Computation of Microstructure Utilizing Young Measure Representations

R. A. Nicolaides
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

and

Noel J. Walkington
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

Research Report No. 92-NA-001

January 1992

**Center for
Nonlinear Analysis**

**Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213-3890**

Sponsored by
U.S. Army Research Office
Research Triangle Park
North Carolina 27709

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

Computation of Microstructure Utilizing Young Measure Representations

R. A. Nicolaides
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

and

Noel J. Walkington
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

Research Report No. 92-NA-001

January 1992

Computation of Microstructure Utilizing Young Measure Representations*

R. A. Nicolaides† and Noel J. Walkington‡

Center for Nonlinear Analysis
Department of Mathematics, Carnegie Mellon University

Abstract

An algorithm is proposed for the solution of non–convex variational problems. In order to avoid representing highly oscillatory functions on a mesh, an associated Young measure, which characterizes such oscillations, is also approximated. Sample calculations demonstrate the viability of this approach.

keywords: Calculus of Variations, Young Measures.

1 Introduction

A recent development in continuum mechanics is the introduction of continuum energy functionals modeling nonlinear effects of crystal thermoelasticity [2, 6, 7, 8]. Among other things these functionals can be used to study displacive phase transformations and shape memory effects [9].

A characteristic feature of the energy functionals is their multiple well structure. Typically, each well represents a potential equilibrium state of the crystal, and at a transformation temperature more than one well is accessible to the crystal as a stable configuration.

The variational approach to finding an overall equilibrium state for the crystal requires that the energy functional be minimized in some suitable sense. In attempting such minimizations, one frequently encounters minimizing sequences of rapidly oscillating functions. These oscillations are usually a mathematical precursor to the formation of microstructure. This microstructure is characterized mathematically by probability distributions which, in principle, can be found by taking certain averages of the oscillatory functions.

In computational practice, the minimizing sequences are often constructed using a finite mesh, for example by finite elements. The oscillations referred to above then show up as grid scale oscillations of the (generally nonunique) minimizer. As the mesh is refined, the oscillations persist becoming more and more rapid while remaining of finite amplitude e.g. [4]. Usually, one

*To appear in Proceedings of the Conference on Recent Advances in Adaptive Sensory Materials and their Applications, Blacksburg VA, 4/92

†Supported by Airforce Grant AFOSR F49620-92-J-0133.

‡Supported by National Science Foundation Grant No. DMS-9002768.

wishes to know the values of macroscopic quantities associated with the deformation. These are computed in two different ways. Essentially, linear functions of the deformation can be obtained as the limits of the same linear functions of the minimizing sequence. On the other hand, nonlinear functions of the deformation (including energy) in general have to be computed as expected values of the probability distribution mentioned in the previous paragraph.

In some situations, the probabilities that are needed for computing nonlinear functions of the deformation are known a priori. In this paper we are interested in the opposite case. Although in principle it must be possible to compute the probabilities from the oscillatory minimizing sequence, in practice this could be very difficult if there were a relatively large number of wells. Also, it is easy to imagine that a rather fine mesh would be necessary to accumulate enough data to permit the evaluation of stable averages. We refer to this as the "microscopic" approach.

An alternative method is to compute with the probabilities as dependent variables. However, this is feasible only if we have some information about the limiting probability distributions which can occur. It turns out that frequently there is enough prior information to permit the computation to be done in that way. Our main goal is to investigate this alternative approach, which we will call the "macroscopic" methodology. Its potential advantage is that since the probabilities are smoothly varying quantities, a relatively coarse mesh can be used to approximate them. In this way we can avoid the need to deal with the oscillations explicitly. Nonlinear functions of the deformation (and linear functions as a special case) may be approximated using the computed probability distributions.

In the sections which follow, we state our algorithm and explain the ideas behind it. Then we present the results of some model computations. The work reported is of a preliminary nature. So far, we do not have sufficient experience with the algorithm to make rational comparisons with other approaches. It is hoped to address these issues in a future report.

2 The macroscopic formulation

We will begin with a simplified presentation of some background information in variational methods which is sufficient for understanding the principles behind the formulation of the algorithm. More detailed accounts can be found in [5, 12, 13].

We consider variational integrals

$$J(u) := \int_\Omega F(x, u, \nabla u)\, dx, \qquad u \in W_0^{1,p}(\Omega)^m, \tag{1}$$

where $\Omega \in R^n$ is a bounded domain and $F(\cdot, \cdot, \cdot)$ is continuous. Inhomogeneous boundary conditions can easily be accommodated if necessary. The case of most interest is when $F(x, u, \cdot)$ is not convex with respect to its last variable. The multiple well property of stored elastic energy functions causes this lack of convexity. In this case, the infimum of $J(\cdot)$ usually cannot be reached in $W_0^{1,p}(\Omega)^m$ and it is necessary to admit generalized solutions.

The standard example to illustrate this is

$$J(u) := \int_0^1 (u'^2 - 1)^2 + u^2\, dx, \qquad u \in W_0^{1,4}(0, 1).$$

The sequence $\{u_k\}$, whose first three members are illustrated in Figure 1b below, gives

$$J(u_k) = \int_0^1 u_k^2\, dx \to 0 = \inf_{W_0^{1,4}} J(u),$$

so that $\{u_k\}$ is a minimizing sequence. Also clear is that $\lim_{k\to\infty} u_k = u \equiv 0$. On the other hand, $u_k' \not\to 0$ in any ordinary sense, and so for this $\{u_k\}$, $\inf_{W_0^{1,4}} J(u)$ is not attained. It is the existence of the two wells at ± 1, illustrated in Figure 1a, which is responsible for the oscillatory behavior of the sequence $\{u_k'\}$.

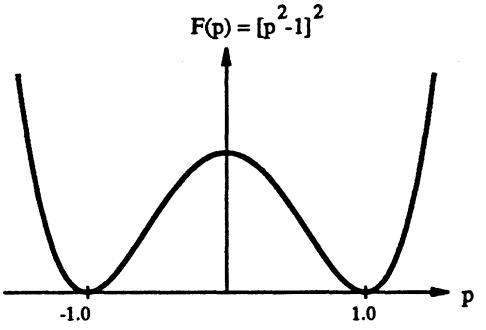


Figure 1a. Double well energy.

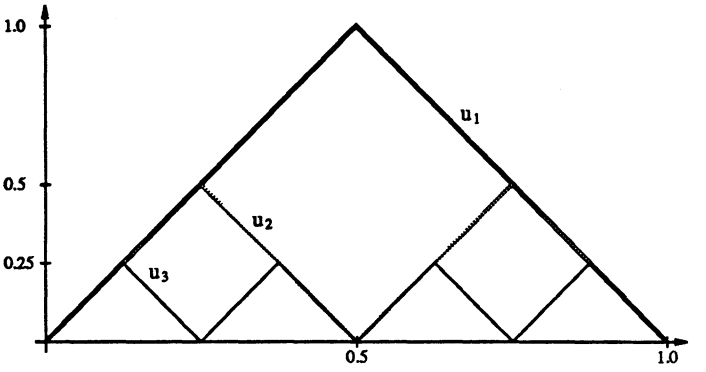


Figure 1b. A minimizing sequence.

The behavior of the sequence $\{u_k'\}$ strongly suggests that its "values" at any point $0 < x_0 < 1$ may be described by the probability distribution $u'(x_0) = \pm 1$ with probability $1/2$. In fact there is a general result from which this can be inferred: for any bounded sequence in $W^{1,p}(\Omega)^m$, $\| u_k \|_{1,p} \leq M$, $\{u_k\}$ contains a subsequence $\{u_{k_j}\}$ such that $u_{k_j} \to u \in L^p(\Omega)^m$. Additionally a subsequence of $\{u_{k_j}\}$ exists (denoted the same way) with the property that for any continuous g which is reasonably behaved at infinity, and for each $x \in \Omega$ there is a probability measure ν_x such that

$$g(\nabla u_{k_j}) \rightharpoonup G \in L^p(\Omega)^m, \tag{2}$$

where

$$G(x) = \int_{R^n} g(y)\, d\nu_x(y) \tag{3}$$

for almost all $x \in \Omega$. A useful version of this result, due to Kinderlehrer and Pedregal [10], states that if the sequence $\{u_k\}$ is a minimizing sequence for a variational problem having non–negative integrand with p-growth, then g may also have p-growth.

A family of probability measures $\{\nu_x\}$ obtained in this way is called a family of *gradient Young measures* [11]. Young measures also exist which are not gradient measures. They are derived in a similar way from bounded sequences in $L^p(\Omega)^m$.

There is a very useful characterization of Young measures ν_x due to Ball [1]. We will state this for gradient Young measures: let $\nu_{x,\delta}^k$ denote the probability distribution of the values of $\nabla u_k(z)$ as z is chosen uniformly at random from $B(x, \delta)$, the open ball with radius δ and center $x \in \Omega$. Then

$$\lim_{\delta\to 0} \lim_{k\to\infty} \left| \int_{R^n} g(y)\, d\nu_x(y) - \int_{R^n} g(y)\, d\nu_{x,\delta}^k(y) \right| \to 0.$$

This result reveals how it is the minimizing sequence that determines the probability distribution ν_x and provides a way to approximate it.

The result (2)-(3) does not give any information about the structure of the measure ν_x, and in particular whether it is discrete. General results on this do not appear to be available. Nevertheless, there is a large class of problems where it is expected on physical grounds that ν_x is indeed discrete. This class includes most, if not all, of the continuum functionals used so far

to model crystal energy. Since we want to make essential use of discreteness, we will introduce it as a hypothesis. Specifically, we will assume that

$$\nu_x = \sum_{l=1}^{L} \lambda_l(x)\delta_{A_l(x)}, \tag{4}$$

$$\sum_{l=1}^{L} \lambda_l(x) = 1, \qquad 0 \le \lambda_l \le 1, \tag{5}$$

where $\delta_{A_l(x)}$ denotes a Dirac mass with pole at $A_l(x)$ and $\lambda_l(x)$ varies measurably with x. References [3, 2, 8] contain examples satisfying the discreteness hypothesis.

Choosing g in (2) to be $F(x, u(x), \cdot)$ and denoting by $\{u_k\}$ a minimizing sequence bounded in $W^{1,p}(\Omega)^m$, we have

$$\lim_{j\to\infty} \int_\Omega F(x, u_{k_j}, \nabla u_{k_j})\, dx = \int_\Omega \langle \nu_x, F(x, u(x), \cdot)\rangle\, dx,$$

where $\langle \nu_x, \cdot\rangle$ denotes the action on the right side of (3). Additionally, choosing g in (2)-(3) to be the identity mapping shows that

$$\nabla u := \sum_{l=1}^{L} \lambda_l(x) A_l(x).$$

These results motivate the following generalized variational problem: minimize

$$I(u) := \int_\Omega \langle \nu_x, F(x, u(x), \cdot)\rangle\, dx, \qquad u \in W_0^{1,p}(\Omega)^m, \tag{6}$$

subject to

$$\nabla u(x) = \sum_{l=1}^{L} \lambda_l(x) A_l(x)$$

over suitable $A_l \in L^p(\Omega)^{mn}$, $\lambda_l \in L^\infty(\Omega)$, $l = 1, 2, \ldots, L$. Solutions to this problem are regarded as generalized solutions to (1). Notice that classical solutions to (1) may be recovered from the generalized formulation by taking, say, $\lambda_1 = 1$.

The variables in the generalized formulation are, in principle, slowly varying or macroscopic.

3 Numerical Algorithm

In this section we consider discretizations of the generalized problem. Basically, we use continuous piecewise linear approximations for u, and piecewise constant approximations for the A_l and λ_l, $1 \le l \le L$. However, it is important to note that the A_l cannot be always be arbitrarily chosen, since the combination on the right of (4) must be a gradient Young measure. We present a general way to handle this issue.

3.1 Computing the Constraints

The algorithm presented above involves several constraints, namely,

$$\nabla u = \sum_{l=1}^{L} \lambda_l A_l, \qquad \sum_{l=1}^{L} \lambda_l = 1, \quad \text{and} \quad 0 \le \lambda_l \le 1, \ 1 \le l \le L$$

(recall that the discrete u is piecewise linear, so its gradient is piecewise constant, as are the discrete A_l and λ_l). In addition to these obvious constraints, when u is vector valued further constraints on the representation of the gradient are required to guarantee that $\nu = \sum_{l=1}^{L} \lambda_l \delta_{A_l}$ is a *gradient* Young measure. The constraints on $\{\lambda_l\}_{l=1}^{L}$ are convex and trivially accommodated; however, the constraints associated with the gradient are not convex. Moreover, since imposing constraints can be computationally taxing, it is imperative to resolve them in an efficient manner. Below we outline an algorithm that effectively eliminates the constraints on ∇u analytically.

We begin by considering the case with $L = 2$, i.e.

$$\nabla u = \lambda A_0 + (1 - \lambda) A_1.$$

Letting $b = A_1 - A_0$, we may write

$$A_0 = \nabla u - (1 - \lambda) b, \qquad \text{and} \qquad A_1 = \nabla u + \lambda b.$$

In this situation,

$$\begin{aligned} \langle F(x, u, .), \nu \rangle &= \lambda F[x, u, A_0] + (1 - \lambda) F[x, u, A_1] \\ &= \lambda F[x, u, \nabla u - (1 - \lambda) b] + (1 - \lambda) F[x, u, \nabla u + \lambda b]. \end{aligned}$$

In the scalar case, $b \in R^n$ can be selected arbitrarily; however, when u is vector valued, $\nabla u \in R^{m \times n}$, and it is necessary and sufficient that $b = A_1 - A_0$ be a rank one matrix, $A_1 - A_0 = \mathbf{a} \otimes \mathbf{n}$, in order to obtain a gradient Young measure ($\mathbf{a} \in R^m$, $\mathbf{n} \in R^n$ may be chosen freely). i.e.

$$A_0 = \nabla u - (1 - \lambda) \mathbf{a} \otimes \mathbf{n}, \qquad A_1 = \nabla u + \lambda \mathbf{a} \otimes \mathbf{n},$$

$$\begin{aligned} \langle F(x, u, .), \nu \rangle &= \lambda F[x, u, A_0] + (1 - \lambda) F[x, u, A_1] \\ &= \lambda F[x, u, \nabla u - (1 - \lambda) \mathbf{a} \otimes \mathbf{n}] + (1 - \lambda) F[x, u, \nabla u + \lambda \mathbf{a} \otimes \mathbf{n}]. \end{aligned}$$

To obtain a representation of the gradient for arbitrarily large L, we repeat the construction as follows. Given A_0 and A_1 as above, write

$$A_0 = \lambda_0 A_{00} + (1 - \lambda_0) A_{01}, \qquad A_1 = \lambda_1 A_{10} + (1 - \lambda_1) A_{11},$$

where

$$A_{01} - A_{00} = b_0, \qquad \text{and} \qquad A_{11} - A_{10} = b_1,$$

if u is scalar valued, and

$$A_{01} - A_{00} = \mathbf{a}_0 \otimes \mathbf{n}_0, \qquad \text{and} \qquad A_{11} - A_{10} = \mathbf{a}_1 \otimes \mathbf{n}_1,$$

it u is vector valued. The representation for the gradient then becomes,

$$\nabla u = \lambda \lambda_0 A_{00} + \lambda (1 - \lambda_0) A_{01} + (1 - \lambda) \lambda_1 A_{10} + (1 - \lambda)(1 - \lambda_1) A_{11}.$$

The quantities A_{00} etc. are determined from b (or $\mathbf{a}$ and $\mathbf{n}$), λ, b_0 (or $\mathbf{a}_0$ and $\mathbf{n}_0$), λ_0, etc., for example,

$$A_{01} = \nabla u - (1-\lambda)b + \lambda_0 b_0$$

in the scalar case, and

$$A_{01} = \nabla u - (1-\lambda)\mathbf{a} \otimes \mathbf{n} + \lambda_0 \mathbf{a}_0 \otimes \mathbf{n}_0$$

in the vector case.

By repeating this process N times, we obtain admissible Young measures consisting of 2^N Dirac masses. This construction is conveniently represented with a binary tree as shown in Figure 2. Each matrix occurring in the representation of the measure corresponds to a leaf on the tree, and is uniquely identified by a binary word of length N.

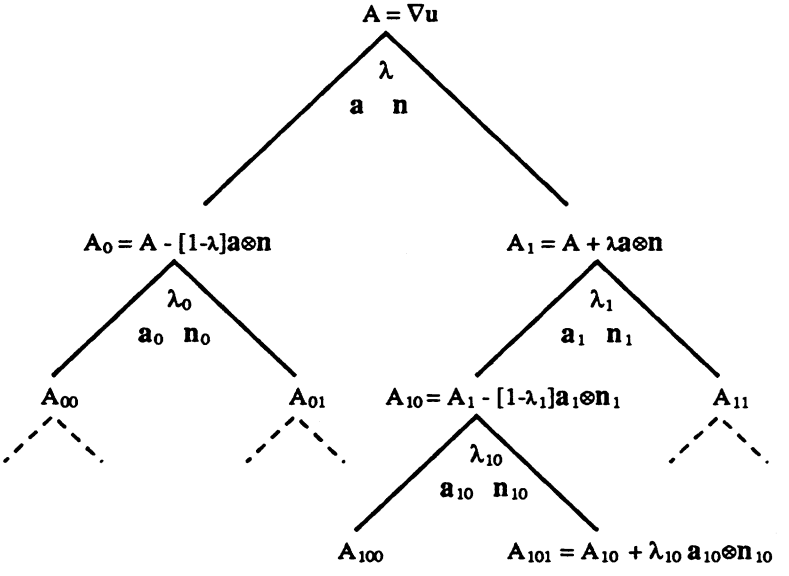


Figure 2. Binary tree representation of the micro structure

This representation of the gradient has the following desirable properties.

- If F is convex in its last variable, then trivially the minimum of $\langle F(x,u,.),\nu\rangle$ is attained with $b = b_0 \ldots \equiv 0$, $A_0 = A_1 = \ldots \equiv \nabla u$, and in this situation the problem reduces to the classical algorithm for approximating the solution of elliptic problems using piecewise linear functions.
- Given a guess for the minimizing function u, minimizing with respect to the piecewise constant functions b, λ b_0 etc. can be done in parallel over each element, suggesting the overhead associated with calculating a Young measure can be minimized by taking advantage of modern computer architectures.

4 Numerical Results

4.1 Computational Considerations

To obtain a solution of the discrete problem, simple relaxation was used in conjunction with the "numerical tricks" discussed below. The idea behind relaxation is to freeze all but one

unknown, ξ (a nodal value of u, or a λ value for an element, etc.), and to make one Newton iteration for the Euler equation $dI/d\xi = 0$, i.e. $\xi^{n+1} = \xi^n - I'(\xi^n)/I''(\xi^n)$. The following embellishments were required for a practical algorithm.

- Clearly it is necessary to restrict λ to lie in $[0,1]$. Moreover, since the algorithm degenerates when $\lambda = 0$, $\lambda = 1$, or $b = 0$, λ was required to satisfy $\epsilon \leq \lambda \leq 1 - \epsilon$ for some $\epsilon > 0$ (typically $\epsilon = 10^{-6}$ or 10^{-7}). Additionally, terms of the form $\epsilon(\lambda - 1/2)^2$ were added to the integrand to give a preferred value of $\lambda = 1/2$ when $b = 0$.

- Except for the Dirichlet data, initial values of $u = 0$, $b = 0$, and $\lambda = 1/2$ were chosen. It was observed that initially oscillations in u might develop before a suitable microstructure was found (this corresponds to computing a minimizing sequence directly). In order to suppress these oscillations while the microstructure developed, an "artificial viscosity" of the form $\mu(\Delta u)^2$ was added to the integrand. In all instances, μ was set to zero for the latter iterations.

- Since relaxation is a local algorithm, it is prone to "getting stuck" in local minima. It was frequently observed that microstructure would be present in one element, but not in an adjacent element. To remedy this problem, the micro–variables (λ, b etc.) were substituted for those in adjacent elements. If this lowered the energy, the modified microstructure was accepted. This non–local move was very effective for avoiding local minima.

This modified relaxation algorithm was found to be very effective for the computation of global minima. As with classical relaxation for the solution of elliptic problems, convergence was slow, especially in the latter iterates.

4.2 One Dimensional Examples

We consider one dimensional examples of the form,

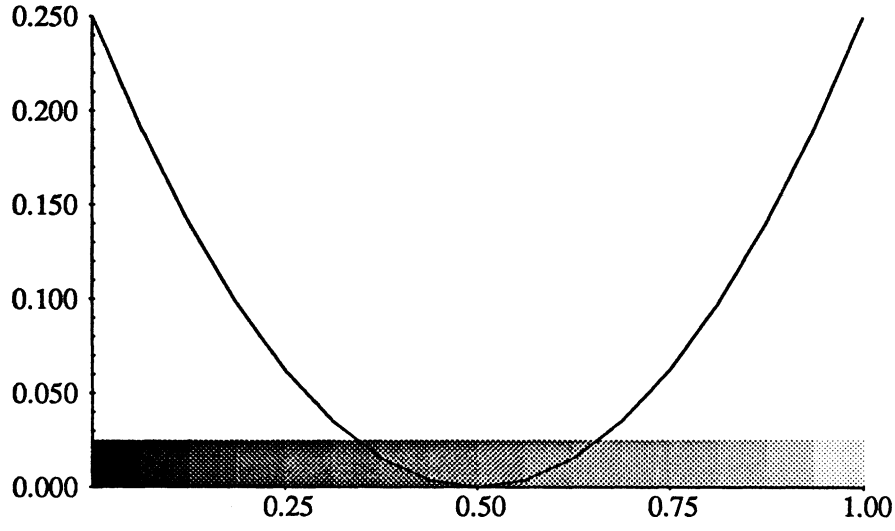
$$I(u) = \int_0^1 F(u') + (u - f)^2, \qquad u(0) = u_0, \;\; u(1) = u_1,$$

where $F(p) = (p^2 - 1)^2$ (see Figure 1) is the classical double well potential, and $f : [0,1] \to R$ is specified. For one dimensional problems, it suffices to consider only one level of the binary tree, i.e.

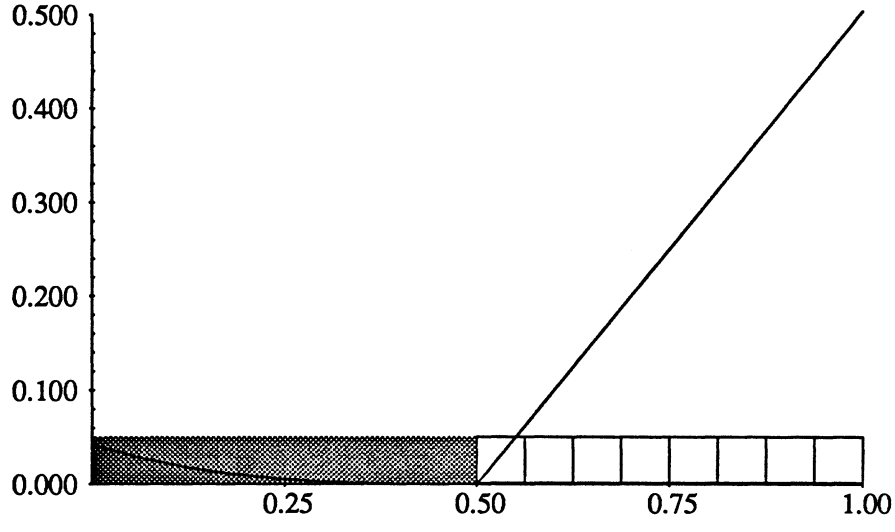
$$u' = \lambda A_0 + (1 - \lambda)A_1, \qquad A_0 = u' - (1 - \lambda)b, \qquad A_1 = u' + \lambda b.$$

4.2.1 Example 1 (Non–Homogeneous Young's Problem)

Setting $f(x) = (x - 1/2)^2$, the generalized solution of the variational problem is $u = f$, $\lambda = 1/2(1 + f')$, $b = 2$. For problems of this type (i.e. $-1 \leq f' \leq 1$), it is possible to show that the discrete solutions $\{u_h\}_{h>0}$ converge to u in $W^{1,p}(0,1)$ at the optimal rate of h. Similarly, $\lambda_h \to \lambda$ in $L^p(0,1)$ at optimal rate h. This is exhibited in Figure 3 where the $L^2(0,1)$ and $H^1(0,1)$ errors for u_h are tabulated. The solution obtained with a 16 element mesh is shown in Figure 4a.



(a) Example 1.



(b) Example 2.

Figure 4. One Dimensional Exapmles (16 Elements).
Black = well at -1, White = well at +1

No. Elements	Example 1		Example 2	
	$\|u-u_h\|_{L^2(0,1)}$	$\|u'-u_h'\|_{L^2(0,1)}$	$\|u-u_h\|_{L^2(0,1)}$	$\|u'-u_h'\|_{L^2(0,1)}$
4	0.010423	0.144338	0.002352	0.029793
8	0.002604	0.072121	0.000582	0.015258
16	0.000654	0.036115	0.000145	0.007673
32	0.000176	0.018454	0.000037	0.003867
$\|u\|_{L^2(0,1)}$ or $\|u'\|_{L^2(0,1)}$	0.111803	0.577350	0.205711	0.719047

Figure 3: Error Norms for One Dimensional Examples

4.2.2 Example 2

We consider a second less trivial example involving a "broken" extremal[1]. The nonhomogeneous term is,

$$f(x) = -3/128(x-1/2)^5 - 1/3(x-1/2)^3,$$

and the solution, given by

$$u(x) = \begin{cases} f(x), & 0 \le x \le 1/2, \\ 1/24(x-1/2)^3 + (x-1/2), & 1/2 \le x \le 1, \end{cases}$$

has microstructure in $(0,1/2)$ and is "elliptic" on $(1/2,1)$. On $(0,1/2)$, $\lambda = 1/2(1+u')$ and $b = 2$. Note that the derivative of u jumps from zero to one at $x = 1/2$. Figure 3 exhibits the optimal rates of convergence observed for $\{u_h\}_{h>0}$ in $L^2(0,1)$ and $H^1(0,1)$. The solution for a 16 element mesh is shown in Figure 4b.

4.3 Two Dimensional Example

We consider examples of the form

$$I(u) = \int_\Omega |\nabla u - \mathbf{w}_1|^2 |\nabla u - \mathbf{w}_2|^2,$$

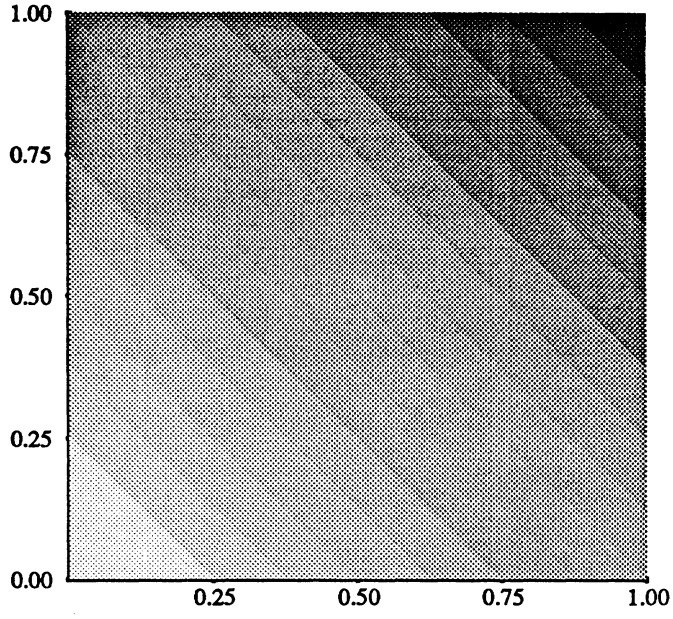
where $\mathbf{w}_1$, $\mathbf{w}_2 \in R^2$ are the locations of the energy wells. We chose $\Omega = (0,1)^2$ to be the unit square, and impose Dirichlet boundary conditions on u. Triangular meshes are constructed by dividing the region into similar squares, and dividing them in two along the diagonal with slope -1. We consider examples where the slope of u lies on the line joining $\mathbf{w}_1$ and $\mathbf{w}_2$, so that the micro–structure can be represented by a gray scale with $\mathbf{w}_1$ colored black and $\mathbf{w}_2$ colored white.

We present two examples, one being obtained from the other by a rotation of 90^0. This illustrates what happens when the mesh is most favorably and least favorably aligned with the contours of the exact solution, u. In Figure 5a, a solution with $\mathbf{w}_1 = (-1,-1)$ and $\mathbf{w}_2 = (1,1)$ is shown having exact solution

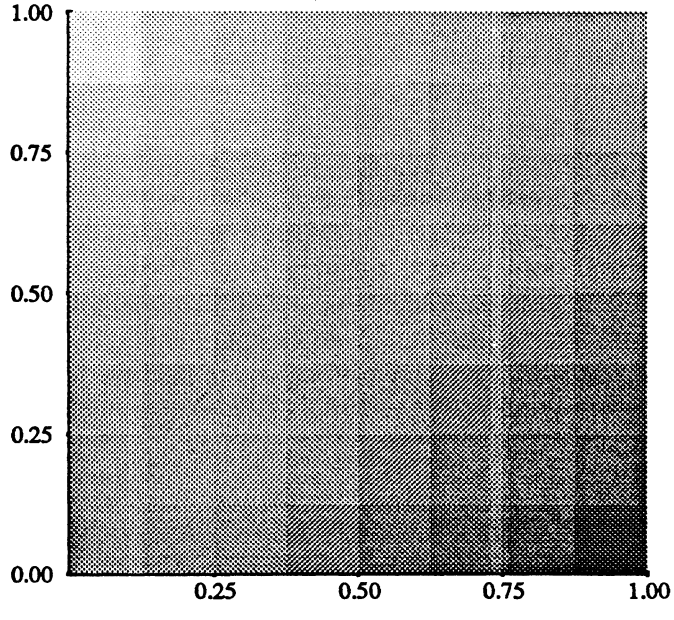
$$u(x,y) = \frac{2}{e^2-1} e^{(x+y)} - \frac{e^2+1}{e^2-1}(x+y),$$

$$\nabla u = \lambda \mathbf{w}_1 + (1-\lambda)\mathbf{w}_2, \qquad \lambda(x,y) = (e^2 - e^{x+y})/(e^2-1).$$

[1]This solution was suggested by Luc Tartar.



(a) Black = (1,1) well, White = (-1,-1) well.



(b) Black = (1,-1) well, White = (-1,1) well.

Figure 5. Two Dimensional Example, 8 x 8 Mesh.

The solution corresponding to a 90^0 rotation is shown in Figure 5b. Here $\mathbf{w}_1 = (1,-1)$, $\mathbf{w}_2 = (-1,1)$ with exact solution

$$u(x,y) = \frac{2e}{e^2-1}e^{(x-y)} - \frac{e^2+1}{e^2-1}(x-y),$$

$$\nabla u = \lambda \mathbf{w}_1 + (1-\lambda)\mathbf{w}_2, \qquad \lambda(x,y) = (e^{1+x-y}-1)/(e^2-1).$$

5 Concluding Remarks

In conclusion, the computations show that the overall approach is a useful one, and that it does produce optimal rates of convergence under mesh refinement. Certainly, more work must be done to implement the vector case, and also to improve the performance of algebraic solvers. It is hoped to address these matters in the future.

Acknowledgment: We thank David Kinderlehrer for many valuable conversations and suggestions.

References

[1] J. M. Ball. A version of the fundamental theorem for young measures. In D. Serre, editor, *Partial Differential Equations and Continuum Models of Phase Transitions.* Springer, May 1988.

[2] J. M. Ball and R. D. James. Fine phase mixtures as minimizers of energy. *Archive for Rational Mechanics and Analysis*, 100:13–52, 1987.

[3] K. Bhattacharya. Wedge–like microstructures in martensites. *Acta Metall. Mater.*, 39(10):2431–2444, 1991.

[4] C. Collins and M. Luskin. The computation of the austenitic-martinsitic phase transition. In M. Rascle, D. Serre, and M. Slemrod, editors, *Partial Differential Equations and Continuum Models of Phase Transitions, Lecture Notes in Physics 344*, pages 34–50. Springer Verlag, 1989.

[5] B. Dacorogna. *Direct Methods in the Calculus of Variations.* Springer Verlag, 1989.

[6] J. L. Ericksen. Stable equilibrium configurations of elastic crystals. *Archive for Rational Mechanics and Analysis*, 94:1–14, 1986.

[7] J. L. Ericksen. Some constrained elastic crystals. In J. M. Ball, editor, *Material Instabilities in Continuum Mechanics and Related Mathematical Problems*, pages 119–135. Oxford University Press, May 1988.

[8] R. D. James and D. Kinderlehrer. Theory of diffusionless phase transitions. In M. Rascle, D. Serre, and M. Slemrod, editors, *Partial Differential Equations and Continuum Models of Phase Transitions, Lecture Notes in Physics 344*, pages 51–84. Springer Verlag, 1989.

[9] R. D. James and D. Kinderlehrer. Frustration in ferromagnetic materials. *Continuum Mechanics and Thermodynamics*, 2:215–239, 1990.

[10] D. Kinderlehrer and P. Pedregal. Weak convergence of integrands and the young measure representation. Technical Report 90-87-NAMS-3, Carnegie Mellon University, Aug. 1990.

[11] D. Kinderlehrer and P. Pedregal. Characterization of young measures generated by gradients. *Archive for Rational Mechanics and Analysis*, Preprint.

[12] C. B. Morrey. *Multiple Integrals in the Calculus of Variations.* Springer, 1966.

[13] L. C. Young. *Lectures on the Calculus of Variations and Optimal Control.* Chelsa, 1980.