Numer. Math. (2000) 85: 647–663
Digital Object Identifier (DOI) 10.1007/s002110000150

Numerische Mathematik

Estimates for numerical approximations of rank one convex envelopes*

G. Dolzmann[1], **N.J. Walkington**[2]

[1] Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22–26, 04103 Leipzig, Germany; e-mail: georg@mis.mpg.de
[2] Department of Mathematical Sciences, Carnegie Mellon University Pittsburgh, PA 15213, USA; e-mail: noelw@andrew.cmu.edu

Received February 1, 1999 / Published online April 20, 2000 –

Summary. We present a convergence analysis of an algorithm for the numerical computation of the rank-one convex envelope of a function $f : \mathbb{M}^{m\times n} \to \mathbb{R}$. A rate of convergence for the scheme is established, and numerical experiments are presented to illustrate the analytical results and applications of the algorithm.

1. Introduction

In recent years, many surprising mechanical properties of crystalline materials, like the shape memory effect, have been successfully explained by analyzing microstructures which are created during a solid to solid phase transition of the material. In their fundamental paper [2] Ball and James developed a mathematical theory in the framework of nonlinear elasticity in which the experimentally observed geometries arise naturally as minimizers (or almost energy minimizing configurations) of a non–convex free energy functional. More precisely, one is led to the variational problem: Minimize

$$I(u) = \int_\Omega f(Du)dx$$

where $u : \Omega \to \mathbb{R}^n$ denotes the elastic deformation of the body $\Omega \subset \mathbb{R}^n$ in its reference configuration and f is the (nonlinear) energy density which we

* Supported in part by National Science Foundation Grant No. DMS–9512142. This work was also supported by the Army Research Office and NSF through the Center for Nonlinear Analysis.
Correspondence to: N.J. Walkington

assume to be non-negative and to depend only on the deformation gradient $Du \in \mathrm{M}^{n\times n}$. The zero set of f has typically a multi-well structure,

$$\{X : f(X) = 0\} = \bigcup_{i=1}^{I} \mathrm{SO}(n)U_i,$$

where the matrices U_i denote the preferred deformation for the i-th phase and U_i satisfies $U_i = U_i^{\mathrm{T}}$ and $\det U_i > 0$. Finding minimizers is a non-trivial problem, since the integrand fails to be quasiconvex and hence I is not weakly lower semicontinuous (see, e.g. [9,22] for further information). Minimizing sequences develop oscillations, so converge weakly, and not strongly, to a deformation which is not energy minimizing. However, the infimum of the energy, the *relaxed* or *effective* energy, can be computed as the minimum of the corresponding relaxed functional

$$I^{\mathrm{qc}}(u) = \int_{\Omega} f^{\mathrm{qc}}(Du)dx, \tag{1.1}$$

where f^{qc} is the quasiconvex [9] envelope of f, i.e., the largest quasiconvex minorant of f. We recall that $\inf I^{\mathrm{qc}} = \inf I$. Generally it is impossible to compute f^{qc} explicitly (or even numerically) and therefore it is natural to ask for lower and upper bounds on the effective energy. Here we focus on replacing f^{qc} by f^{rc}, the rank–one [9] convex envelope of f, which is defined to be the largest rank–one convex minorant of f. Since every quasiconvex function is rank–one convex (the converse is not true, see [26]), we conclude $f \geq f^{\mathrm{rc}} \geq f^{\mathrm{qc}}$ and therefore $\inf I^{\mathrm{rc}} = \inf I^{\mathrm{qc}}$. However, minimizing I^{rc} can be as ill-posed as the original problem, unless f^{qc} and f^{rc} coincide. On the other hand, for all functions, for which f^{qc} and f^{rc} are known explicitly, these two envelopes do coincide [16,17] and thus the computation of I^{rc} is certainly of interest, and this is the focus of this paper.

During the past decade a variety of numerical techniques have been proposed for various non–convex problems where oscillations play an important role, and the survey paper by Luskin [19] gives an extensive overview of the progress made to date. Much of the numerical work on non–convex variational problems has centered around the computation of f^{qc} at a single point $F \in \mathrm{M}^{m\times n}$, [5–8,11,12]. Typically this is attempted by numerically estimating the minima of $I(.)$ subject to affine boundary conditions. Luskin et. al. and Carstensen et. al. [4,15,18,20] have also considered a variety of other problems. The theory of Young measures [14,27] has been developed to characterize the minimizing sequences of $I(.)$ and their weak limits [1,13, 24,27], and Nicolaides and Walkington [23] developed numerical schemes which exploited this connection. In spite of all of this activity, there are no robust and efficient algorithms for the approximation of non–convex variational problems. In this paper we consider an alternative approach. Instead of using minima of $I(.)$ to characterize the relaxed energy, the procedure considered here exploits the convexity properties mentioned in the previous paragraph. This approach is expensive in the sense that it necessitates computations on a mesh in high dimensions ($m \times n$ for functions $f : \mathbb{R}^m \to \mathbb{R}^n$); however, this one computation provides all of the information required to solve many variational problems of the form (1.1) with a variety of boundary conditions and various low order terms etc.

In the next Sections we recall the definition and various characterizations of rank–one convex functions, and the algorithm proposed in [10] for the computation of rank–one envelopes. Section 4 establishes a rate of convergence for this algorithm and in Sect. 5 numerical examples are presented and compared with the results in Sect. 4. We also consider examples where the discrete rank–one convexification is used to estimate (relaxed) deformations that minimize integrands of the form (1.1).

We finish this introductory section with a brief overview of the notation utilized below. If $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$, then $|a|$ and $|b|$ will denote their Euclidean norm. $\mathrm{M}^{m\times n}$ denotes the space of real $m \times n$ matrices, for $F \in \mathrm{M}^{m\times n}$, $|F|$ denotes the norm induced from the Euclidean norms on $\mathbb{R}^m$ and $\mathbb{R}^n$. The tensor product of vectors $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$, denoted $a \otimes b \in \mathrm{M}^{m\times n}$, is the matrix having entries $(a \otimes b)_{ij} = a_i b_j$; observe that $|a \otimes b| = |a|\,|b|$. The maximum norm on finite dimensional spaces will be indicated by ℓ^∞ and the Lipschitz constant of a real valued function f is denoted by $|f|_{\mathrm{Lip}}$. Greek letters λ and μ will be used to indicate real numbers in the interval $[0, 1]$.

2. Rank–one convex envelopes

Recall that a function $f : \mathrm{M}^{m\times n} \to \mathbb{R}$ is said to be rank-one convex if $f(\lambda F_1 + (1-\lambda)F_2) \le \lambda f(F_1) + (1-\lambda) f(F_2)$ for all matrices $F_1, F_2 \in \mathrm{M}^{m\times n}$ with $\mathrm{rank}(F_1 - F_2) = 1$ and all $\lambda \in [0,1]$. If $f \in C^2$, then this is equivalent to $D^2 f(F; R, R) \ge 0$ for all $F, R \in \mathrm{M}^{m\times n}$ with $\mathrm{rank}(R) = 1$.

There are two representations for the rank-one convex envelope f^{rc} of a given function f. The first formula is analogous to the inequality $f(\sum_{i=1}^n \lambda_i F_i) \le \sum_{i=1}^n \lambda_i f(F_i)$ for convex functions f where $\lambda_i \in [0,1]$ and $\sum_{i=1}^n \lambda_i = 1$ (see Dacorogna [9]).

Definition 2.1. The sequence of pairs $(\lambda_i, F_i) \in (0,1) \times \mathrm{M}^{m\times n}$, $i = 1 \dots N$, satisfies the condition $\mathcal{H}_N$ (we write $(\lambda_i, F_i) \in \mathcal{H}_N$) if the following holds: $\sum_{i=1}^N \lambda_i = 1$ and

i) if $N = 2$ then $\mathrm{rank}(F_1 - F_2) = 1$,

ii) if $N > 2$, then there exists $k \in \{1, \ldots, N-1\}$ such that, up to relabeling the matrices,

$$\left(\frac{\lambda_i}{\lambda_1 + \ldots + \lambda_k}, F_i\right)_{1,\ldots,k} \in \mathcal{H}_k, \quad \left(\frac{\lambda_i}{\lambda_{k+1} + \ldots + \lambda_N}, F_i\right)_{k+1,\ldots,N} \in \mathcal{H}_{N-k},$$

and if

$$\mu_1 = \lambda_1 + \ldots + \lambda_k, \quad \mu_2 = \lambda_{k+1} + \ldots + \lambda_N,$$

and

$$H_1 = \frac{\lambda_1 F_1 + \ldots + \lambda_k F_k}{\lambda_1 + \ldots + \lambda_k}, \quad H_2 = \frac{\lambda_{k+1} F_{k+1} + \ldots + \lambda_N F_N}{\lambda_{k+1} + \ldots + \lambda_N},$$

then $(\mu_i, H_i) \in \mathcal{H}_2$.

(Note that there are no additional conditions if $N = 1$).

Then

$$f^{\rm rc}(F) = \inf\Big\{\sum_{i=1}^{N} \lambda_i f(F_i) : \,(\lambda_i, F_i) \in \mathcal{H}_N,\, F = \sum_{i=1}^{N} \lambda_i F_i\Big\}. \tag{2.1}$$

Here the infimum is taken over all $N \in \mathbb{N}$ and all $(\lambda_i, F_i) \in \mathcal{H}_N$. Note that it is not possible to bound N in the definition of $f^{\rm rc}$.

The second representation has a more algorithmic flavor and was obtained in [17]. Let $f_0 = f$ and define iteratively

$$f_{k+1}(F) = \inf\Big\{\lambda f_k(F_1) + (1-\lambda) f_k(F_2) : \, F = \lambda F_1 + (1-\lambda) F_2, \\ \mathrm{rank}(F_1 - F_2) = 1\Big\}.$$

Then $f^{\rm rc}(F) = \lim_{k\to\infty} f_k(F)$. Our analysis takes advantage of both representations: while the algorithm is based on the second representation, the convergence analysis relies on the first representation.

3. The algorithm

Let $\mathcal{G}_h = h\mathbb{Z}^{m\times n}$ be a uniform grid in $\mathrm{M}^{m\times n}$ and denote by $\mathcal{G}_{h,r} = \mathcal{G}_h \cap B(0,r)$ the intersection of the grid with the ball $B(0,r)$. Choose a subset $\mathcal{R}_h \subset \{hR : \, R \in \mathbb{Z}^{m\times n}, \, \mathrm{rank}(R) = 1\}$ of rank-one matrices in $\mathcal{G}_h$. The algorithm is now defined in the flow chart in Fig. 2.1. We use the notation $\{f_i^h\} = \mathcal{A}(f, \mathcal{G}_{h,r})$ to denote the sequence of functions generated in the algorithm (with $EPS = 0$).

Remarks. 1). Assume that $\mathcal{D} \subset \mathrm{M}^{m\times n}$. We say that a function $f_h : \mathcal{G}_h \to \mathbb{R}$ is $\mathcal{D}$-convex if $\tilde{f}_h$, the restriction of f_h to $\{F + tR\} \cap \mathcal{G}_h$ is convex as a

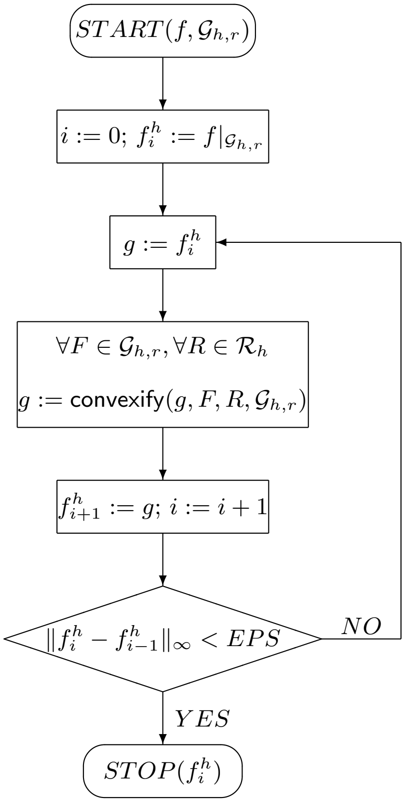

Fig. 2.1 The algorithm for the computation of the rank-one convex envelope

function of one variable for all $F \in \mathcal{G}_h$ and $R \in \mathcal{D}$ (see [21] for a discussion of general properties of $\mathcal{D}$-convex functions). The function $\tilde{f}_h$, which we may consider to be defined on $\mathbb{Z}$, is convex if $\tilde{f}_h(i-1) - 2\tilde{f}_h(i) + \tilde{f}_h(i+1) \geq 0$ for all $i \in \mathbb{Z}$. This is equivalent to asking that the piecewise affine interpolation of f_h is convex as a function on $\mathbb{R}$. Rank-one convexity corresponds to $\mathcal{D} = \{R \in \mathrm{M}^{m \times n} : \operatorname{rank}(R) = 1\}$.

2). In the algorithm we use a procedure convexify $(g, F, R, \mathcal{G}_{h,r})$ which computes the one dimensional convexification of g restricted to $\ell_h = \{F + tR\} \cap \mathcal{G}_{h,r}$. Let us call this restriction g_h. In the special situation at hand the points are sorted, so the convex envelope of g_h can be obtained with $\mathcal{O}(k)$ operations where k is the number of points on ℓ_h.

3). An interesting (and open) question is how to order the one-dimensional convexifications to achieve rapid convergence. Some examples seem to indicate that it is crucial to use a random order to propagate the information obtained in one step faster through the grid (see Sect. 5).

4. Convergence of the algorithm

The following convergence result was established in [10]. Recall that $\mathcal{R}_h$ is a subset of the rank one matrices on the grid $\mathcal{G}_h = h\mathbb{Z}^{m\times n}$ and that a function $f : \mathrm{M}^{m\times n} \to \mathbb{R}$ is $\mathcal{R}_h$-convex if its restriction to $\mathcal{G}_h$ is convex along all directions in $\mathcal{R}_h$.

Theorem 4.1. *Assume that f is Lipschitz continuous and that there exists a rank-one convex function $g^{\mathrm{rc}} : \mathrm{M}^{m\times n} \to \mathbb{R}$ such that $f \geq g^{\mathrm{rc}}$ on $B(0,r)$ and $f = g^{\mathrm{rc}}$ on $\mathrm{M}^{m\times n} \setminus B(0,r)$. Then there exists an $\mathcal{R}_h$-convex function $f^h : \mathcal{G}_h \to \mathbb{R}$ such that the functions $\{f_i^h\} = \mathcal{A}(f, \mathcal{G}_{h,2r})$ defined in the algorithm in Fig. 2.1 converge to f^h.*

Remark. It is necessary to compute the rank-one convexification on a domain which contains $B(0,r)$ as a proper subset; otherwise, the computed function could be $\mathcal{R}_h$-convex on $B(0,r)$ but fail to be $\mathcal{R}_h$-convex on all $\mathrm{M}^{m\times n}$.

If the rank-one convex envelope of f can be computed with N bounded from above by some constant in formula (2.1), then we have the following quantitative estimate for the error $f - f^h$:

Theorem 4.2. *Assume that the hypotheses in Theorem 4.1 are satisfied and that f^{rc} can be computed by formula (2.1) with $N \leq N_0$. Suppose that*

$$\mathcal{R}_h = \{h(a \otimes b) : \; a \in \mathbb{Z}^m, \, b \in \mathbb{Z}^n, |a|_{\ell^\infty}, |b|_{\ell^\infty} < h^{-1/3}\}.$$

Then there exists a constant C which depends only on m, n and r (the radius appearing in Theorem 4.1) such that

$$\|f^{\mathrm{rc}} - f^h\|_{L^\infty(\mathcal{G}_h)} \leq C|f|_{\mathrm{Lip}} h^{1/3}.$$

The proof of the theorem is based on the following approximation lemma which is interesting in its own right. While the explicit formulae given for the constants appearing in the lemma are not of particular interest, we state them here for ease of exposition in the proof given below.

Lemma 4.3. *Assume that $0 < h \leq 1$ and $(\lambda_i, F_i) \in \mathcal{H}_N$ with $F = \sum_{i=1}^N \lambda_i F_i$. Suppose that $F^h \in \mathcal{G}_h$ satisfies $|F - F^h| \leq c_0 h^{1/3}$. Then there exist $(\lambda_i^h, F_i^h) \in \mathcal{H}_N$ and a constant c_1 which depends only on m, n, and $\max_i |F_i|$ such that*

i) $F_i^h \in \mathcal{G}_h$ *and* $F^h = \sum_{i=1}^N \lambda_i^h F_i^h$*;*
ii) $|F_i^h - F_i| \leq (c_0 + (N-1)c_1)h^{1/3}$ *for* $i = 1, \ldots, N$*;*
iii) we have the estimate

$$\Big| \sum_{i=1}^N (\lambda_i f(F_i) - \lambda_i^h f(F_i^h)) \Big| \leq (c_0 + (N+1)c_1) \, |f|_{\mathrm{Lip}} h^{1/3}.$$

Given this approximation lemma, Theorem 4.2, follows almost immediately.

Proof (of Theorem 4.2). Fix $F \in \mathcal{G}_h$. By hypothesis there exits $(\lambda_i, F_i) \in \mathcal{H}_N$ such that $f^{\text{rc}} = \sum_{i=0}^N \lambda_i f(F_i)$. Next let $F^h \in \mathcal{G}_h$ be a closest point to F, so that $|F - F^h| \le \sqrt{mnh}/2$, and observe that

$$f^{\text{rc}}(F) - f^h(F) = f^{\text{rc}}(F) - f^h(F^h) \le \sum_{i=0}^N \lambda_i f(F_i) - \lambda_i^h f(F_i^h)$$

where $(\lambda_i^h, F_i^h) \in \mathcal{H}_N$ is the sequence guaranteed by Lemma 4.3 and we have used the fact that $f^h(F^h)$ is the minimum over all sequences $(\lambda_i^h, F_i^h) \in \mathcal{H}_N$ in $\mathcal{G}_h$. □

To establish Lemma 4.3 we will first develop some elementary approximation properties of vectors and matrices on equi–spaced lattices. Given a lattice of points with spacing $\sqrt{h}$ and a vector $a \in \mathbb{R}^m$, in order to find a vector $a^h \in \sqrt{h}\mathbb{Z}^m$ whose angle from a is less than ϵ it will, in general, be necessary for a^h to have length of order $\sqrt{h}/\epsilon$. If we then approximate a by an integer multiple of a^h, $a \sim ka^h$, then the error in the magnitude is of order $\sqrt{h}/\epsilon$. This elementary argument illustrates that there is a trade off between the accuracy of the angle and the accuracy of the length, and shows that the following lemma (where we put $\epsilon \sim h^{1/3}$) is sharp.

Lemma 4.4. *Let* $0 < h \le 1$, $0 \ne a \in \mathbb{R}^m$ *and for integer* $d > 0$ *let* $\mathbb{Z}_d = \{i \in \mathbb{Z} : |i| \le d\}$. *Then there exists* $d > 0$ *and* $a^h \in \sqrt{h}\,\mathbb{Z}_d^m$ *such that*

– $d \le h^{-1/3}$,
– $|a^h| \le \sqrt{m}\, h^{1/6}$,
– *The angle between the vectors* a^h *and* a *is bounded by* $(\pi/2)\sqrt{m}\, h^{1/3}$.

Proof. Let $d > 0$ be an integer and define the box $B = \{\sqrt{h}x | x \in \mathbb{R}^m,\ |x|_{\ell^\infty} \le d\}$ and its discrete counterpart $B_h = \{\sqrt{h}x | x \in \mathbb{Z}^m,\ |x|_{\ell^\infty} \le d\}$, and denote their boundaries by ∂B and ∂B_h respectively. Let $a^0 = ta$, $t > 0$, be the point at which the ray generated by a meets ∂B, and set a^h to be the point on ∂B_h closest to a^0. Clearly $\sqrt{h}\, d \le |a^h| \le \sqrt{hm}\, d$ and since each component of the error $a^h - a^0$ is no bigger than $\sqrt{h}/2$ it follows that $|a^h - a^0| \le \sqrt{hm}/2$.

Consider now the angle, θ, formed between the vectors a^h and a^0. Since $0 \le \theta \le \pi/2$ it follows that

$$\theta \le (\pi/2)\sin(\theta) \le \frac{\pi}{2}\frac{|a^h - a^0|}{|a^h|} \le \frac{\pi\sqrt{m}}{4d}.$$

Selecting d to be the largest integer less than or equal to $1/h^{1/3}$ establishes the lemma. □

The proof of Lemma 4.3 will follow by induction upon the “number of laminates” N. The following lemma and its corollary will not only initialize the induction argument but will also be used for the inductive step.

Lemma 4.5. *Let* $H = \mu H_1 + (1-\mu)H_2 \in \mathrm{M}^{m\times n}$, $\mathrm{rank}(H_2 - H_1) \le 1$ *and* $H^h \in \mathcal{G}_h$ *satisfy* $|H - H^h| \le c_0 h^{1/3}$. *Then there exits* $\{(\mu^h, H_1^h), (1-\mu^h, H_2^h)\} \in \mathcal{H}_2$ *satisfying*

- $H_1^h, H_2^h \in \mathcal{G}_h$ *and* $H^h = \mu^h H_1^h + (1-\mu^h)H_2^h$,
- $H_i^h - H \in \mathbb{Z}\mathcal{R}_h$, $i = 1, 2$, *where*

$$\mathcal{R}_h = \{h(a \otimes b) : a \in \mathbb{Z}^m,\ b \in \mathbb{Z}^n, |a|_\infty, |b|_\infty < h^{-1/3}\},$$

- $|H_i - H_i^h| \le (c_0 + c_1)h^{1/3}$, $i = 1, 2$, *and*
- $|\mu - \mu^h| \le c_2 h^{1/3}/|H_1 - H_2|$,

where $c_1 = (\pi|H_2 - H_1|(\sqrt{m} + \sqrt{n})/2 + \sqrt{mn})$ *and* $c_2 = 2c_1$.

Proof. Let $H_2 - H_1 = a \otimes b$ with $a \in \mathbb{R}^m$, $b \in \mathbb{R}^n$, and without loss of generality let $|a| = |b|$. Let a^h, b^h be the vectors guaranteed by the Lemma 4.4, and note that $a^h \otimes b^h \in \mathcal{G}_h$.

If $|a^h||b^h| \ge |a||b|$ set $H_1^h = H_2^h = H^h$ and $\mu^h = \mu$. The error in H_1 is bounded by

$$\begin{aligned} |H_1 - H_1^h| &= |H - (1-\mu)(a \otimes b) - H^h| \\ &\le |H - H^h| + (1-\mu)|a||b| \\ &\le |H - H^h| + (1-\mu)|a^h||b^h| \\ &\le (c_0 + (1-\mu)\sqrt{mn})h^{1/3} \\ &\le (c_0 + c_1)h^{1/3} \end{aligned}$$

and the error in H_2^h is similarly bounded.

Suppose now that $|a^h||b^h| < |a||b|$. Let $k \in \mathbb{Z}$ and compute

$$\begin{aligned} &(1-\mu)(a \otimes b) - k(a^h \otimes b^h) \\ &= (1-\mu)\frac{|a||b|}{|a^h||b^h|}(\tilde{a} \otimes \tilde{b}) - k(a^h \otimes b^h) \\ &= (1-\mu)\frac{|a||b|}{|a^h||b^h|}\Big((\tilde{a} - a^h) \otimes \tilde{b} + a^h \otimes (\tilde{b} - b^h)\Big) \\ &\qquad + \left((1-\mu)\frac{|a||b|}{|a^h||b^h|} - k\right)(a^h \otimes b^h), \end{aligned}$$

where $\tilde{a} = |a^h|a/|a|$ and $\tilde{b} = |b^h|b/|b|$, so that $|\tilde{a}| = |a^h|$ and $|\tilde{b}| = |b^h|$. Selecting $k \ge 0$ to be the integer closest to $(1-\mu)|a||b|/(|a^h||b^h|)$ and noting that $|\tilde{a} - a^h| \le |a^h|\theta \le (\pi/2)|a^h|\sqrt{m}h^{1/3}$ gives the estimate

$$|(1-\mu)(a \otimes b) - k(a^h \otimes b^h)| \le \big(\pi(1-\mu)|a||b|(\sqrt{m} + \sqrt{n}) + \sqrt{mn}\big)\, h^{1/3}/2.$$

A similar computation shows that there is an integer $l \geq 0$ such that

$|\mu(a\otimes b)-l(a^h\otimes b^h)| \leq \left(\pi\mu|a||b|(\sqrt{m}+\sqrt{n})+\sqrt{mn}\right)h^{1/3}/2 \leq c_1 h^{1/3}$

Setting $H_1^h = H^h - k(a^h \otimes b^h)$ gives an error

$|H_1 - H_1^h| = |H - H^h - (1-\mu)(a\otimes b) + k(a^h \otimes b^h)| \leq (c_0 + c_1)h^{1/3}$

and setting $H_2^h = H^h + l(a^h \otimes b^h)$ gives a similar estimate.

Next, set $\mu^h = l/(l+k)$ (at least one of k or l is positive since $|a^h||b^h| < |a||b|$) and observe that $H^h = \mu^h H_1^h + (1-\mu^h)H_2^h$. Adding the error estimates for $(1-\mu)(a\otimes b)$ and $\mu(a\otimes b)$ gives

$$|(a\otimes b) - (k+l)(a^h \otimes b^h)| \leq c_1 h^{1/3}.$$

The triangle inequality implies

$$\Big|\mu|a||b| - l|a^h||b^h|\Big| \leq c_1 h^{1/3},$$
$$\Big||a||b| - (l+k)|a^h||b^h|\Big| \leq c_1 h^{1/3}$$

and the identity

$$\mu - \mu^h = \frac{\left(\mu|a||b| - l|a^h||b^h|\right) - \mu^h\left(|a||b| - (k+l)|a^h||b^h|\right)}{|a||b|}.$$

establishes the bound on the error $|\mu - \mu^h|$.

Finally, recall that $a^h \in \sqrt{h}\mathbb{Z}_d^m$ and $b^h \in \sqrt{h}\mathbb{Z}_d^n$, where $d < h^{-1/3}$, thus $H^h - H_1^h = k(a^h \otimes b^h) \in \mathbb{Z}\mathcal{R}_h$. □

Corollary 4.6. *The matrices in the lemma satisfy*

$$\begin{aligned}&|\mu f(H_1) + (1-\mu)f(H_2) - \mu^h f(H_1^h) - (1-\mu^h)f(H_2^h)| \\ &\qquad \leq (c_0 + c_1 + c_2)|f|_{\text{Lip}} h^{1/3}\ .\end{aligned}$$

Proof. Applying the triangle inequality to the identity

$$\begin{aligned}&\mu f(H_1) + (1-\mu)f(H_2) - \mu^h f(H_1^h) - (1-\mu^h)f(H_2^h) \\ &= (\mu-\mu^h)\left(f(H_1) - f(H_2)\right) + \mu^h\Big(f(H_1) - f(H_1^h)\Big) \\ &\qquad +(1-\mu^h)\Big(f(H_2) - f(H_2^h)\Big)\end{aligned}$$

gives

$$\begin{aligned}&\Big|\mu f(H_1) + (1-\mu)f(H_2) - \mu^h f(H_1^h) + (1-\mu^h)f(H_2^h)\Big| \\ &\leq \Big(|\mu-\mu^h||H_1 - H_2| + \mu^h|H_1 - H_1^h| + (1-\mu^h)|H_2 - H_2^h|\Big)|f|_{\text{Lip}}\end{aligned}$$

and the result follows. □

Proof. (of Lemma 4.3) Observe that we may assume without loss of generality that $c_0 \geq \sqrt{m}h/2$, and set

$$c_1 = \left(\pi \max_i |F_i| \, (\sqrt{m} + \sqrt{n}) + \sqrt{mn} \right).$$

We now induct on N. Lemma 4.5 and Corollary 4.6 establishe the case for $N = 2$. Suppose that the result holds for integers smaller than N, and let $\{(\lambda_i, F_i)\}_{i=1}^N \in \mathcal{H}_N$ and $H = \sum_i \lambda_i F_i$. Definition 2.1 guarantees the existence of $k \in \{1, \dots, N-1\}$ such that

$$\left(\frac{\lambda_i}{\lambda_1 + \cdots + \lambda_k}, F_i \right)_{i=1,\dots k} \in \mathcal{H}_k, \qquad \left(\frac{\lambda_i}{\lambda_{k+1} + \cdots + \lambda_N}, F_i \right)_{i=k+1,\dots N} \in \mathcal{H}_{N-k}$$

and $\{(\mu, H_1), (1-\mu, H_2)\} \in \mathcal{H}_2$ where $\mu = \lambda_1 + \cdots + \lambda_k$, and

$$H_1 = \frac{\lambda_1 F_1 + \cdots + \lambda_k F_k}{\lambda_1 + \cdots + \lambda_k}, \qquad H_2 = \frac{\lambda_{k+1} F_{k+1} + \cdots + \lambda_N F_N}{\lambda_{k+1} + \cdots + \lambda_N}.$$

Observe that $H = \mu H_1 + (1-\mu) H_2$, $\text{rank}(H_2 - H_1) \leq 1$, and there is $H^h \in \mathcal{G}_h$ satisfying $|H - H^h| \leq \sqrt{m}h/2 \leq c_0 h^{1/3}$ so, an application of Lemma 4.5 gives $\{(\mu_h, H_1^h), (1-\mu^h, H_2^h)\} \in \mathcal{H}_2$, and H_1^h, $H_2^h \in \mathcal{G}_h$ satisfying $|H_i - H_i^h| \leq (c_0 + c_1) h^{1/3}$. The inductive hypothesis when applied to each half of the splitting (with $c_0 + c_1$ playing the role of c_0) gives $\{(\mu_i^h, F_i^h)\}_{i=1}^k \in \mathcal{H}_k$ and $\{(\mu_i^h, F_i^h)\}_{i=k+1}^N \in \mathcal{H}_{N-k}$ satisfying

- $F_i^h \in \mathcal{G}_h, i = 1, 2, \dots, N$,and $H_1^h = \sum_{i=1}^k \mu_i F_i^h$; $H_2^h = \sum_{i=k+1}^N \mu_i F_i^h$,
- $|F_i - F_i^h| \leq ((c_0 + c_1) + \max(k-1, N-k-1) c_1) \, h^{1/3} \leq (c_0 + (N-1) c_1) h^{1/3}$,
-

$$\begin{aligned} \left| \sum_{i=1}^k \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_k} f(F_i) - \mu_i^h f(F_i^h) \right| &\leq ((c_0 + c_1) + (k+1) c_1) \, |f|_{\text{Lip}} h^{1/3} \\ &\leq (c_0 + (N+1) c_1) |f|_{\text{Lip}} h^{1/3}, \end{aligned}$$

$$\begin{aligned} \left| \sum_{i=k+1}^N \frac{\lambda_i}{\lambda_{k+1} + \cdots + \lambda_N} f(F_i) - \mu_i^h f(F_i^h) \right| &\leq ((c_0 + c_1) \\ &\quad + (N-k+1) c_1) \, |f|_{\text{Lip}} h^{1/3} \\ &\leq (c_0 + (N+1) c_1) |f|_{\text{Lip}} h^{1/3}. \end{aligned}$$

Defining $\lambda_i^h = (\lambda_1 + \cdots + \lambda_k) \mu_i^h$ for $i = 1, 2, \dots, k$ and $\lambda_i^h = (\lambda_{k+1} + \cdots + \lambda_N) \mu_i^h$ for $i = (k+1), \dots, N$, gives $\{(\lambda_i^h, F_i^h)\}_{i=0}^N \in \mathcal{H}_N$, and an application of the triangle inequality to the two equations above establishes the bound on the error in the average of the f values. $\square$

5. Numerical experiments

5.1. Kohn and Strang example

Our first experiment uses a slight modification of the following example computed in [17]. Define

$$f(X) = \begin{cases} 1 + |X|^2 & \text{if } X \neq 0, \\ 0 & \text{if } X = 0. \end{cases}$$

Then

$$f^{\mathrm{rc}}(X) = \begin{cases} 1 + |X|^2 & \text{if } \rho(X) \geq 1, \\ 2(\rho - D) & \text{if } \rho(X) \leq 1, \end{cases}$$

where $\rho(X) = \sqrt{|X|^2 + 2D}$, $D = |\mathrm{det} X|$. Now let

$$\tilde{f}(X) = \begin{cases} 1 + |X|^2 & \text{if } |X| \geq \sqrt{2} - 1, \\ 2\sqrt{2}|X| & \text{if } |X| \leq \sqrt{2} - 1. \end{cases}$$

Clearly, $f^{\mathrm{rc}} \leq \tilde{f} \leq f$ and therefore $\tilde{f}^{\mathrm{rc}} = f^{\mathrm{rc}}$. Moreover, $\tilde{f}$ satisfies the assumptions of Theorem 4.1 (f does not satisfy the assumptions since f is not continuous; however, we got the same numerical results for f and $\tilde{f}$). For our computations we used the following sets of rank-one matrices in $\mathcal{G}_{h,r}$:

$$\mathcal{R}_k = \{h(a \otimes b) : \, a \in \mathbb{Z}^m, \, b \in \mathbb{Z}^n, |a|_{\ell^\infty}, |b|_{\ell^\infty} \leq k\}.$$

The results of the computations are summarized in Table 5.1. We used a uniform grid in the cube $[\,-1, 1\,]^4$ with NGRID points on the one dimensional axes and computed the error in L^∞ at the grid points. We used the sets $\mathcal{R}_k$ with $k = 1, 2, 3$ which contain 16, 64 and 256 elements, respectively. The algorithm was implemented with an alphabetical ordering of the points on the grid, a fixed index k for the set of matrices $\mathcal{R}_k$, and the basic loop in the algorithm was iterated until the error stabilized. The number of iterations required for each case is given in the table in parenthesis. The different rows in Table 5.1 show the error for fixed parameters NGRID and k.

As expected, the error decreases as the grid is refined and the number of directions increases. While one could not expect to verify the rate of $h^{1/3}$ predicted by Theorem 4.2 with so few k values, it is clear that in order to get convergence that k and NGRID must increase together, as hypothesized in the theorem. Moreover, the table suggests that for large k the rate of convergence is about linear in $h = 2/\mathrm{NGRID}$ for this particular case.

Table 5.1 Numerical results for the (modified) Kohn Strang example

	$\|f^h - f^{\mathrm{rc}}\|$ (iterations)		
NGRID	$k=1$	$k=2$	$k=3$
5	0.085 786 (1)	0.085 786 (1)	0.085 786 (1)
9	0.043 861 (1)	0.043 861 (1)	0.043 861 (1)
17	0.067 187 (1)	0.031 250 (1)	0.031 250 (1)
33	0.076 384 (1)	0.021 354 (2)	0.012 649 (5)
65	0.076 488 (2)	0.026 795 (2)	0.009 730 (2)

5.2. Kohn's example

Kohn [16] explicitly computes the quasiconvex hull of functions of the form

$$f(X) = \frac{1}{2}\min(|X - A_1|^2, |X - A_2|^2) \tag{5.1}$$

to be

$$f^{\mathrm{rc}}(X) = \begin{cases} f_1 & f_1 - f_2 \le -\lambda/2 \\ f_2 & f_1 - f_2 \ge \lambda/2 \\ f_2 - (f_2 - f_1 + \lambda/2)^2/(2\lambda) & |f_2 - f_1| < \lambda/2 \end{cases}$$

where $f_1 = |X - A_1|^2/2$, $f_2 = |X - A_2|^2/2$ and λ is the maximal eigenvector of $(A_2 - A_1)^{\mathrm{T}}(A_2 - A_1)$. If $\mathrm{rank}(A_2 - A_1) = 1$ then f^{rc} is the convexification of f. Notice that f can not be trivially modified to satisfy the hypotheses of Theorem 4.1. However, Kohn's analysis reveals that f^{qc} is realized as a "simple laminate", that is, for each $X \in \mathrm{M}^{m\times n}$ there exists $\{(\mu, X_1), (1-\mu, X_2)\} \in \mathcal{H}_2$ such that

$$f^{\mathrm{rc}}(X) = \mu f(X_1) + (1-\mu) f(X_2).$$

Moreover, $|X - X_1|$, $|X - X_2| \le |A_2 - A_1|$, thus if $r > 0$ and $R > r + |A_2 - A_1|$, the restriction the results from algorithm $\mathcal{A}(f, \mathcal{G}_{h,R})$ to the grid $\mathcal{G}_{h,r}$ will converge at the established rate as $h \to 0$.

For our numerical experiments we selected the the matrices,

$$A_1 = \begin{bmatrix} 5/4 & 0 \\ 0 & 3/4 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 3\sqrt{3}/8 & 3/8 \\ -5/8 & 5\sqrt{3}/8 \end{bmatrix}$$

for which $\mathrm{rank}(A_2 - A_1) = 2$, $\lambda = |A_2 - A_1| = 3(7 - 5\sqrt{3})/8 \simeq 1.01$. We computed f_h^{rc} on $[-2,2]^4 \cap \mathcal{G}_h$, and computed the error on $[-1,1]^4 \cap \mathcal{G}_h$. It is clear that this problem will require a finer mesh than the previous example to resolve the variations of f, and this is indicated in Table 5.2. For this example there was no difference between the errors when $k = 2$ (64 rank one directions) and $k = 3$ (256 directions), and for these parameters the rate of convergence appeared to be approximately two.

Table 5.2 Numerical results for Kohn's example

	$\|f^h - f^{\rm rc}\|$ (iterations)		
NGRID	$k=1$	$k=2$	$k=3$
5	0.125 150 (0)	0.125 150 (0)	0.125 150 (0)
9	0.075 128 (1)	0.075 128 (1)	0.075 128 (1)
17	0.024 149 (1)	0.024 149 (1)	0.024 149 (1)
33	0.011 956 (1)	0.006 863 (1)	0.006 863 (2)
65	0.009 306 (1)	0.001 733 (1)	0.001 733 (1)

5.3. Tartar's infinite rank laminate

Our next experiment is based on ideas from [25] and [3]. We identify the diagonal matrices with points in $\mathbb{R}^2$ by $\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} = (x, y)$. Let $K = \{(\frac{1}{2}, 1), (1, -\frac{1}{2}), (-\frac{1}{2}, -1), (-1, \frac{1}{2})\}$ and $f(X) = \mathrm{dist}(X, K)$, where $\mathrm{dist}(.,.)$ is the Frobenius distance. While the quasi convexification of this function is not known, it is known that the zero set of $f^{\rm rc}$, is given by the square $[-\frac{1}{2}, \frac{1}{2}]^2$ and the four line segments parallel to the axes connecting the points in K and the square (see Fig. 5.1). Indeed, it is easy to see that this set must be contained in the zero set of $f^{\rm rc}$, since if $f^{\rm rc}(\frac{1}{2}, \frac{1}{2}) = \max\{f^{\rm rc}(\pm\frac{1}{2}, \pm\frac{1}{2})\}$, then by the rank-one convexity of $f^{\rm rc}$ we conclude

$$f^{\rm rc}(\tfrac{1}{2}, \tfrac{1}{2}) \leq \frac{2}{3} f^{\rm rc}(\tfrac{1}{2}, 1) + \frac{1}{3} f^{\rm rc}(\tfrac{1}{2}, -\tfrac{1}{2}) = \frac{1}{3} f^{\rm rc}(\tfrac{1}{2}, -\tfrac{1}{2}),$$

a contradiction unless $f^{\rm rc}(\frac{1}{2}, \frac{1}{2}) = f^{\rm rc}(\frac{1}{2}, -\frac{1}{2}) = 0$.

Using one iteration of the basic loop in the algorithm with NGRID$=$ 21 and $\mathcal{R}_1$ as the set of discrete rank-one directions, the level set of the computed (nonnegative) function restricted to the subspace of all diagonal matrices at level 0.01 is shown in Fig. 5.1. Note that the algorithm performs $882 = 2 \cdot (21)^2$ one dimensional convexifications in the subspace of all diagonal matrices. The third plot in Fig. 5.1 shows the level set of the level 0.0001 where we now performed the 882 rank-one convexifications in the subspace of all diagonal matrices in a stochastic order, i.e. we guessed 882 diagonal matrices and picked randomly one of the two rank-one directions $(0, 1)$ and $(1, 0)$. This example clearly indicates that one should analyze a stochastic version of the algorithm.

5.4. Solving a relaxed variational problem

For our final example we approximate the solution of a (relaxed) variational problem using the relaxed energy computed by our algorithm. Consider the

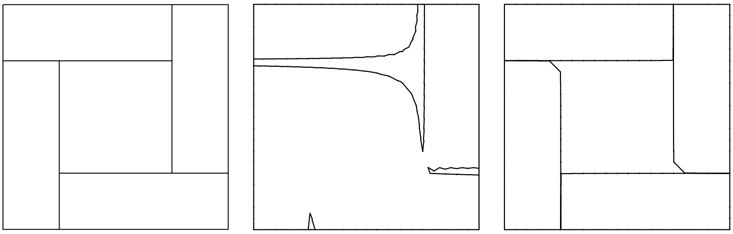

Fig. 5.1 Computation of a nontrivial zero set of a rank-one convex function in diagonal matrices; the left plot shows the exact zero set, the middle one the level set $\{f^{\mathrm{rc}} > 0.01\}$ of a discrete rank-one convexification computed with a deterministic algorithm while the right one shows the level set $\{f^{\mathrm{rc}} > 0.0001\}$ computed with a stochastic version. The little triangles at the points $(-0.5, 0.5)$ and $(0.5, -0.5)$ are artifacts generated by the triangulation used by the plotting algorithm

problem of minimizing

$$I(u) = \int_{\Omega} f^{\mathrm{rc}}(Du)\, dx \qquad u|_{\partial\Omega} = u_0$$

where $\Omega = (0,1)^2$ is the unit square. To approximate this problem we replace f^{rc} by the computed approximation f_h^{rc} and minimize over the class of functions which are piecewise linear on a triangulation of Ω. We consider the simplest situation where the unit square is divided into a uniform square mesh and the triangles formed by adding the diagonals of the small squares. Since we do not compute any gradients of f_h^{rc}, we use the very simple relaxation algorithm to find local minima:

- Initialize $\epsilon = 1$, $u|_{\Omega} = 0$, $u|_{\partial\Omega} = u_0$
- While $\epsilon > 10^{-6}$
  - For every basis function ϕ replace u by $u \pm \epsilon\phi$ if this lowers the energy $I(.)$.
  - If $u \pm \epsilon\phi$ never lowers the energy let $\epsilon \leftarrow \epsilon/2$.

Since f_h^{rc} is only known at the grid points $\mathcal{G}_h \subset \mathrm{M}^{m\times n}$ it is necessary to use an interpolation procedure to estimate $f_h^{\mathrm{rc}}(Du)$ for arbitrary gradients. Clearly it would be desirable to have an interpolation scheme that computed a rank one convex function from data that was "discretely" rank one convex. However, we do not know of any such scheme (or if this is even possible) so we used the natural multi–linear interpolation on the four dimensional cubes (i.e. tensor products of x_i and $(1 - x_i)$). This procedure is not completely satisfactory in the sense that the relaxation algorithm would typically fail to minimize $I(.)$. In an attempt to circumvent this problem we initially added a term of the form $\nu|Du|^2$ to the integrand, to compensate for the lack of (quasi) convexity of the interpolant. As the iterations proceeded

Table 5.3 Energy $I(u_h)$ and $\|\nabla u_h - B\|_{L^2}$ for affine boundary data; Kohn's energy

	$f^{qc} = f^{rc}$	f		f_h^{rc}	
$h\backslash\nu$	0.0	0.0	10.0	0.0	10.0
1/8	0.199 761	0.252 027	0.277 667	0.202 507	0.200 710
	0.000 060	0.352 011	0.000 011	0.082 143	0.000 013
1/16	0.199 761	0.240 885	0.245 972	0.208 264	0.200 716
	0.000 218	0.387 229	0.324 721	0.193 382	0.000 237
1/32	0.199 761	0.236 125	0.236 678	0.245 633	0.230 610
	0.000 939	0.402 100	0.395 681	0.468 030	0.372 291

(and u converged) ν was reduced to zero. In the two examples below we considered the relaxation of Kohn's bulk energy function (5.1) computed on a mesh having $\text{NGRID}= 65$ (see Table 5.2).

We first consider an example with affine boundary conditions for which the minimizer of the relaxed problem is the extension of the affine boundary values to the interior. On the boundary we specified $u(x) = Bx + c$ where $c = (0, 1/4)^{\mathrm{T}}$ and

$$B = \begin{bmatrix} 1/2 & 1/4 \\ -1/4 & 15/32 \end{bmatrix}.$$

The relaxation is achieved as $0.199761 \simeq f^{rc}(B) = \lambda f(B_1) + (1-\lambda)B_2$ where

$$\lambda \simeq 0.394493, \quad B_1 \simeq \begin{bmatrix} 0.871019 & 0.035792 \\ 0.121019 & 0.254542 \end{bmatrix}, \quad B_2 \simeq \begin{bmatrix} 0.258278 & 0.389558 \\ -0.491722 & 0.608308 \end{bmatrix}.$$

In Table 5.3 we tabulate the energy values $I(u_h)$ and the norm $\|\nabla u_h - B\|_{L^2}$ computed using various mesh sizes, bulk energy functions and parameter values ν. As expected, when $f^{rc} = f^{qc}$ was used for the bulk energy function the affine solution was always found, and computations with the original function f exhibit some form of oscillation. For the coarser meshes, $h = 1/8$ and $h = 1/16$ the numerical relaxation f_h^{rc} was adaquate for computation of the affine solution. However, when $h = 1/32$ the mesh appears to be sufficiently fine to detect errors in the interpolant of f_h^{rc}. For example, even when ν is initially set to 100.0 (so that the solution is initially driven to the affine function) the computed local minima had curved ℓ^1 contours[1] as indicated in Fig. 5.2. This suggests that the lack of quasi–convexity in the interpolant of the discrete data is playing a role.

For a second example we selected the boundary data to have degree one. While this is not particularly interesting from a physical stand point (corresponding to an inversion), it does force the minimizer to take on many gradient values (unlike the previous example), so this computation may be

[1] $|u|_{\ell^1} = |u_1| + |u_2|$

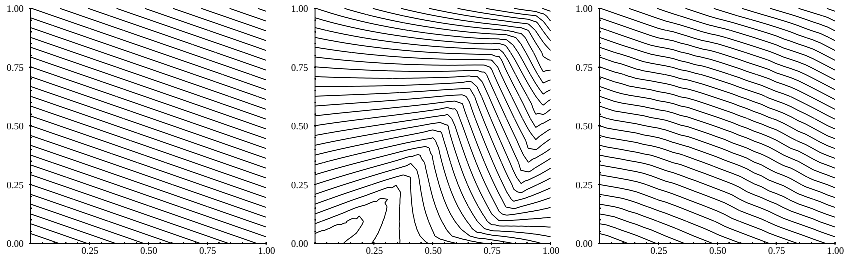

Fig. 5.2 Contour Plots of $|u|_{\ell^1}$. (a) $f^{\mathrm{rc}} = f^{\mathrm{qc}}$, $\nu = 0$, (b) f, $\nu = 0$, (c) f_h^{rc}, $\nu = 100$

Table 5.4 Energy $I(u_h)$ for degree one variational problem; Kohn's energy

	$f^{\mathrm{qc}} = f^{\mathrm{rc}}$	f		f_h^{rc}		
$h\backslash\nu$	0.0	0.0	10.0	0.0	10.0	100.0
1/8	0.173 627	0.200 565	0.189 265	0.175 216	0.175 104	0.175 103
1/16	0.170 198	0.188 119	0.187 324	0.173 716	0.172 598	0.172 528
1/32	0.169 017	0.183 546	0.184 685	0.183 599	0.172 637	0.172 126

more typical of what one may encounter. If

$$r = r(x, y) = \sqrt{(x - 1/2)^2 + (y - 1/2)^2 + 1/4}$$

the boundary values we consider are $u(x, y) = (1/r)(x - 1/2, y - 1/2)^{\mathrm{T}}$. Energy values computed with these boundary values are listed in Table 5.4, and reproduce most of the trends appearing in Table 5.3. In particular, the energy computed with $f^{\mathrm{qc}} = f^{\mathrm{rc}}$ decreased in a monotone fashion as the mesh was refined, and the energy computed using f_h^{rc} initially decreased before becoming sensitive to the lack of convexity in the interpolant of f_h^{rc}.

Acknowledgement. Much of work was done while GD held a postdoctoral research fellowship at the Center for Nonlinear Analysis at Carnegie Mellon University in Pittsburgh, and while NW was visiting the Max Planck Institute for Mathematics in the Sciences in Leipzig. The hospitality and stimulating environment provided by these institutions are greatfully acknowledged.

References

1. Ball, J.M. (1988): A version of the fundamental theorem for Young measures. In: Serre, D. (ed.) Partial Differential Equations and Continuum Models of Phase Transitions. Springer
2. Ball, J.M., James, R.D. (1992): Proposed experimental tests of a theory of fine microstructure and the two-well problem. Phil. Trans. Soc. London, A **338**, 389–450
3. Bhattacharya, K., Firoozye, N.B., James, R.D., Kohnm, R.V. (1994): Restrictions on microstructure. Proc. Royal Soc. Edinburgh **124**, 843–878

4. Carstensen, C., Plechac, P. (1998): Numerical solution of the scalar double–well problem allowing microstructure. Preprint
5. Chipot, M. (1990): Numerical analysis of oscillations in nonconvex problems, Numer. Math. **59**, 747–767
6. Chipot, M., Collins, C. (1992): Numerical approximations in variational problems with potential wells. SIAM Journal on Numerical Analysis **29**, 1002–1019
7. Collins, C., Kinderlehrer, D., Luskin, M. (1991): Numerical approximation of the solution of a variational problem with a double well potential. SIAM Journal on Numerical Analysis **28**, 321–332
8. Collins, C., Luskin, M. (1991): Optimal order error estimates for the finite element approximation of the solution of a nonconvex variational problem. Mathematics of Computation **57**, 621–637
9. Dacorogna, B. (1989): Direct Methods in the Calculus of Variations. Springer Verlag
10. Dolzmann, G. (1997): Numerical computation of rank–one convex evelopes. SIAM Journal on Numerical Analysis, submitted
11. Gremaud, P.A. (1994): Numerical analysis of a non–convex variational problem related to solid–solid phase transitions. SIAM Journal on Numerical Analysis **31**, 111–127
12. Gremaud, P.A. (1995): Numerical optimization and quasiconvexity. European J. Appl. Math. **6**, 69–82
13. Kinderlehrer, D., Pedregal, P. (1992): Weak convergence of integrands and the Young measure representation. SIAM Journal on Mathematical Analysis **23**, 1–19
14. Gremaud, P.A. (1994): Weak convergence of integrands and the Young measure representation. The Journal of Geometric Analysis **4**, 59–90
15. Kloucek, P., Luskin, M. (1994): Computational modeling of the martensitic transformation with surface energy. Mathematical and Computer Modelling **20**
16. Kohn, R.V. (1991): The relaxation of a double well energy, Contin. Mech. Thermodyn **3**, 193–236
17. Kohn, R.V., Strang, G. (1986): Optimal design and relaxation of a variational problems, I–III. Communications on Pure and Applied Mathematics **XXXIX**, 113–137, 139–182, 353–377
18. Li, B., Luskin, M. (1998): Finite element analysis of microstructure for the cubic to tetragonal transformation. SIAM Journal on Numerical Analysis **35**, 376–392
19. Luskin, M. (1996): On the computation of crystalline microstructure. Acta Numerica **5**, 191–258
20. Luskin, M., Ma, L. (1992): Analysis of the finite element approximation of microstructure in micromagnetics. SIAM Journal on Numerical Analysis **29**, 320–331
21. Matousek, J., Plechac, P. (1998): On functional seperately convex hulls. Disc. Comput. Geometry (to appear)
22. Morrey, C.B. (1966): Multiple Integrals in the Calculus of Variations. Springer
23. Nicolaides, R.A., Walkington, N.J. (1993): Computation of microstructure utilizing Young measure representations. J. Intelligent Material System and Structures **4**, 457–462
24. Tartar, L. (1983): The compensated compactness method applied to systems of conservation laws. In: Ball, J.M. (ed.) Systems of Nonlinear Partial Differential Equations. Reidel, pp. 263–285
25. Tartar, L. (1983): Some remarks on seperately convex functions. In: Kinderlehrer (ed.) Microstructure and Phase Transitions. Springer, pp. 191–204
26. Šverák, V. (1993): Rank one convexity does not imply quasiconvexity. Proc. Roy. Soc. Edinburgh, Sect. A **120**, 185–189
27. Young, L.C. (1980): Lectures on the Calculus of Variations and Optimal Control. Chelsa