NAMT
95-003

Beyond Young Measures

Luc Tartar
Carnegie Mellon University

Research Report No. 95-NA-003

January 1995

Sponsors

U.S. Army Research Office
Research Triangle Park
NC 27709

National Science Foundation
1800 G Street, N.W.
Washington, DC 20550

Beyond YOUNG measures

Luc TARTAR

Department of Mathematics, CARNEGIE-MELLON University
Pittsburgh, PA 15213-3890

Dedicated to Jerry ERICKSEN on the occasion of his seventieth birthday

Abstract: YOUNG measures and their limitations are discussed. Some relations between YOUNG measures and H-measures are described and used to analyze an example from micromagnetics. The need to improve H-measures and semi-classical measures is stressed.

At the end of the 70s, I had imagined a new mathematical approach for solving the nonlinear partial differential equations of Continuum Mechanics. I had first taught this new point of view in my cours PECCOT at Collège de France in Paris in March 1977, describing some results on what was later called Homogenization and on the method of Compensated Compactness which I had developed partly in collaboration with François MURAT; I never wrote the corresponding lecture notes, but a few did refer to these lectures, where they had first learned about some concepts that later became classical. A year later in July 1978, at the invitation of Robin KNOPS and John M. BALL, I had lectured at HERIOT-WATT University in Edinburgh on some applications of the Compensated Compactness method ([1]), and the framework was already more complete, unchanged until recently when I introduced H-measures ([2]). I had learned from J. M. BALL to attribute to Laurence C. YOUNG ([3]) these parametrized measures which I had first heard about in the one dimensional setting of Control Theory, and then in a more general setting of Optimization ([4]). My use of YOUNG measures was definitely different, as no derivatives were involved in previous applications, and I was using them in a setting of partial differential equations in order to relate the information which I could deduce from the linear differential balance equations by applying the Compensated Compactness method with the information which I could deduce from the pointwise nonlinear constitutive relations. My approach for the particular situation of Elasticity was quite different from the one that J. M. BALL was advocating ([5]), and instead of studying the class of quasiconvex functionals, related to sequential weak lower semicontinuity and therefore adapted to minimizing energy functionals, I was more interested in identifying which YOUNG measures were adapted to the system of equilibrium equations, adding of course the information that any gradient be curl free. Of course, I did not believe so much in minimization of energy which I considered a too simple approach to Physics or Continuum Mechanics, and I wanted to solve more general evolution problems ([6]), which give rise to quasilinear hyperbolic systems where shocks may occur and such equations are unfortunately not so well understood until now, but I had found a way to use "entropy" conditions in my framework: I was only able to use this new approach for a scalar equation, and it was Ron DIPERNA who opened the way for the case of systems ([7]) (he was lecturing on another subject during the same period, but I had told him about my approach before, when I had found the way to apply it to hyperbolic systems during a visit to the Mathematics Research Center in Madison). From the small amount of references to my work in articles where YOUNG measures appear to be the main tool, one can deduce that many were not able to understand the general framework that I had introduced for handling a large class of the problems in Continuum Mechanics or Physics that I knew at the time. My idea of characterizing the YOUNG measures associated to a given set of linear differential balance equations together with a nonlinear constitutive relation was too difficult, and it is still not so well understood now even in the simple case of YOUNG measures related to gradients which some consider the unique possibility, and I had tried then to create new mathematical tools for that task and other questions as well, as I knew some problems for which one had to go beyond YOUNG measures.

The problem that I want to use here to show how to go beyond YOUNG measures is not one of those where I had first found YOUNG measures to be inadequate, and it is in some way related to ideas of Jerry ERICKSEN. It was around 1983 that some mathematicians learned about J. ERICKSEN's ideas involving Elasticity and Crystals, and since then such questions have appeared often in the work of J. M. BALL, Irene FONSECA and David KINDERLEHRER, sometimes in collaboration with Dick JAMES.

A precise knowledge of how real materials behave is often useful to a mathematician interested in developing adequate mathematical tools for handling in a rigourous way equations from Continuum Mechanics and Physics. In 1985, after a talk where I had explained my general construction of multilayered materials for computing effective coefficients in Homogenization ([8]), J. ERICKSEN had showed me a book containing a photograph of a geometry which looked a little like the ones that I had used (for purely mathematical reasons): it was showing the microstructure of a material which had experienced stress hardening. Of course, my computations were only valid for a diffusion equation, although F. MURAT and Gilles FRANCFORT were busy computing the analogous formulas for linearized elasticity at the time ([9]), but on another occasion J. ERICKSEN had answered one of my questions about the variation of YOUNG's modulus for steel (a much earlier YOUNG, of course) by describing how drastic the stress hardening effect was in the case of a monocrystal of aluminum. J. ERICKSEN's comments had confirmed my general idea that materials like to optimize their microstructures, an idea which I had derived from an earlier conjecture of Daniel D. JOSEPH and Michael RENARDY, conjecture which was actually false because of an effect of apparition of microstructures in optimal design problems according to some results that I had obtained with F. MURAT ten years before. I had also read about this idea of optimality in books by D. JOSEPH ([10]), where he was discussing about turbulent flows then, and so after J. ERICKSEN's comments I had definitely adopted the idea that real materials, in any of the various phases that they may exhibit, are trying to optimize something (and the optimization process should also select in what phase the material should be), an idea which made a perfect fit with my earlier discovery with F. MURAT that microstructures may appear naturally in some optimal design problems ([11]). The problem remains to find what materials are really trying to optimize – and that idea of optimizing something might be only an approximation, of course – but I guessed that better mathematical tools would be necessary then, as I do not like to imagine that the world is so simple that it fits into the mathematical framework that I have already understood. I think that mathematicians should not become too attached to a particular class of problems, even when it is a particularly successful area like the questions about crystalline materials, where the ideas of J. ERICKSEN have been so helpful: after a while, mathematicians should look for what lies beyond, and in that case it might mean to understand why crystals do occur in the first place.

1. YOUNG's measures

When more than twenty years ago I started working with F. MURAT on an academic problem of optimal design, it was not clear that what we were doing had anything to do with composite materials or microstructures. We were looking at a nonclassical problem of Optimization, involving an elliptic equation in variational form, which required using SOBOLEV's spaces, the classical LAX-MILGRAM lemma and a few classical tools of Functional Analysis like weak convergence, but it led us to an improved version of what Sergio SPAGNOLO had called G-convergence a few years before. Some articles of Enrique SANCHEZ-PALENCIA, who was working with materials having a periodic microstructure and was deriving effective properties by asymptotic expansions, led us to the new point of view that weak convergence was a mathematical way to describe the relations between microscopic and macroscopic levels, a notion which I had only heard before associated with some awkward probabilistic argument (my use of the term microscopic level corresponds to what others name a mesoscopic level, as one must be quite above the atomic scale in order to invoke the modeling on a continuum level using partial differential equations).

After this first step, where weak convergence replaces the average over a period in the periodic case (or the expectation in those probabilistic schemes that I dislike), one immediately realizes that if a sequence u_n concerges weakly to u_∞, then $f(u_n)$ does not generally converges to $f(u_\infty)$, except if f is an affine function. For example a sequence of characteristic functions χ_n, taking therefore only the values 0 or 1 (a.e., i.e. almost everywhere), may converge weakly (or weakly $\star$ in L^∞) to a function θ taking values in the interval $[0, 1]$, and therefore the relation $\chi_n = \chi_n^2$ a.e. does not imply the same relation $\theta = \theta^2$ a.e. as θ may well take values inside $(0, 1)$. If one works on a LEBESGUE measurable set Ω and if one considers disjoint measurable subsets $\omega_{i,n}$ of Ω, $i = 1, \ldots, k$, whose characteristic functions $\chi_{i,n}$ converge in $L^\infty(\Omega)$ weak $\star$ to θ_i for $i = 1, \ldots, k$, then one can only deduce that $0 \le \theta_i(x)$ a.e. for $i = 1, \ldots, k$, and $\sum_{i=1}^k \theta_i(x) \le 1$ a.e., or $\sum_{i=1}^k \theta_i(x) = 1$ a.e. if the $\omega_{i,n}$, $i = 1, \ldots, k$, form a partition of Ω for every n. This framework is useful for describing a fine mixture of k different materials, with material $\#i$ occupying the region $\omega_{i,n}$ for $i = 1, \ldots, k$, and it is certainly a mathematical idealization to let n go to ∞ and to consider that a fine mixture is described by the weak $\star$ limits θ_i, $i = 1, \ldots, k$, $\theta_i(x)$ representing the local proportion of material $\#i$ in the mixture around the point x (which only makes sense a.e.).

All the methods – at least those known to me – which are used for describing the relations between microscopic and macroscopic levels, contain such an idealization, but the interesting feature of my approach using various types of weak convergence (as some other types are necessary for Homogenization), is that the weak or weak $\star$ topologies which are used are metrizable on bounded sets and therefore the mathematical idealization of letting n tend to ∞ is not really necessary and actually means "if $\chi_{i,n}$ is near enough to θ_i, then some equations are approximately satisfied", and the possibility remains to improve this kind of result by explaining how near enough needs to be, and what the next corrections are. It seems better then than the probabilistic approach of averaging on an ensemble of realizations, which leaves the possibility that our universe has not been designed to look like all the others that could have been designed as well. In a periodic setting with a small characteristic length ε, it is equivalent to say "if ε is small enough", but I maintain that the hypothesis of a periodic structure is a restriction that Nature does not really follow, even in the case of crystals! It remains that some models derived for crystals, like those that J. ERICKSEN has introduced, may predict in a quite precise manner the behaviour of these crystals in various circumstances, but my interpretation of the physical reality (reality which some consider to be an illusion, anyway) is that it is probably described by partial differential equations (at least this is the best mathematical explanation at the moment), and some partial differential equations may quite well produce an almost perfect piece of periodic pattern without having themselves any periodic structure, and I would like to understand how this happens; I agree that this point of view is quite opposite to the physicists' point of view, who are too much used to put in a model what has been already observed, but I only want to point out that mathematicians should be more aware of the logical defect of that approach.

A different questionable practice of physicists is that of extending to a general nonperiodic situation some computations made in a periodic setting, and they forget usually to wonder why this approach works as it is not the tradition among them to worry about why one obtains "experimentally right" results using "logically wrong" arguments. Mathematicians are supposed to clarify these matters, and I have obtained some partial successes in that direction in my work on Homogenization, Compensated Compactness and H-measures, and there are now ways of explaining some of these lucky successes. Of course, a lot remains to be done, and I will show some limitations of these more or less recently introduced mathematical tools, but before explaining what these new tools are and why they can be useful on questions intractable with the sole use of YOUNG measures, I first recall what YOUNG measures are.

If a sequence U_n of functions defined on a measurable set Ω of R^N and taking (a.e.) values in a subset K of R^p converges in $L^\infty(\Omega)$ weak $\star$ to U_∞, it is not true in general that U_∞ takes its values in K, and it is almost the best possible result that U_∞ takes (a.e.) its values in $\overline{conv}(K)$, the closed convex hull of K. What YOUNG measures do is a little more than describing what the weak $\star$ limit of U_n can be, as they describe simultaneously the weak $\star$ limit of every function of U_n, at the expense of extracting a subsequence.

Theorem 1. a) If U_n is a bounded sequence in $L^\infty(\Omega; R^p)$, taking (a.e.) its values in K, then there exists a subsequence U_m and a (weakly measurable) family of probability measures ν_x on R^p, $x \in \Omega$, with support in $\overline{K}$, the closure of K, such that for every real continuous function F on $\overline{K}$, $F(U_m)$ converges in $L^\infty(\Omega)$ weak $\star$ to the function l_F defined by $l_F(x) = \langle \nu_x, F \rangle$ a.e. $x \in \Omega$. The family ν_x, $x \in \Omega$, is called the YOUNG measure associated to the subsequence U_m.
b) If K is bounded, then for any (weakly measurable) family of probability measures ν_x, $x \in \Omega$, with support in $\overline{K}$, there exists a sequence U_n bounded in $L^\infty(\Omega; R^p)$ and taking its values in K, such that for every real continuous function F on $\overline{K}$, $F(U_m)$ converges in $L^\infty(\Omega)$ weak $\star$ to l_F defined by $l_F(x) = \langle \nu_x, F \rangle$ a.e. $x \in \Omega$.

In particular if U_n converges to U_∞ in $L^\infty(\Omega; R^p)$ weak $\star$, then $U_\infty(x)$ is (a.e.) the center of mass of ν_x and therefore $U_\infty(x) \in conv(\overline{K}) \subset \overline{conv}(K)$. In the example of the characteristic functions $\chi_{i,n}$ of disjoint measurable subsets $\omega_{i,n}$ of Ω, which converge in L^∞ weak $\star$ to θ_i as n tends to ∞ for $i = 1, \ldots, k$, then one defines the function U_n from Ω into R^p by $(U_n)_i = \chi_{i,n}$, $i = 1, \ldots, k$, and one finds that – without extracting a subsequence – the YOUNG measure ν_x in that case is $\nu_x = \sum_{i=1}^{k} \theta_i(x) \delta_{e_i} + \Big(1 - \sum_{i=1}^{k} \theta_i(x)\Big) \delta_0$, where c«, i = 1,..., A, is the canonical basis of *RP*. If the u>„m, s = 1,..., *k*, form a partition of ft for every *n*, then one has $\prod_{i=1}^{k}$ T^Xi,n 00 = 1 a.e. x G ft, and therefore $\sum_{i=1}^{k}$]£}0i(*) = 1 a.e. x € ft.

If /f is unbounded, the result in part b) can be false, but only because the weak • topology is not metrizable, and it is true if one uses the language of general topology, replacing sequences by filters (or nets).

For sequences which are not bounded in *L°°(Q;RP)*, one needs to be more careful; in the case of a sequence bounded in L^r*(Q\RP)* with 1 ≤ r < oo for example, it is better to use only continuous functions *F* such that $\lim_{|t/Hoo}$ |F(v')| / ||t;||^r = 0, in order to avoid concentration effects, that YOUNG measures are not designed to take into account, because they are based on the use of TV-dimensional LEBESGUE measure and cannot therefore see what happens on sets of //-dimensional measure zero. One can easily extend Theorem 1 for general sequences (i.e. not necessarily bounded in *L°°(u;* R^p*))*, in the following way: for *p* > 0 one defines the mapping $*_p$ from *RP* into itself by 4,(v) = *v* if ||v|| ≤ *p* and $*_p$(v) = J *pv*/||v|| if ||v|| ≥ *p*, and one extracts then a (diagonal) subsequence U_m such that for any integer *p* the sequence $\$_p(U_m)$ defines a YOUNG measure v_p\ because $<b_p = \$_p$ o $\$_e$ for *a* ≥ *p*, one deduces that for *a* ≥ *p* the probability measure $i/_{PtX}$ is the image by $\$_p$ of $i/_{fffX}$, i.e. (*>,*,V>) = ("*,*, V> ° $p) for every continuous function ^> a.e. x £ ft, and as *p* tends to infinity the sequence $i/_{PtX}$ converges vaguely to a probability measure *v*^,* °B the compactification of *RP* with a sphere at infinity, which I denote #? = *RP* U {^S**[1]}. The function *, can be extended to /2? by defining $\$_p$(°°*z*) = *pz* for every *z* € $S^{p\sim1}$, and the probability measure u_{PiX} is the image by $<b_p$ of i/oo_{tX} for almost every *x* € ft. A real function *F* is continuous on *RP*, if and only if its restriction to *RP* is continuous and has radial limits at infinity (corresponding to the restriction of *F* to the sphere at infinity °°$S^{p\sim l}$), i.e. for any sequence v_n converging to °°z at infinity, or equivalently ||t;$_n$|| —• oo and v/ll^nll —• 2, one has $F(v_n)$ —* *F(°°z);* a particular example of such a function is any function *F* which is continuous and tends to 0 at infinity (a situation related to the one point ALEXANDROFF's compactification of *RP)*, which was the choice made by J. M. BALL ([12]).

Theorem 2. a) If U_n is a sequence of measurable functions from fi into *RP* taking (a.e.) its values in *K*, then there exists a subsequence U_m and a (weakly measurable) family of probability measures VQQ_{JX} on *RP*, *x* € ft, with support in 7f~, which denotes the closure of *K* in *RP*, such that for every real continuous function *F* on i? (or on 7T"), $F(U_m)$ converges in L°°(ft) weak • to the function *lp* defined by /F(X) = *{1/^ x,F)* a.e. x€ft.
b) If lim (limsupmeas{x € ft : ||t/^(x)|| > *p]*) = 0, then i/^,* does not charge $^{00}5^{p-1}$, a.e. x € ft, and one obtains a YOUNG measure in ft. If U_n —^£ t/00 in i^1(ft) weak and defines a family of probability measure "oo,* on /#>, then j/00,* does not charge the sphere at infinity, the function *U* •-* *U*\\ is $i/oo_{>r}$ integrable and the center of mass of 1/00,* is *Uoo(x)*, a.e. x € ft. If *G* is a continuous mapping from *RP* to *Ri* and the sequence U_n defines a family of probability measures $i/oo_{>r}$ on *RP*, then $G(U_n)$ defines a family of probability measures *«>,* on *Rfi*, and Xoo,r is the image by *G* of *>«>,*, a.e. x € ft.

Of course, one can also introduce some variants that can take care of concentration effects: for example if a sequence U_n is bounded in L^r*{il\RP)*, then for a subsequence U_m one can describe all the weak • limits of functions $G(U_m)$ with *G(U)* = *U*\\$^r$*F(U)* for *U* € *RP* and *F* is continuous on *RP*, and these limits are described by a nonnegative measure on ft x $^{OO}S^{P\sim1}$, which sees concentration effects, the choice made by R. DIPERNA *k* Andrew MAJDA ([13]).

2. H-Measures

YOUNG measures cannot notice that the set ft on which our functions are defined is an open set of R^N (or a manifold) and therefore has a differential structure, as the main property needed to define them is that the LEBESGUE measure *dx* has no atoms. Of great importance in Continuum Mechanics are the constitutive relations and the balance equations, and YOUNG measures can handle the constitutive relations which might be nonlinear but are pointwise relations, while they are inadequate for handling linear balance equations because they are given by partial differential equations. I have designed H-measures as an extension of the Compensated Compactness method which I had partly developed with F. MURAT, and they can handle some linear "pseudo-differential" equations and predict the weak $\star$ limit of some quadratic quantities, but they are not designed to handle general nonlinear relations. One expects then that by using both YOUNG measures and H-measures one can obtain a better understanding of the equation of Continuum Mechanics, and indeed I will describe a problem where both these measures are needed, although it should be emphasized that the relations between these two objects are not completely understood yet, and that more general objects should certainly be developed.

If Ω is an open subset of R^N and U_n is a sequence converging to 0 in $L^2(\Omega; R^p)$ weak, then after extraction of a subsequence U_m one can define the H-measure associated to the subsequence U_m, which is a nonnegative Hermitian $p \times p$ matrix of RADON measures μ in (x,ξ) with $x \in \overline{\Omega}$ and $\xi \in S^{N-1}$, the unit sphere in R^N. It permits to compute the weak $\star$ limits (in the space of RADON measures) of products of the type $L_1(U_{m,i})\overline{L_2(U_{m,j})}$ where U_m has been extended by 0 outside Ω and L_1 and L_2 are some "pseudo-differential" operators of order 0.

The "pseudo-differential" operators that I chose to use in the theory are not the ones with smooth symbols which are classical in some circles, as smoothness would preclude many applications to a Continuum Mechanics framework. These operators have the property to map $L^2(R^N)$ into itself (which is why I call them operators of order zero) and their symbols are of the form

$$s(x,\xi) = \sum_{k=1}^{\infty} a_k(\xi) b_k(x)$$

with $a_k \in C(S^{N-1})$, the space of continuous functions on the unit sphere and $b_k \in C_0(R^N)$, the space of continuous functions converging to 0 at infinity, $k \geq 1$, and one imposes that

$$\sum_{k=1}^{\infty} ||a_k||_{L^\infty(S^{N-1})} ||b_k||_{L^\infty(R^N)} < \infty.$$

Defining the FOURIER transform $\mathbf{F}$ by

$$\mathbf{F}f(\xi) = \int_{R^N} e^{-2i\pi(x.\xi)} f(x)\, dx,$$

one associates to an admissible symbol s the standard operator S (that one can denote $s\left(x, \frac{D}{2i\pi}\right)$) by

$$\mathbf{F}(Su)(\xi) = \sum_{k=1}^{\infty} a_k\left(\frac{\xi}{||\xi||}\right) \mathbf{F}(b_k u)(\xi), \text{ a.e. } \xi \in R^N, \text{ for } u \in L^2(R^N).$$

To a symbol $b \in C_0(R^N)$ is associated the operator M_b of multiplication by b, while to a symbol $a \in C(S^{N-1})$ extended to be homogeneous of degree zero in R^N is associated the operator $\mathbf{F}^{-1} M_a \mathbf{F}$. A linear continuous operator L from $L^2(R^N)$ into itself is said to have symbol s if $L - S$ is a compact operator from $L^2(R^N)$ into itself.

Theorem 3. If the subsequence U_m defines a H-measure μ and L_1, L_2 are operators with symbols s_1 and s_2, then

$$L_1(U_{m,i})\overline{L_2(U_{m,j})} \rightharpoonup \nu \text{ in the weak } \star \text{ convergence of measures,}$$

and one has

$$\langle \nu, \varphi \rangle = \langle \mu^{ij}, \varphi s_1 \overline{s_2} \rangle \text{ for every } \varphi \in C_c(\Omega),$$

the space of continuous functions with compact support in Ω.

An important remark, which I called the Localization Principle, expresses how H-measure are constrained by any linear "pseudo-differential" information on the subsequence U_m.

Theorem 4. If the subsequence U_m defines a H-measure μ and satisfies

$$\sum_{i=1}^{N}\sum_{j=1}^{p}\frac{\partial}{\partial x_i}(A_{ij}U_{m,j}) \to 0 \text{ strongly in } H^{-1}_{loc}(\Omega),$$

where A_{ij} are continuous functions in Ω, one has

$$\sum_{i=1}^{N}\sum_{j=1}^{p}\xi_i A_{ij}(x)\mu^{jk} = 0 \text{ in } \Omega \text{ for } k = 1,\dots,p.$$

Actually these two conditions are equivalent: if for $\varphi \in C_c(\Omega)$ one defines V_m by

$$V_m = \varphi\sum_{i=1}^{N}\sum_{j=1}^{p}R_i(A_{ij}U_{m,j}),$$

where R_i denotes the RIESZ transform (the operator of symbol ξ_i), then the first condition means that $V_n \to 0$ in $L^2(R^N)$ strong for every $\varphi \in C_c(\Omega)$, and the limit of $V_m\overline{V_m}$ is $\sum_{i,i'=1}^{N}\sum_{j,k=1}^{p}\langle\mu^{jk}, |\varphi|^2\xi_i\xi_{i'}A_{ij}\overline{A_{i'k}}\rangle = 0$ and therefore the first condition is equivalent to $\sum_{j,k=1}^{p}\mu^{jk}\Big(\sum_{i=1}^{N}\xi_i A_{ij}\Big)\overline{\Big(\sum_{i=1}^{N}\xi_i A_{ik}\Big)} = 0$, which by the Hermitian nonnegative character of H-measures is equivalent to the second condition. The Localization Principle improves the results given by the Compensated Compactness method, which could only handle the case of constant coefficients and discuss the possible weak $\star$ limits of quadratic quantities, and H-measures also give new results in other directions.

As an example, which is used in the following problem, let a sequence M_n converge to 0 in $L^2(\Omega; R^3)$ weak and correspond to a H-measure μ, which is a 3×3 Hermitian nonnegative matrix whose entries are scalar (eventually complex) RADON measures in $R^3 \times S^2$. If one solves the equation $div(grad(u_n) + M_n) = 0$ in R^3, then the mapping $M \mapsto grad(u)$ where u is the solution of $div(grad(u) + M) = 0$ (which is a 3×3 matrix of admissible operators from $L^2(R^3)$ into itself) has symbol $-\xi \otimes \xi$, and the weak $\star$ limit of $||grad(u_n)||^2$ is the measure π defined by $\langle\pi,\varphi\rangle = \sum_{i,j=1}^{3}\langle\mu^{ij}, \varphi\otimes\xi_i\xi_j\rangle$ for every $\varphi \in C_c(R^3)$. The condition $\sum_{i,j=1}^{3}\mu^{ij}\xi_i\xi_j = 0$ is equivalent to $div(M_n) \to 0$ in $H^{-1}_{loc}(\Omega)$ strong.

3. A problem in Micromagnetics, involving YOUNG measures and H-measures

The solutions of many important problems seem to require the use of a mathematical object, yet to be developed, which will encompass both YOUNG measures and H-measures. Partial results about the relations between YOUNG measures and H-measures have already been obtained with F. MURAT and I will show how they can be used on the following example.

The motivation for this example is the model of micromagnetics of William BROWN [14], for a crystal occupying a bounded open domain Ω of R^3, a problem studied by R. JAMES & D. KINDERLEHRER [15], using only the tool of YOUNG measures, so they could only obtain partial results. In their work, they considered the equation

$$-div\Big(grad(u) + m\chi_\Omega\Big) = 0 \text{ in } R^3,$$

where χ_Ω is the characteristic function of Ω and m satisfies the constraint

$$||m(x)|| = 1 \text{ a.e. in } \Omega,$$

and they sought m minimizing the quantity $J_0(m)$ defined by

$$J_0(m) = \int_{R^3} ||grad(u)||^2\, dx + \int_\Omega \Big(\varphi(m) - (H_0.m)\Big)\, dx.$$

In that model, where a few physical parameters have been suppressed, m corresponds to a macroscopic spin variable, the function φ corresponds to an anisotropic energy due to the crystalline nature of the body and H_0 is an applied magnetic field, the magnetic field H being $grad(u)$ and the magnetic induction field B being $H + m$. An exchange energy $E_\varepsilon(m)$, of a form similar to

$$E_\varepsilon(m) = \varepsilon^2 \int_\Omega \sum_{i,j=1}^{3} \left|\frac{\partial m_i}{\partial x_j}\right|^2 dx,$$

involving a small length scale ε, has been neglected, and an initial question is to minimize the total energy $J_\varepsilon(m)$ defined as

$$J_\varepsilon(m) = J_0(m) + E_\varepsilon(m).$$

By a classical argument of compactness, the functional J_ε certainly attains its global minimum for at least one configuration m_ε under the minimal assumption that φ is a continuous function on the sphere S^2 and $H_0 \in L^2(\Omega; R^3)$. It might well be that $J_\varepsilon(m)$ has numerous local minima, but the physical intuition behind the modeling corresponds to a competition between three tendencies for making the various parts of the total energy small: the first tendency is to minimize the magnetostatic energy $\int_{R^3} ||grad(u)||^2\, dx$ by annuling the magnetic field, which happens if $div(m) = 0$ in Ω and the normal component $(m.n)$ of m is zero on $\partial\Omega$; the second tendency is to minimize the anisotropic energy $\int_\Omega (\varphi(m) - (H_0.m))\, dx$; the third tendency is to minimize the exchange energy by having m constant. In the case where $\varepsilon = 0$ and $H_0 = 0$ one can sometimes satisfy the first two tendencies by having m oriented along an easy axis which minimizes φ – which is chosen as an even function on S^2 – the orientation along the easy axis being chosen so that the conditions $div(m) = 0$ in Ω and $(m.n) = 0$ on $\partial\Omega$ can be satisfied; R. JAMES & D. KINDERLEHRER studied when such classical solutions do exist, and noticed that when they do not exist one can use instead a sequence m_n creating a YOUNG measure charging the two directions of an easy axis with equal weights and having a weak $\star$ limit $m_\infty = 0$ annuling therefore the magnetic field in the limit, a scheme which is the mathematical expression of classical arguments on how to create a piecewise constant divergence free field m. Unfortunately, the case of a nonzero applied field H_0 could not be handled as easily by the same argument, which is the reason why I solved the problem using both YOUNG measures and H-measures, although it was shown afterwards by Antonio DE SIMONE how one could extend the classical argument ([16]).

When ε is small the physical intuition is that some magnetic domains are formed where the orientation of m is almost uniform, the variation of m being confined to walls between these domains, the only places where the exchange energy must be taken into account; the walls have a thickness of the order of ε. I was made aware of contributions of physicists like BLOCH, LANDAU and NÉEL on the question of the structure of these walls, and I decided to attack the mathematical problem of describing in a mathematical way the apparition of magnetic domains and walls in the sequence m_ε of minimizers of $J_\varepsilon(m)$, in order to predict the typical size of the domains, the orientations of the walls and the energy attached to them, hoping to derive a precise understanding of the microstructures exhibited. I only suceeded in understanding what happened in the limit $\varepsilon \to 0$ and identified what the relaxed functional of J_0 was, characterizing $\lim_{\varepsilon\to 0} J_\varepsilon(m_\varepsilon)$, by using a combination of YOUNG measures and H-measures and a partial result obtained with F. MURAT concerning the relations between YOUNG measures and H-measures ([17]).

The mathematical difficulty arises from the fact that the functional J_0 is not lower semicontinuous for the natural topology for m, the $L^\infty(\Omega; R^3)$ weak $\star$ topology. Minimizing sequences might then develop oscillations, and this is in qualitative agreement with the experimentally observed formation of small magnetic domains, although the model has been considered dubious because of quantitative discrepancies between the theoretical predictions and the experimental measurements. As pointed out to me by R. JAMBS, the measurements can be rather different if one uses a polycrystal instead of a monocrystal for which the theory was derived, and therefore I want to mention that the computations which I will show can be extended to the case of polycrystals; there might actually be some questions of Homogenization to be considered in the case of heterogeneous materials.

If a sequence m^e converges to m° in $L^{\circ\circ}(Q\backslash R^3)$ weak •, the computation of the limit of $<p(m^e)$ requires more than the weak • limit of m^e (as *(p* is not affine in general) and can be derived from the YOUNG measure *v* associated to a subsequence. On the other hand the limit of $||\wedge racf(ti^e)||^2$ cannot be computed in general by using only the YOUNG measure i/, but it can be derived from the H-measure /i associated to a subsequence of *m* — m°*, the reason being that the mapping *m >-• grad(u)* is given by an admissible "pseudo-differential" operator of order zero. One has

$$/ \varphi(m'M^*)\, dx - / (u_{Xi}\, v)t(x)\, dx$$

and

$$\int_{R^3} \|grad(u<)\|?tl>(x)dx-+ f \quad \|_9 rad(u^\circ)\|\wedge(x)\, dx + \sum_{i,j=1}^{3} \langle \mu^{ij}, \psi \otimes \xi_i \xi_j \rangle$$

for every bounded continuous function V> where u° denotes the solution corresponding to m^0, which only satisfies $||m^\circ(x)|| \leq 1$ a.e. x £ Q, as *m°(x)* is the center of mass of the probability v_x which lives on the sphere S^2. The crucial question is then to understand what relations link the YOUNG measure *v* and the H-measure *fi.* Without answering this question, one can describe an abstract relaxation problem where one seeks to minimize the functional *J* defined in the following way: let *X* be the space of all pairs *(v,n)* for which there exists a sequence m^c satisfying the constraint $||m^r|| = 1$ a.e. x € ft, and such that m^e defines a YOUNG measure *v* and such that m^e — m° defines a H-measure /x, m°(x) being the center of mass of v_x a.e. x € ft; one defines the functional *J* on *X* by the formula

$$J_1(\nu,\mu) = \int_r \|grad(u^\circ)\|^2 dx + \sum^{3} j > ' M \circledR 6 \,\&\,) + \int^{t} (<*,*) - (\#o.m^0))\, dx,$$

where u° is the solution of

$$-div(grad(u^\circ) \text{ -f } m^\circ xn) = 0 \text{ in } iZ^3.$$

By following the construction of *v* and /i, one can put a topology on *X* that makes it compact and renders *J* continuous, so that *J* does attain its minimum on *X.* The initial problem is imbedded into this new one and corresponds to u_x being the DIRAC mass at m°(x), a.e. x € ft, and /i being 0. Of course, the preceding result is only a change of language for recasting the problem, and nothing important has been done yet at this level.

A precise description of *X* is not yet available, but a partial result obtained with F. MURAT shows that for a given YOUNG measure */, there is a pair (f,/i) belonging to *X* such that $\wedge^{3}$ *{/***,* 1 ® &{;) = 0. This gives rise to a second relaxed problem defined on the set *Y* of all YOUNG measures, where one defines the functional *J2* by the formula

$$J_2(\nu) = \int_{JR^*} \|9rad(u^\circ)\|fdx + \int_{Ja} \Big(\langle \nu_x, \varphi\rangle - (H_0.m^0)\Big)\, dx,$$

with m° and «° defined as before. If *Y* is equipped with the weak • topology, then *Y* is compact and *32* is lower semicontinuous and does attain its minimum. The initial problem is imbedded into this new one and corresponds to each v_x being the DIRAC mass at m°(x), a.e. x € ft.

Finally one defines *Z* to be the convex set of functions m° satisfying

$$|m^\circ(x)| \leq 1 \text{ a.e. in f2},$$

and one defines the functional J_3 by

$$J_3(m^0) = \int_{R^3} ||grad(u^0)||^2\,dx + \int_\Omega \Big(\varphi^{**}(m^0) - (H_0.m^0)\Big)\,dx,$$

where u^0 is defined as before and where φ^{**} is the convex function defined on the unit ball by

$$\varphi^{**}(m) = \inf_\nu \langle \nu, \varphi\rangle \text{ for all probability measures } \nu \text{ on } S^2 \text{ with center of mass } m.$$

[φ^{**} is the convex envelope of the function equal to φ on S^2 and $+\infty$ elsewhere]. If one equips Z with the weak $\star$ topology, then Z is compact and J_3 is lower semicontinuous and does attain its minimum. The initial problem is imbedded into this new one and corresponds to m^0 taking its values on the unit sphere (a.e.).

If no solution of this last problem satisfies $||m^0(x)|| = 1$ a.e. $x \in \Omega$, then there is no classical solution mimimizing J_0, and minimizing sequences tend to create somewhere in Ω some tiny magnetic domains, the statistics of orientations for m being described by some YOUNG measure ν appearing in the definition of φ^{**}; the H-measure μ, sees another kind of information, like the orientations of the walls of these magnetic domains, and the purpose of the result with F. MURAT has been to assert that one can construct a sequence satisfying $div(m^\epsilon) \rightarrow div(m^0)$ in $H^{-1}_{loc}(R^3)$ strong. Of course, having neglected the exchange energy, there is nothing that limits the size of the magnetic domains in this simplified model. The preceding analysis also shows that for any minimizing sequence m^ϵ, any subsequence m^η converging in $L^\infty(\Omega; R^3)$ weak $\star$ to m^0 must be such that m^0 minimizes J_3 and that one has $div(m^\eta) \rightarrow div(m^0)$ in $H^{-1}(R^3)$ strong.

I have therefore used H-measures in a way which might not seem so crucial as they did appear in the problem of minimizing J_1 but not in that of minimizing J_2, and the same remark can be applied to YOUNG measures as they did appear in the problem of minimizing J_2 but not in that of minimizing J_3, but this is the result of the analysis, which used YOUNG measures and H-measures as an internal tool, to have discovered what is really important in the situation considered. There seems to be realistic problems, however, like a similar question involving magnetostriction, where the role of the interaction of H-measures and YOUNG measures is not yet completely clarified.

It is worth repeating my opinion expressed so many years ago that another usefulness of studying oscillations at the theoretical level which I have chosen is also to avoid computing these oscillations, which is a time/money consuming task and might not even show clearly what is hapening. In the preceding problem, it would be useless to invent a numerical approach to approximate H-measures, or to use one of the numerical approaches to YOUNG-measures which have been devised already, as the preceding analysis tells that once the function φ^{**} has been computed – which might be an interesting numerical question if no closed form formula is available – one has to solve a convex problem for which efficient methods have been derived some time ago (actually there is another difficulty in the numerical solution of the problem at hand, as it is set in the entire space R^3).

4. Relations between YOUNG measures and H-measures

The more important piece of information used in the analysis of the preceding model was a result obtained with F. MURAT that some pairs (ν, μ) of a YOUNG measure and a H-measure could be obtained for a sequence satisfying some constraints. Of course, we had not been working in view of the preceding example, and we had in mind a more general question, corresponding to the program that I had devised many years ago for the Compensated Compactness method, program built around trying to understand more about the structure of the partial differential equations of Continuum Mechanics. On an open set Ω of R^N one considers a sequence $U_n \in L^\infty(\Omega; R^p)$ satisfying two types of constraints: the first one, related to constitutive relations is of the form

$$U_n(x) \in K \text{ a.e. } x \in \Omega,$$

where K is a nonempty subset of R^p, and the second one, related to balance equations is of the form

$$\sum_{j=1}^p \sum_{k=1}^N \frac{\partial}{\partial x_k}\Big(A_{ijk}(x)(U_n)_j\Big) \in \text{ compact of } H^{-1}_{loc}(\Omega) \text{ for } i = 1, \ldots, q,$$

where the functions A_{ijk} are continuous (the Compensated Compactness method could only handle constant coefficients). One then looks at a subsequence U_m converging to U_∞ in $L^\infty(\Omega;R^p)$ weak $\star$ and defining a YOUNG measure ν, and such that $U_m - U_\infty$ defines a H-measure μ, and the problem is to characterize which pairs (ν,μ) can be obtained in this way. Of course, for some applications where sequences are not necessarily bounded, one should then relax the hypothesis of boundedness in $L^\infty(\Omega;R^p)$.

Due to the fact that H-measures can see the balance equations in a quite transparent way, the natural problem was then to take only into account the constitutive relations, because if one could characterize the admissible pairs (ν_0,μ_0) that could be obtained by taking only into account the constitutive relations, one would obtain a solution of the complete problem by considering the pairs (ν_0,μ_0) satisfying also

$$\sum_{j=1}^{p}\sum_{k=1}^{N} A_{ijk}\xi_k\mu_0^{jl} = 0 \text{ for } i = 1,\ldots,q, \text{ and } l = 1,\ldots,p.$$

For a given YOUNG measure ν_0 compatible with the constitutive relations, i.e. having support in $\overline{K}$, we looked then for the H-measures μ_0 which are such that the pair (ν_0,μ_0) is admissible, and we first proved that the set of such μ_0 is convex (closed for the weak $\star$ topology, of course). There are some obvious necessary conditions that the admissible pairs (ν_0,μ_0) must satisfy: as both YOUNG measures and H-measures can predict the weak $\star$ limit of real quadratic quantities (or sesquilinear quantities), these predictions must be the same, and this gives the following relations, where $U_\infty(x)$ is the center of mass of ν_{0x} and Q is an arbitrary sesquilinear form given by $Q(U) = \sum_{i,j=1}^{p} q_{ij}U_i\overline{U_j}$,

$$\int_\Omega\Big(\langle\nu_{0,x},Q\rangle - Q(U_\infty)\Big)\varphi(x)\,dx = \sum_{i,j=1}^{p} q_{ij}\langle\mu^{ij},\varphi\otimes 1\rangle \text{ for every } \varphi \in C_c(\Omega).$$

It is an open question to derive other general necessary conditions. We focused our attention then on improving our understanding of which pairs we could construct, and we tried to construct as many pairs as possible for which the YOUNG measure is a finite combination of DIRAC masses $\nu_0 = \sum_{i=1}^{r}\theta_i\delta_{V^{(i)}}$ in an open set ω, with $V^{(i)} \in K$, $\theta_i \geq 0$, $i = 1,\ldots,r$ and $\sum_{i=1}^{r}\theta_i = 1$. In that case, denoting by V^G the center of mass of ν_0, i.e.

$$V^G = \sum_{i=1}^{r}\theta_i V^{(i)},$$

the compatibility condition becomes

$$Proj_1\mu_0 = \Big(\sum_{i=1}^{r}\theta_i(V^{(i)} - V^G)\otimes(V^{(i)} - V^G)\Big)dx,$$

where $Proj_1$ is the first projection transforming a measure in (x,ξ) into a measure in x, i.e. integration in ξ. Then by a limiting process one can construct admissible pairs where ν_0 is a general probability measure with support in a bounded set of $\overline{K}$.

Our method of construction was based on the result of small amplitude Homogenization for which I had first introduced H-measures. Let ω be a bounded open set of R^N, let $A_0 \in C^0\big(\omega;\mathcal{L}(R^N,R^N)\big)$ be uniformly elliptic (i.e. there exists $\alpha > 0$ such that $(A_0(x)\lambda.\lambda) \geq \alpha||\lambda||^2$ for all $\lambda \in R^N$ a.e. $x \in \omega$), let $B_n \rightharpoonup 0$ in $L^\infty\big(\omega;\mathcal{L}(R^N,R^N)\big)$ weak $\star$ and corresponding to a H-measure μ, then for γ small the operator $\mathcal{A}_n$ defined by

$$\mathcal{A}_n v = -div\Big((A_0 + \gamma B_n)grad(v)\Big) \text{ for } v \in H_0^1(\omega),$$

is an isomorphism from $H_0^1(\omega)$ onto $H^{-1}(\omega)$ and, after extracting a subsequence $\mathcal{A}_m$, the sequence of inverse operators $\mathcal{A}_m^{-1}$ converge weakly to the inverse $\mathcal{A}_{eff}^{-1}$ of an effective operator $\mathcal{A}_{eff}$ of the form

$$\mathcal{A}_{eff}v = -div\Big((A_{eff}(x;\gamma)grad(v)\Big) \text{ for } v \in H_0^1(\omega),$$

where $A_{eff}(x;\gamma) = A_0(x) + \gamma^2 C_2(x) + \ldots$ is analytic in γ and the first corrector C_2 can be computed using A_0 and μ. For example if $B_n = b_n I$ with $b_n \rightharpoonup 0$ in $L^\infty(\omega)$ weak $\star$ and corresponding to a scalar nonnegative H-measure μ, then the formula is

$$C_2(x)\,dx = Proj_1\Big(\mu\frac{\xi\otimes\xi}{(A_0(x)\xi.\xi)}\Big).$$

For each microgeometry for which one can compute an effective conductivity, one can deduce a corresponding H-measure; on the other hand, in order to find optimal bounds for effective conductivity I had been led to derive a method for computing the effective conductivity of multilayered materials, and this method gives then a way to compute the H-measure corresponding to a multilayered geometry. Of course, the main difficulty for the method of deriving bounds for the effective conductivity is that the reiteration of the method gives rise to inextricable computations; however it is not so for the case of H-measures because one is only interested in a TAYLOR expansion at order two and the computations become somewhat more tractable, although quite intricate so that we have not been able to derive yet a simple characterization of all the H-measures that one can construct by following our algorithm. Nevertheless we had encountered early a particular geometry giving rise to a relatively symmetric formula, and this is the particular case that I used for the model of micromagnetics: for $V^{(i)} \in K$, $\theta_i \geq 0$, $i = 1,\ldots,r$ and $\sum_{i=1}^r \theta_i = 1$, let $V^G = \sum_{i=1}^r \theta_i V^{(i)}$ and let π_i, $i = 1,\ldots,r$, be (independent) probabilities on the unit sphere S^{N-1}, then there exists a sequence U_n of the form

$$U_n(x) = \sum_{i=1}^r \chi_{n,i}(x)V^{(i)},$$

where the $\chi_{n,i}$ are characteristic functions of disjoint measurable subsets $\omega_{n,i}$ of ω, with

$$\chi_{n,i} \rightharpoonup \theta_i \text{ in } L^\infty(\omega) \text{ weak } \star,$$

and such that the sequence corresponds to the YOUNG measure

$$\nu_0 = \sum_{i=1}^r \theta_i \delta_{V^{(i)}},$$

and $U_n - V^G$ corresponds to the H-measure μ_0 defined by

$$\mu_0 = \sum_{i=1}^r \theta_i\Big((V^{(i)} - V^G)\otimes(V^{(i)} - V^G)\Big)\Big(\pi_i \otimes dx\Big).$$

One can see on this formula then how in the example (where $N = p = 3$), μ_0 can satisfy $\sum_{j,k=1}^p \mu_0^{jk}\xi_j\xi_k = 0$, as the left hand side is

$$\sum_{j,k=1}^p \sum_{i=1}^r \theta_i(V^{(i)} - V^G)_j(V^{(i)} - V^G)_k\xi_j\xi_k(\pi_i \otimes dx),$$

i.e. $\sum_{i=1}^r \theta_i\Big(\sum_{j=1}^p (V^{(i)} - V^G)_j\xi_j\Big)^2\Big(\pi_i \otimes dx\Big)$, and this can be made 0 by choosing the probability π_i such that $\sum_{j=1}^p (V^{(i)} - V^G)_j\xi_j = 0$ on the support of π_i (if $\theta_i > 0$, of course), i.e. by taking π_i supported by the equator perpendicular to $V^{(i)} - V^G$ (or anywhere if $V^{(i)} = V^G$). Of course, one must extend this argument to a general YOUNG measure ν, and this is done by noticing that if one has constructed admissible pairs (ν_i, μ_i) in disjoint open sets ω_i, $i = 1, \ldots, I$, whose union is Ω up to a set of measure 0, then one can glue the sequences together and obtain an admissible pair (ν, μ) on Ω, because the functions used are bounded in L^∞ and therefore no concentration effect in x can occur, and nothing happens then on the set of measure zero which is missing. Then using the fact that the set of admissible pairs corresponding to a bounded K is (sequentially) closed for the natural weak $\star$ topology, one can extend the result to a more general ν.

Actually the same argument works, without even needing the points $V^{(i)}$ to be in any special position, when the balance equations are such that

$$\text{for every } \lambda \in R^p \text{ there exists } \xi \in S^{N-1} : \sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \xi_k \lambda_j = 0 \text{ for } i = 1, \ldots, q,$$

i.e. $\Lambda = R^p$ in the notations of the Compensated Compactness method, where

$$\Lambda = \Big\{ \lambda \in R^N : \exists \xi \in S^{N-1}, \sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \xi_k \lambda_j = 0 \text{ for } i = 1, \ldots, q \Big\}.$$

In the case $\Lambda \neq R^p$, the argument works if $V^{(i)} - V^G \in \Lambda$ for $i = 1, \ldots, r$. Indeed, in order to have $\sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \xi_k \mu^{jl} = 0$ for $i = 1, \ldots, q$, and $l = 1, \ldots, p$, one needs to have $\sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \xi_k (V^{(s)} - V^G)_j = 0$ on the support of π_s (if $\theta_s > 0$), and therefore as $\lambda = V^{(s)} - V^G \in \Lambda$, there is a ξ associated with λ in the definition of Λ, and one can take $\pi_s = \delta_\xi$.

Actually the result obtained with F. MURAT tells a little more as it also describes the H-measures of functions of U_n: if F is a function from R^p to a finite dimensional vector space E, and F^G denotes the center of mass of the points $F(V^{(i)})$, i.e.

$$F^G = \sum_{i=1}^{r} \theta_i F(V^{(i)}),$$

then the H-measure μ_F associated to $F(U_n)$ is

$$\mu_F = \sum_{i=1}^{r} \theta_i \Big((F(V^{(i)}) - F^G) \otimes (F(V^{(i)}) - F^G) \Big) \Big(\pi_i \otimes dx \Big),$$

and this formula shows more clearly the nature of a H-measure associated to a sequence taking its values in E: it maps continuous functions with compact support in (x, ξ) into $\mathcal{L}(E', E)$.

5. Generalizations

In the preceding problem, one could change the pointwise constraint on m into $m \in K$, where K is any bounded set of R^3 (and even into $m(x) \in K(x)$ with $K(x)$ varying in a reasonable way at the expense of invoking a few more classical results of approximation from Measure theory), and of course the YOUNG measures in that case are supported by $\overline{K}$. One can also consider an heterogeneous material, like a polycrystal instead of a monocrystal, using then an anisotropic energy $\varphi(x, m)$ (satisfying CARATHÉODORY conditions), and a more general relation $B(x) = M(x)\big(H(x) + m(x)\big)$ with a symmetric positive definite matrix $M(x)$, the change being minor if M is continuous, and a little more technical if M is only bounded measurable as one must then use the fact that m is bounded in $L^\infty(R^3)$ so the projection $Proj_1 \mu$ of the corresponding H-measures have a $L^\infty(\Omega)$ density with respect to the LEBESGUE measure dx, and one can therefore use operators of multiplication by functions of x with symbols in $L^\infty(\Omega)$.

For reasons that may seem purely mathematical, i.e. for testing the tools that I have successfully used on the preceding example, suppose now that one is interested in maximizing the functional J_0 instead of minimizing it.

This problem is actually not so academic, as a similar situation occurs when one wants to extend the preceding analysis to questions involving magnetostrictive effects, which also involve an elastic energy, but the albegraic manipulations in that case are not so clear, and it is therefore useful to analyze a similar question on a model analog to the one already used.

One first notices that J_0 is bounded above because m is bounded in $L^\infty(\Omega)$ and the mapping $m \mapsto grad(u)$ is linear continuous from $L^2(\Omega; R^3)$ into $L^2(R^3; R^3)$, although J_ε is not bounded above for $\varepsilon > 0$ as one can create oscillations in m and obtain arbitrarily large norms for the gradient of m. Taking a maximizing sequence and analyzing the situation as before, one finds a first relaxed problem where one must maximize the same functional J_1 already encountered, on the same set X of admissible pairs (ν, μ), which has not been characterized yet. The next question, however, is different because among all the H-measures μ that can be associated with a given YOUNG measure ν (set which we know to be convex and weak $\star$ closed), one must now maximize the quantity $\sum_{i,j=1}^{3} \langle \mu^{ij}, 1 \otimes \xi_i \xi_j \rangle$, and the situation is not as simple as when I was minimizing it because the minimization problem is completely solved once one has constructed an admissible μ for which that quantity is 0, while it is not clear a priori if the maximization problem will have one of its solutions inside the special subset of admissible μ that we already know how to construct. We are actually lucky: as μ is Hermitian nonnegative and $\xi \in S^{N-1}$, one has $\sum_{i,j=1}^{3} \langle \mu^{ij}, 1 \otimes \xi_i \xi_j \rangle \le \sum_{i=1}^{3} \langle \mu^{ii}, 1 \otimes 1 \rangle$, which is the limit of $\int_\Omega \sum_{i=1}^{3} |m_{\varepsilon,i} - m_{0,i}|^2\, dx$, which itself can be expressed in terms of the YOUNG measure ν, i.e. $\int_\Omega \big(\langle \nu_x, \|\cdot\|^2 \rangle - \|m_0\|^2 \big)\, dx$, and this bound can be obtained in the case already described where ν is a finite combination of DIRAC masses by taking for π_i the DIRAC mass at $\dfrac{V^{(i)} - V^G}{\|V^{(i)} - V^G\|}$ (if $V^{(i)} \neq V^G$, and any probability measure if $V^{(i)} = V^G$), for $i = 1, \ldots, r$. A similar choice of an optimal μ can be done for a more general ν, and this analysis leads to a second variational problem on the set Y of all YOUNG measures, where instead of minimizing the functional J_2 one maximizes the functional $\widetilde{J_2}$ defined as

$$\widetilde{J_2}(\nu) = \int_{R^3} \|grad(u^0)\|^2\, dx + \int_\Omega \Big(\langle \nu_x, \varphi + \|\cdot\|^2 \rangle - \|m_0\|^2 - (H_0.m^0) \Big)\, dx,$$

with m^0 and u^0 defined as before. The initial problem is imbedded into this new one and corresponds to each ν_x being a DIRAC mass. If Y is equipped with the weak $\star$ topology, then Y is compact and $\widetilde{J_2}$ is indeed upper semicontinuous and attains its maximum, but proving this property of upper semicontinuity for $\widetilde{J_2}$ without using H-measures is not as straightforward as proving the lower semicontinuity for J_2 which is convex. The map $m \mapsto \int_{R^3} \|grad(u)\|^2\, dx - \int_\Omega \|m\|^2\, dx$ is actually concave because it is quadratic and nonpositive as can be seen by FOURIER transform, using the inequality $\left| \sum_{i=1}^{3} \mathbf{F} m_i(\xi) \xi_i \right|^2 \le \sum_{i=1}^{3} |\mathbf{F} m_i(\xi)|^2$, which is of course a disguised analog of the same inequality used for H-measures, but if one is dealing with an heterogeneous material, this quick proof by FOURIER transform is not available because one has to localize in x and in ξ and that is exactly what the H-measures have been designed for (in that case the corresponding functional is a compact perturbation of a concave function).

Finally one defines Z to be the convex set of functions m^0 satisfying

$$|m^0(x)| \le 1 \text{ a.e. in } \Omega,$$

and one defines the functional $\widetilde{J_3}$ by

$$\widetilde{J_3}(m^0) = \int_{R^3} \|grad(u^0)\|^2\, dx + \int_\Omega \Big(\widetilde{\varphi}(m^0) - \|m_0\|^2 - (H_0.m^0) \Big)\, dx,$$

where u^0 is defined as before and where $\widetilde{\varphi}$ is the concave function defined on the unit ball by

$$\widetilde{\varphi}(m) = \sup_\nu \langle \nu, \varphi + \|\cdot\|^2 \rangle \text{ for all probability measures } \nu \text{ on } S^2 \text{ with center of mass } m.$$

The initial problem is imbedded into this new one and corresponds to m^0 taking its values on the unit sphere (a.e.). If one equips Z with the weak $\star$ topology, then Z is compact and $\widetilde{J_3}$ is upper semicontinuous and does attain its maximum. Notice that the choice of keeping separate the terms $\widetilde{\varphi}$ and $||\cdot||^2$ instead of giving a name to the function $\widetilde{\varphi} - ||\cdot||^2$ comes from the desire to show clearly what is the way to decompose the functional in order to prove directly the weak upper semicontinuity property.

6. H-measures and characteristic lengths

As ε tends to 0, the minimum I_ε of $J_\varepsilon(m)$ under the constraint $||m(x)|| = 1$ a.e. $x \in \Omega$ converges to I_0, the infimum of $J_0(m)$ under the same constraint. I_0 has been shown to be equal to the minimum of $J_3(m^0)$ for m^0 satisfying the constraint $||m^0(x)|| \leq 1$ a.e. $x \in \Omega$. The question is then to estimate the difference $I_\varepsilon - I_0$ in term of the characteristic length ε. Physical intuition tells that the width of the walls between magnetic domains is of the order of ε, but what could be a mathematical statement proving that there are indeed magnetic domains in the minimizers of J_ε? A mathematician cannot mistake the precise mathematical properties of a model with the experimental measurements of some physical effect, even when the introduction of this particular mathematical model has been motivated by this special effect. If one finds a mathematical way for asserting the existence of magnetic domains and thin walls between them (in this particular model with a small exchange energy), can one then account for the excess energy between the case $\varepsilon > 0$ and $\varepsilon = 0$, and how much will be located inside the walls and how much will be located elsewhere?

H-measures are obviously inadapted for solving this kind of questions, as they are defined without any use of characteristic lengths and perform an analysis in direction but not in frequency, adding the quadratic contributions corresponding to oscillations or concentration effects at various characteristic lengths: they were natural for the small amplitude Homogenization question for which I had introduced them (Homogenization problems do not need to have only one characteristic length as those who have only been able to understand the method of asymptotic expansion wrongly believe), and even for describing the propagation of energy in the limit of Geometrical Optics, which H-measures can also do, a characteristic length is not necessary as the phase only plays a secondary role when the frequency tends to infinity (but a primary one in order to study the corrections to Geometrical Optics like the Geometric Theory of Diffraction of Joe KELLER, which has not been completely analyzed from a mathematical point of view yet).

Patrick GÉRARD, who had introduced H-measures independently and called them microlocal defect measures ([18]), introduced a variant of H-measures using one characteristic length in their definition, which he called semiclassical measures ([19]). Localization in x is done exactly like for H-measures, but for what concerns localization in ξ, H-measures are defined by using the limits of quantities of the form

$$\int_{R^N} \mathbf{F}\Big((U_n)_i\Big)(\xi)\overline{\mathbf{F}\Big((U_n)_j\Big)(\xi)}\psi\Big(\frac{\xi}{||\xi||}\Big)\,d\xi,$$

with $\psi \in C(S^{N-1})$, while P. GÉRARD defined semiclassical measures by using the limits of quantities of the form

$$\int_{R^N} \mathbf{F}\Big((U_n)_i\Big)(\xi)\overline{\mathbf{F}\Big((U_n)_j\Big)(\xi)}\psi(\varepsilon_n\xi)\,d\xi,$$

with $\psi \in S(R^N)$. In that way H-measures are measures in $(x,\xi) \in (\overline{\Omega} \times S^{N-1})$ while semiclassical measures are measures in $(x,\xi) \in \overline{\Omega} \times R^N)$ (in order to avoid losing too much information about oscillations using characteristic lengths far too different from ε_n, I have extended the definition to functions ψ which behave like $\psi\Big(\frac{\xi}{||\xi||}\Big)$ near 0 and are bounded uniformly continuous away from 0, which corresponds to ξ belonging to a compactification of $R^N \setminus 0$ obtained by adding a sphere at 0 and something more subtle than a sphere at infinity).

The idea behind the definition of semiclassical measures is that if $U_n(x) = V\Big(\frac{x}{\varepsilon_n}\Big)$, with V periodic, then for $\varphi \in \mathcal{D}(\Omega)$ the FOURIER transform of φU_n is only important for $||\xi||$ of the order of $\frac{1}{\varepsilon_n}$, and one performs then a rescaling to look at the shape of $||\mathbf{F}(\varphi U_n)(\xi)||^2$. It is not clear however if semiclassical measures are adequate for discussing the problem of micromagnetics with a small exchange energy, at it seems to involve at least two characteristic lengths: the width of the walls, which is expected to be of the order of ε, and the characteristic size of the magnetic domains, for which I have not heard of any particular conjecture (most technological applications seem to be done with very small crystals, so that only one domain appears, and the question of estimating the size of magnetic domains for large monocrystals has probably not been considered important by physicists).

Semiclassical measures require the choice of a characteristic length ε_n and it is an important restriction that they can only work at one characteristic scale and may lose information about oscillations and concentration effects that use a smaller characteristic scale than ε_n (and the corresponding information is lost at infinity) or a larger scale than ε_n (and the corresponding information is accumulated at 0), and that is why I had worked out the extension already mentioned. As will be shown later on, some variants are needed in order to analyze situations where only two scales occur.

Contrary to what I once heard a physicist mumble, the tool of H-measures is not just the same set of ideas which are attributed to WIGNER, and it is quite surprising that Pierre-Louis LIONS & Thierry PAUL could have written that one can recover H-measures from semiclassical measures (which they renamed WIGNER measures) ([20]), a statement which is not only false as they could have noticed if they had tried to prove it (or if they had understood what P. GÉRARD had done before them), but stupid: when a mathematician has obviously read so many articles written by physicists that he insists in making his constructions obscure to almost all mathematicians by following lengthy considerations of no interest except to physicists instead of giving a quick mathematical proof of the main statement (like the one that I derived easily in a discussion with P. GÉRARD), how can one explain that he understands so little about Physics after all that reading that he could think of a world with only one characteristic length?

I had been told a long time ago by George PAPANICOLAOU about WIGNER transform, but it was only after discussing with P. GÉRARD and deriving a simplified proof of the main result obtained by P.-L. LIONS & T. PAUL that I understood the idea behind it. If YOUNG measures are but a mathematical way to discuss about one point statistics, H-measures do not exactly see the two-point correlation statistics but the part of it which is invariant by scaling, as was pointed out to me by Graeme MILTON. Actually one cannot even define correlations without using a characteristic length (except for the probabilistic schemes which I have decided to avoid): the intuitive idea is to use a characteristic length ε_n and to consider (for a subsequence) the weak $\star$ limits of quantities of the form $(U_n)_i(x+\varepsilon_n y)\overline{(U_n)_j(x+\varepsilon_n z)}$, which actually only depend upon $y-z$ so that one denotes them $C_{i,j}(x;y-z)$; one notices then that the FOURIER transform in y of $C(x;y)$ is a nonnegative Hermitian matrix μ of RADON measures in (x,ξ) by applying BOCHNER's theorem (and one must use the generalization by Laurent SCHWARTZ ([21]), of course, as what I describe here is just the idea and ideas must often be supplemented by classical technical details), and that μ is indeed the semiclassical/WIGNER measure associated to the subsequence. In discussion with P. GÉRARD, we have looked at a similar approach for n-point correlations, but we have not been able to identify what kind of microlocal object was hiding behind the formulas that we had proved, and we do not know in that situation what should replace FOURIER transform and BOCHNER's theorem.

In order to see the limitations of these different tools, and suggest that one might need an object seeing many different scales together with their interactions, it seems useful to replace the three-dimensional problem of micromagnetics that one would like to understand by a simpler one dimensional model.

For $u \in W_0^{1,4}(0,L)$, define $K_0(u)$ by the formula

$$K_0(u) = \int_0^L \left((1-|u'|^2)^2 + \frac{|u|^2}{a^2}\right) dx,$$

where $a > 0$ is a characteristic length. The problem of minimizing the functional K_0 is usually taken as an example for showing the need to introduce YOUNG measures, as the infimum of K_0 is equal to 0 but is not attained, and minimizing sequences u_n converge uniformly to 0 and have the property that u_n' converges weakly to 0 in $L^4(0,L)$ while $|u_n'|^2$ converges strongly to 1 in $L^2(0,L)$, so that u_n' defines the YOUNG measure $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. There is no small characteristic length in this problem and minimizing sequences may exhibit sawtooth patterns (as u_n mostly wants to have derivative ± 1), on many different scales. If for $\varepsilon > 0$ one considers the regularized functional K_ε defined on $W_0^{1,4}(0,L) \cap W^{2,2}(0,L)$ by

$$K_\varepsilon(u) = K_0(u) + \varepsilon^2 \int_0^L |u''|^2\,dx,$$

it was conjectured by J. M. BALL, as Nick OWEN had told me, that the minimizers would be periodic and therefore select a small length scale. I suppose that J. M. BALL had selected that question as a model for understanding why one observes regularly spaced twins in martensitic transitions of some crystals, while the model of minimization of energy following the ideas of J. ERICKSEN correctly predicts oscillations in some precise directions but cannot propose a length scale (as no small parameter appears in the formulation), and therefore it was natural to investigate the effect of a small surface energy term or of a penalization on higher order derivatives. The intuition tells that the characteristic scale to use in a transition layer where u' changes between ± 1 will be ε, but it was not obvious to me what the characteristic scale between two of these transitions would be; I had introduced the characteristic length a (which most mathematicians like to take equal to 1), and therefore there was a variety of candidates of the form $\varepsilon^\theta a^{1-\theta}$ for some $\theta \in (0,1)$ and I was surprised to discover after a simple but formal computation that the characteristic scale for the domains was $\varepsilon^{1/3}a^{2/3}$ and that the minimum of K_ε was of the order of $C\varepsilon^{2/3}a^{1/3}$ for a precise constant C. The conjecture was later proved by Stefan MÜLLER ([22]).

In the three-dimensional problem, one expects ε to be the characteristic scale for the walls but the characteristic scale for the domains is not clear, assuming that it does not even depend upon how far one is from the boundary $\partial\Omega$, as was pointed out to me by R JAMES. Should one use semiclassical/WIGNER measures with the characteristic length ε, or with the characteristic length conjectured for the size of the domains, or even some other scaling, or should one simply invent a different mathematical tool that would be more adapted to that precise purpose? I certainly would prefer to derive a mathematical method where one would not have to guess what scales appear in a problem and where it would be the method that would determine which scales do appear in the problem, but I am still searching for such a method.

My final point will be to present a simple computation done with P. GÉRARD: it shows that some of the preceding ideas are too crude, and that there might be a large class of variants that should be developed. Let $\varphi \in \mathcal{D}(R)$ and for any integer n define the function $u_n \in L^2(0,1)$ by

$$u_n(x) = n^{\lambda/2} \sum_{k \in Z} \varphi\Big(n^\lambda(nx - k)\Big),$$

with $\lambda > 0$. The sequence u_n converges weakly to 0, while $|u_n|^2$ converges weakly $\star$ to the constant function $C = \int_R |\varphi|^2\,dx$ (in the sense of RADON measures). As u_n is obtained by placing at distances $\frac{1}{n}$ rescaled versions of φ with the scaling factor $n^{1+\lambda}$, one is easily convinced that the sequence u_n exhibits the two scales $\frac{1}{n}$ and $\frac{1}{n^{1+\lambda}}$. However, if one computes the semiclassical/WIGNER measure associated to the sequence u_n when one uses the characteristic length $\varepsilon_n = \frac{1}{n}$, one has the surprise to find 0: the scale $\frac{1}{n}$ seems to be absent. Of course, there is some information lost at infinity because of the smaller scale $\frac{1}{n^{1+\lambda}}$, but if one computes the semiclassical/WIGNER measure associated to the sequence u_n when one uses the characteristic length $\varepsilon_n = \frac{1}{n^{1+\lambda}}$, one finds a nonzero contribution, but no information is crushed at 0, as one would expect for oscillations at much coarser scales like $\frac{1}{n}$: the scale $\frac{1}{n}$ still seems to be absent. We were indeed puzzled for a few minutes before discovering where the information was hidden, as the scale $\frac{1}{n}$ is indeed present but is not showing at the place where, quite naïvely, we expected to find it.

The intuition behind the definition of semiclassical measures is that a periodic pattern with period ε will show after FOURIER transform for $||\xi||$ of the order of $\frac{1}{\varepsilon}$, so that the FOURIER transform shows the scale $\frac{1}{\varepsilon}$; if one adds two functions which have periodic patterns of periods ε and η much smaller than ε, the

FOURIER transform does show some effects at two different places, one for $||\xi||$ of the order of $\frac{1}{\varepsilon}$ and the other for $||\xi||$ of the order of $\frac{1}{\eta}$, but if the two scales ε and η interact, as in our example, it may happen that the FOURIER transform only shows something happening for $||\xi||$ of the order of $\frac{1}{\eta}$, but what it shows there may contain oscillations on the missing scale $\frac{1}{\varepsilon}$.

Much before doing this computation, P. GÉRARD had proposed to introduce a variant of microlocal quadratic measures, where to symbols $a(\xi)$ and $b(x)$ one associates the operators $A_\varepsilon, B_\varepsilon$ given by

$$\mathbf{F}(A_\varepsilon u)(\xi) = a\Big(\frac{\xi}{\varepsilon}\Big)\mathbf{F}u(\xi),$$

and

$$(B_\varepsilon u)(x) = b\Big(\frac{x}{\varepsilon^{1/3}}\Big)u(x),$$

and for other reasons, I had advanced another idea which appeared quite similar.

7. Conclusion

Whatever it is that one knows, one must certainly go beyond it, but once one has succeeded in attaining a higher level of understanding, one must try to analyze what has been gained and explain it as much as one can in the simpler framework that is possible.

If one is interested in understanding the partial differential equations of Continuum Mechanics or Physics, one quickly finds that one must go beyond YOUNG measures as they can only handle constitutive relations, and that balance equations must be treated with another tool adapted to differential structures, and the Compensated Compactness method was such a complementary approach, improved recently by the introduction of H-measures. However, the information given by the Compensated Compactness method can be translated into properties of YOUNG measures, and I have shown an example where one needed to use both YOUNG measures and H-measures, but that after a slightly improved understanding of the relations between these two types of measures, one could then deduce the correct result, mentioning only YOUNG measures. Other tools must be developed to take into account a variety of different scales, and it might happen that in each case where such improved tools will permit to describe the solution of a problem, one will also deduce how to state a correct result in terms of the sole YOUNG measures.

In some way, the reason why after going beyond YOUNG measures one tries to state the results that one has obtained using only YOUNG measures, is linked to the traditional point of view in Continuum Mechanics: when dealing with some materials like fine mixtures, one does not try to use a mathematical object that can describe all possible microstructures and their evolution (object which has not been defined yet), but one tries to use a limited number of variables and one invokes some general principles to select a class of constitutive relations and balance equations that will describe the effective behaviour of these materials. In order to test the validity of some of these general principles (like Thermodynamics, which I hope to understand some day), one needs to develop better mathematical tools, which should show the limits of validity of these methods, but returning to YOUNG measures will always be necessary: after all, YOUNG measures is but a mathematical way of talking about statistics, without a probabilistic framework!

As a mathematician interested in improving the understanding of the partial differential equations of Continuum Mechanics and Physics, I am grateful to Jerry ERICKSEN for having shared his deep understanding of the structure of materials and expressed his results in a way that I could understand. I hope to be able in the future to improve the mathematical tools which are needed to go beyond the actual knowledge in some of the many directions for which he has opened the way.

Acknowledgements

The author's research has been supported by CARNEGIE-MELLON University, the National Science Fundation (grant DMS-9100834) and the Army Research Office through a grant to the Center for Nonlinear Analysis.

I want to thank for some fruitful discussions about the matters discussed here Patrick GÉRARD, Dick JAMES, François MURAT, Stefan MÜLLER and Antonio DE SIMONE. Some of the missing technical details in this text will hopefully appear either in some joint work, or in the lecture notes that I have been preparing ([23]).

References

[1] TARTAR L., "Compensated compactness and applications to partial differential equations," in *Nonlinear Analysis and Mechanics, HERIOT-WATT Symposium IV, ed. R. J. KNOPS, Research Notes in Mathematics* 39, pp. 136-212 (*London: Pitman*, 1979).

[2] TARTAR L., "H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations," Proc. Roy. Soc. Edinburgh, Sect. A. 115, (1990), 193-230.

[3] YOUNG L. C., *Lectures on the Calculus of Variation and Optimal Control Theory*, (Philadelphia: W.B. Saunders, 1969).

[4] BERLIOCCHI H. & LASRY J.-M., "Intégrandes normales et mesures paramétrées en calcul des variations," Bull. Soc. Math. France, 101, 1973, p. 129-184.

[5] BALL J. M., "Convexity conditions and existence theorems in nonlinear elasticity," Arch. Rational Mech. Anal., **63** (1977), 337-403.

[6] TARTAR L., "Nonlinear constitutive relations and homogenization," in *Contemporary developments in continuum mechanics and partial differential equations, (Proc. Internat. Sympos., Inst. Mat., Univ. Fed. Rio de Janeiro, Rio de Janeiro*, 1977), pp. 472-484. *North-Holland Math. Studies, 30, North-Holland, Amsterdam*, 1978.

[7] DIPERNA R. J., "Convergence of approximate solutions to conservation laws," Arch. Rational Mech. Anal., **82** (1983), 27-70.

[8] TARTAR L., "Estimations fines de coefficients homogénéisés," in *Ennio DE GIORGI Colloquium, ed. P. KRÉE, Research Notes in Mathematics* 125, pp. 168-187 (*London: Pitman*, 1985).

[9] FRANCFORT G., & MURAT F., "Homogenization and optimal bounds in linear elasticity," Arch. Rational Mech. Anal., **94** (1986), 307-334.

[10] JOSEPH D., *Stability of Fluid Motions, I, II*, Springer Tracts in Natural Philosophy, vol. 27, 28, Springer, 1976.

[11] MURAT F. & TARTAR L.C., "Calcul des variations et homogénéisation," *Les Méthodes de l'Homogénéisation: Théorie et Applications en Physique, Coll. de la Dir. des Etudes et Recherches de Elec. de France* 57, 319-369, *Eyrolles, Paris* 1985.

[12] BALL J. M., "A version of the fundamental theorem for YOUNG measures," in *P.D.E's and Continuum Models of Phase Transitions, Lect. Notes in Phys.* 344, *RASCLE M. & SERRE D. & SLEMROD M. eds.*, 207-215, *Springer*, 1989.

[13] DIPERNA R. J., & MAJDA A., "Oscillations and concentration in weak solutions of the incompressible fluid equations," Comm. Math. Phys., 108 (1987), 667-689.

[14] BROWN W. F., *Micromagnetics*, Interscience, New York, 1963.

[15] JAMES R. D. & KINDERLEHRER D., "Frustration in ferromagnetic materials," Cont. Mech. Therm. 2, (1990), 215-239.

[16] DE SIMONE A., Ph.D. thesis, U. of Minnesota, 1992.

[17] MURAT F., & TARTAR L., "On the relation between YOUNG measures and H-measures," in preparation.

[18] GÉRARD P., "Microlocal defect measures," Comm. Partial Differential Equations 16 (1991), no. 11, 1761-1794.

[19] GÉRARD P., "Mesures semi-classiques et ondes de BLOCH," in *Equations aux Dérivées Partielles, Exposé XVI, Séminaire* 1990-1991, *Ecole Polytechnique, Palaiseau.*

[20] LIONS P.-L. & PAUL T., "Sur les mesures de WIGNER," Revista Matemática Iberoamericana, **9** (1993), 261-270.

[21] SCHWARTZ L., *Théorie des distributions*, Hermann, Paris, 1966.

[22] MÜLLER S., "Mimimizing sequences for nonconvex functionals, phase transitions and singular perturbations," in *Problems involving change of type, KIRCHGÄSSNER K. ed., Lect. Notes in Phys.*, 359, pp. 31-44, Springer-Verlag, 1990.

[23] TARTAR L., *Homogenization, Compensated Compactness and H-Measures*, (CBMS-NSF Conference, Santa Cruz, June 1993), Lectures Notes in preparation.