Discontinus*, Séminaire de Mathématiques Supérieures, No. 16 (Été, 1965) Les Presses de l'Université de Montréal, Montreal, Que. 1966 326 pp. MR0251373 (40 #4603)
[Ta1] Tartar Luc, "Imbedding theorems of Sobolev spaces into Lorentz spaces." Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) 1 (1998), no. 3, 479–500. MR1662313 (99k:46060)
[Ta2] Tartar Luc, "Compensation effects in partial differential equations." Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5) 29 (2005), no. 1, 395–453. MR2305083 (2007k:35341)
[Ta3] Tartar Luc, *An Introduction to Navier–Stokes Equation and Oceanography*, Lecture Notes of the Unione Matematica Italiana, 1. Springer-Verlag, Berlin; UMI, Bologna, 2006. xxviii+245 pp. ISBN: 978-3-540-35743-8; 3-540-35743-2. MR2258988 (2007h:35001).
[Ta4] Tartar Luc, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Lecture Notes of the Unione Matematica Italiana, 3. Springer, Berlin; UMI, Bologna, 2007. xxvi+218 pp. ISBN: 978-3-540-71482-8; 3-540-71482-0. MR2328004 (2008g:46055).
[Ta5] Tartar Luc, *From Hyperbolic Systems to Kinetic Theory: A Personalized Quest* , Lecture Notes of the Unione Matematica Italiana, 6. Springer-Verlag, Berlin; UMI, Bologna, 2008. xxviii+279 pp. ISBN: 978-3-540-77561-4. MR2397052 (2010j:35003).
[Ta6] Tartar Luc, *The General Theory of Homogenization: A Personalized Introduction*, Lecture Notes of the Unione Matematica Italiana, 7. Springer-Verlag, Berlin; UMI, Bologna, 2009. xxii+470 pp. ISBN: 978-3-642-05194-4. MR2582099.
[Wi] Willis John R., "Second-order effects of dislocations in anisotropic crystals." Internat. J. Engrg. Sci. 5 (1967): 171–190.