NAMT
95-007

Numerical Methods for a Nonconvex Optimization Problem Modeling Martensitic Phase Transitions

Roy A. Nicolaides
Noel Walkington
Han Wang
Carnegie Mellon University

Research Report No. 95-NA-007

February 1995

Sponsors

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

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

(CP) Find $U \in \mathcal{V} = \{\mathbf{w} : \Omega \to \mathbf{R}^2; \mathbf{w} \text{ is continuous and piecewise linear}\}$ such that:

$$E(U) = \inf_{\mathbf{w} \in \mathcal{V}} E(\mathbf{w}) \tag{2.4}$$

and

$$U(\mathbf{x}) = \frac{1}{2}(F_0 + RF_1)\mathbf{x} \qquad \mathbf{x} \in \partial\Omega \tag{2.5}$$

where $E(.)$ is defined as (2.1) with W defined in (2.3).

3. Square Mesh

In this section we will investigate some numerical problems which occur when a uniform square mesh is used to discretize the problem (CP) in section 2.

Denote by $U^h : \Omega \to \mathbf{R}^2$ a continuous piecewise bilinear vector valued function on a uniform square mesh Ω^h of size $h = 1/n, n \in \mathbf{N}$, where

$$\Omega^h = \{\mathbf{x}_{ij} \in \mathbf{R}^2 : \mathbf{x}_{ij} = (ih, jh), i, j = 0, 1, 2, ..., n\}.$$

Ω_{ij} denotes the square element with corners $\mathbf{x}_{ij}, \mathbf{x}_{i+1,j}, \mathbf{x}_{i+1,j+1}$ and $\mathbf{x}_{i,j+1}$. Since U^h is bilinear in each element Ω_{ij}, the energy density function $W(\nabla U^h)$ in (2.3) is an 8th degree polynomial in $\mathbf{x}$. To reduce the computational cost of integration of such a high degree polynomial, the usual approach is to compute the energy $E(U^h)$ approximately by calculating ∇U^h at the center of Ω_{ij}. This constant matrix, denoted by $\overline{\nabla U^h_{ij}}$, can be obtained also by averaging directional gradients of U^h on the edges of each element Ω_{ij}. We denote this approximated energy by $E_h(U^h)$:

$$E_h(U^h) = \sum_{i,j=0}^{n-1} W(\overline{\nabla U^h_{ij}}) \times h^2 \tag{3.1}$$

where W is defined in (2.3).

Now we consider the following discretized problem:

(DP) Find $U^h \in \mathcal{V}^h = \{\mathbf{w} : \Omega \to \mathbf{R}^2; \mathbf{w} \text{ is continuous in } \Omega \text{ and bilinear in } \Omega_{ij},$

$i, j = 0, 1, ..., n\}$ such that:

$$E_h(U^h) = \inf_{\mathbf{w} \in \mathcal{V}^h} E_h(\mathbf{w}) \tag{3.2}$$

and

$$U^h(\mathbf{x}) = \frac{1}{2}(F_0 + RF_1)\mathbf{x} \qquad \mathbf{x} \in \partial\Omega \tag{3.3}$$

where $E_h(.)$ is defined in (3.1).

Using the boundary condition (2.5), if $E(U)$ is relatively small we expect the gradient distribution of U to have the structure illustrated in Figure 1. In the discretized problem with the square mesh, the best numerical result to (DP) we can expect is a staircase structure illustrated in Figure 2. The number of diagonal bands observed should be dependent on the mesh size and should increase as the mesh is refined.

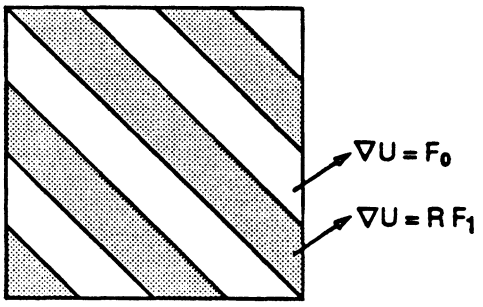


Figure 1

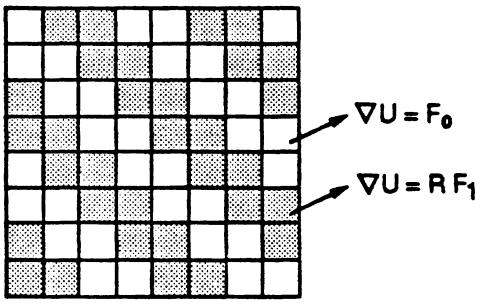
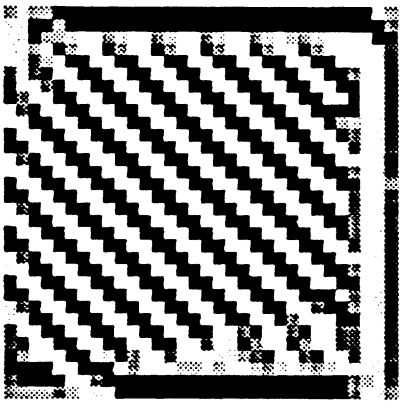


Figure 2



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10, h = 1/32$
$E_h(U^h) = 4.81, \quad E(U^h) = 598.3$

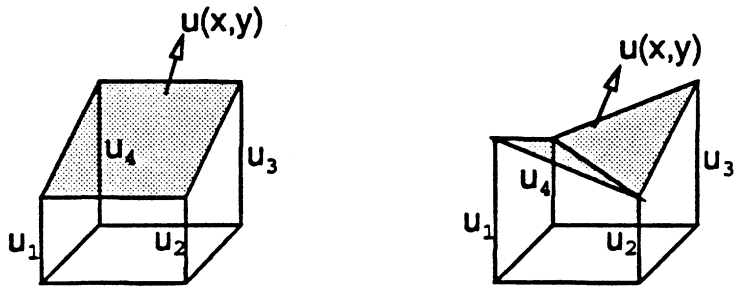
Figure 3

Figure 3 plots a computed gradient distribution (projected onto energy wells) of U^h obtained by minimizing the energy $E_h(U^h)$ in (DP). White colored squares denote well F_0 and black squares denote well RF_1. Intermediate shades measure closeness to the wells.

While the solution depicted in Figure 3 appears reasonable, it is in fact a representation

of a numerical artifact. To see this we can compute the exact energy $E(.)$ for the bilinear function U^h by using high order Gauss quadrature. We find that $E(U^h)$ is not small at all. It is only $E_h(U^h)$ that is (relatively) small.

To explain this phenomenon, consider a bilinear vector valued function $U^h = \begin{pmatrix} u(x,y) \\ v(x,y) \end{pmatrix}$ defined on an element Ω_{ij}. The energy function $E(U^h)$ only depends on $\nabla U^h = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$, where each entry is a linear function since U^h is bilinear. On the element Ω_{ij}, u and v each have four (dependent) corner values which uniquely determine the bilinear function U^h. In the energy $E_h(U^h)$, ∇U^h is approximated by its value at the center of the element. Alternatively this matrix can be computed by averaging the horizontal and vertical slopes on the boundary $\partial\Omega_{ij}$ since all entries in it are linear. This approximate gradient matrix, denoted by $\overline{\nabla U^h}$, does not change if U^h is symmetrically bent along a diagonal line of the square. For instance, if u equals u_1, u_2, u_3, u_4 on four corners over this square element (see Figure 4), then u_x is approximated by $\overline{u_x} = \frac{1}{2}(\frac{u_2-u_1}{h} + \frac{u_3-u_4}{h})$ and u_y is approximated by $\overline{u_y} = \frac{1}{2}(\frac{u_4-u_1}{h} + \frac{u_3-u_2}{h})$. If u_1 and u_3 are simultaneously replaced by u_1+1 and u_3+1, neither u_x nor u_y are affected. This means that the skewing ('bending' along diagonals of squares) of U^h does not affect $E_h(U^h)$ at all since $E_h(U^h)$ only involves $\overline{\nabla U^h}$. The microstructure in the Figure 3 was obtained in this way. Thus while $E_h(U^h)$ for this example is 4.81, the exact energy $E(.)$ for the bilinear function U^h is 598.3. Figure 5 plots $u(x,y)$ over part of Ω. The skewing is clear in this figure.



Strict linear function Skewing

Figure 4

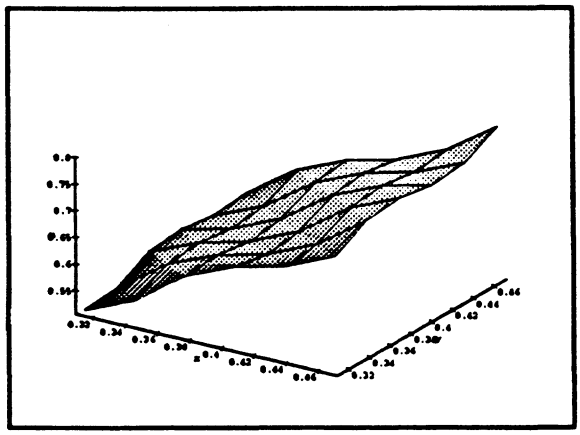


Figure 5

The microstructure orientation is not the reason for the skewing. The skewing is generated by the method in which we compute $E_h(U^h)$. If we rotate the domain by 45° to align the mesh with the normal n the effect of skewing can be more serious. For instance, letting $Q = \begin{pmatrix} \cos\frac{\pi}{4} & -sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & cos\frac{\pi}{4} \end{pmatrix}$, we expect to get a U^h which has roughly the following microstructure (ignoring boundary layers) :

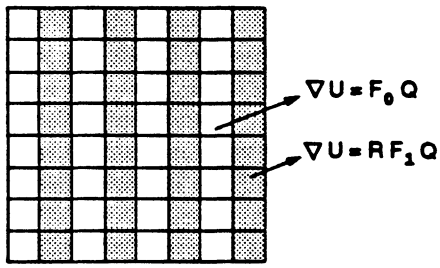


Figure 6

Now consider Figures 7 and 8. In Figure 7 we plot $u(x, y)$ over part of Ω adjacent to the boundary $y = 0$. Figure 8 is the projection of $\overline{\nabla U^h}$ onto the energy wells. In this figure, there should be a 'boundary layer' near $y = 0$, but it does not form. These figures suggest

that the skewing allows $\overline{\nabla U^h}$ to be in a well, even though U^h itself is not likely to be a good minimizer.

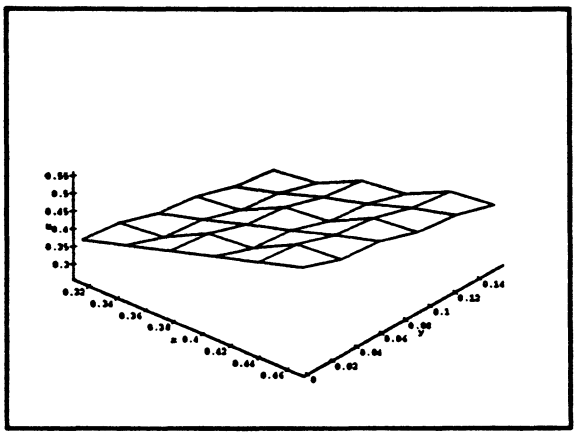
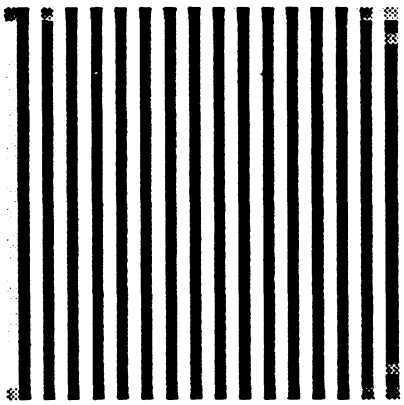


Figure 7



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10, h = 1/32$
$E_h(U^h) = 0.501 \quad E(U^h) = 140.64$

Figure 8

To check this hypothesis we introduced a penalty term in the energy function which

penalizes skewing. It is as follows (restricted to one element) :

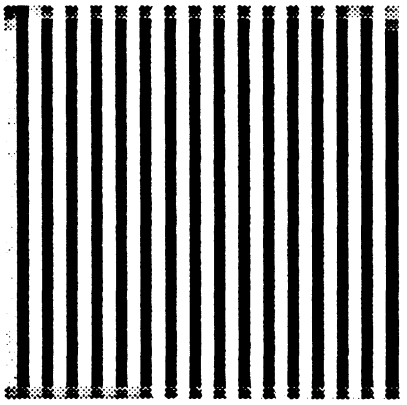
$$p[(u_2 - u_1) - (u_3 - u_4)]^2 + p[(v_2 - v_1) - (v_3 - v_4)]^2,$$

where p is a constant weight. Minimizing $E_h(U^h)$+ penalty term produces a solution without noticeable skewing. Next we formed bilinear vector fields U^h with the computed values and computed the exact energy $E(U^h)$ of these fields. The results are as follows

	Standard Approach	Penalized Approach
Exact $E(U^h)$	140.64	18.33
Approx. $E_h(U^h)$	0.501	1.13

The energy $E(U^h)$ of the skewed solution is relatively large, indicating that skewing exacts a severe energy penalty, even though the discretized energy suggests that the skewed solution is better.

Figure 9 shows the microstructure of the penalized solution. The microstructure is similar to that of the skewed solution (Figure 8) except it exhibits a 'boundary layer'. Analogous with Figure 7, Figure 10 plots the penalized $u(x, y)$ over the same region in Ω. It is indeed unskewed except near the boundary.



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10, W = 1000, h = 1/32$
$E_h(U^h) = 1.13$

Figure 9

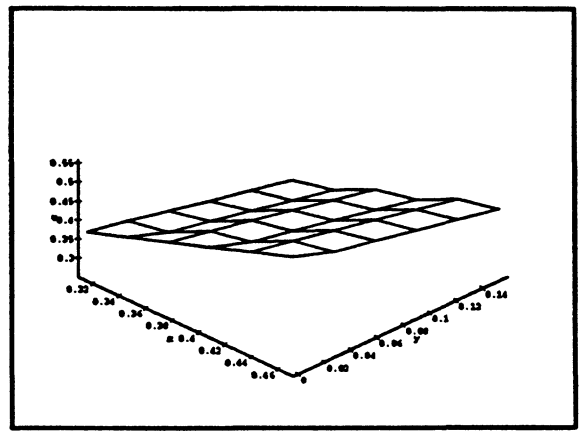


Figure 10

We emphasize that we are not advocating the penalty technique as a general approach. It is merely a device to obtain an unskewed solution to compare with the skewed solution for the particular test problem considered here.

4. Triangular Mesh

In this section we will discuss the use of uniform triangular meshes for discretizing (2.1).

Consider a triangulation of Ω:

$$\Omega^h = \{\mathbf{x}_{ij} \in \mathbf{R}^2 : \mathbf{x}_{ij} = (ih, jh), i, j = 0, 1, 2, ..., n\}.$$

In Figure 11 mesh A, let UT_{ij} denote the upper triangle with vertices $\mathbf{x}_{i+1,j}$, $\mathbf{x}_{i+1,j+1}$, $\mathbf{x}_{i,j+1}$ and LT_{ij} denote the lower triangle with vertices $\mathbf{x}_{i+1,j}$, $\mathbf{x}_{ij}$, $\mathbf{x}_{i,j+1}$. Let U^h be a continuous piecewise linear vector valued function defined on Ω^h. In this case the constant gradient matrix ∇U^h can be computed exactly on each triangular element of Ω^h.

We solve the following discretized problem on Ω^h:

(DP-T) Find $U^h \in \mathcal{V}^h = \{\mathbf{w} : \Omega \to \mathbf{R}^2; \mathbf{w}$ is continuous in Ω and linear in UT_{ij} and $LT_{ij}\ \forall i, j\}$ such that:

$$E(U^h) = \inf_{\mathbf{w} \in \mathcal{V}^h} E(\mathbf{w}) \tag{4.1}$$

and

$$U^h(\mathbf{x}) = \frac{1}{2}(F_0 + RF_1)\mathbf{x} \qquad \mathbf{x} \in \partial\Omega \tag{4.2}$$

where $E(.)$ is defined in (2.1).

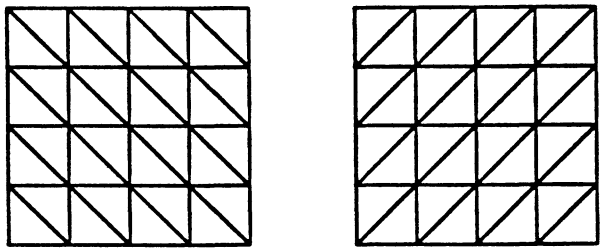
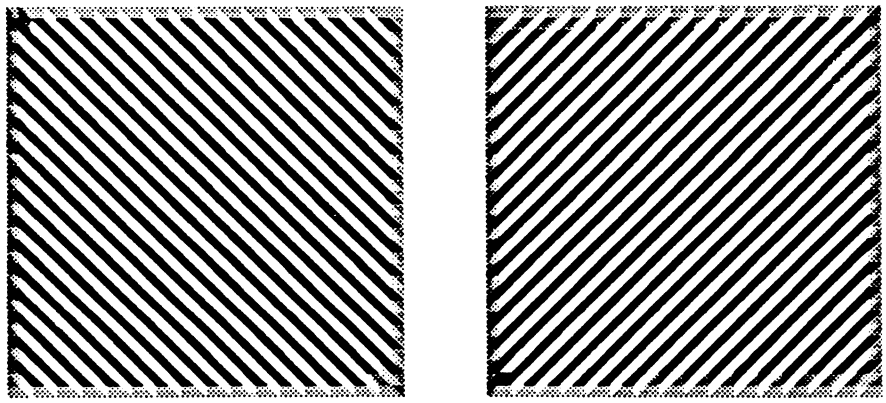


Figure 11



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10$. Mesh size: 32x32

$E(U^h) = 5.92$ $\qquad$ $E(U^h) = 13.56$

Figure 12 $\qquad$ **Figure 13**

With Mesh A we obtained the expected microstructure (Figure 12) by minimizing $E(U^h)$ in (DP-T). Notice that the normal $\mathbf{n} = (1,1)^t$ between F_0 and RF_1 matches the orientation of Mesh A so that the alternating layers of U^h aligned with mesh lines are perpendicular to $\mathbf{n}$.

A difficulty arises if a triangular mesh whose orientation does not fit the expected mi-

crostructure is used in the discretization. Mesh B illustrated in Figure 11 is an example of such case. If we try to solve (DP-T) by using this mesh, a continuous piecewise linear function U^h with a small energy is found. But this U^h displays a fine structure which has gradient layers aligned with the mesh instead of being perpendicular to the normal $\mathbf{n}$.

Figure 13 plots the gradient structure of a such example. In this picture the gradient of U^h forms alternating layers fitting the mesh with two matrices G_0 and G_1, which are totally different from F_0 and RF_1. But G_0 and G_1 are very close to the two energy wells SO_2F_0 and SO_2F_1 respectively, in the sense that $W(G_0)$ and $W(G_1)$ are very close to zero and hence the energy $E(U^h)$ is relatively small. In fact, G_0 is very close to RF_0 and G_1 is very close to F_1. We also find that

$$(G_0 + G_1)/2 = (F_0 + RF_1)/2,$$

and

$$G_0 - G_1 = \mathbf{b} \otimes \mathbf{m}, \quad \text{where } \mathbf{m} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

In other words, G_0 and G_1 satisfy the boundary condition in an average sense and are rank-1 connected with the normal $\mathbf{m} = (1, -1)^t$. An interesting fact is that RF_0 and F_1 are also on the energy wells and satisfy the rank-1 connection with the normal $\mathbf{m}$, but their average does not match the boundary condition. Figure 14 demonstrates these relationships. Circles denote energy wells and lines between matrices represent rank-1 connections.

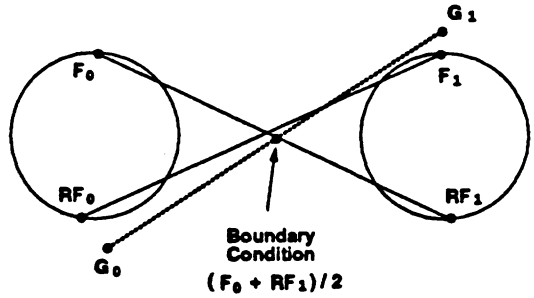


Figure 14

The question now is whether a minimizing sequence to the problem (DP-T) can be constructed using Mesh B. To answer this question, we form the following table which lists minimized energy of functions U^h computed with same initial and boundary conditions in Mesh A and Mesh B for different mesh sizes.

Mesh Size	Mesh A		Mesh B	
	Energy $E(U^h)$	Rate of Convergence	Energy $E(U^h)$	Rate of Convergence
8	28.20		33.47	
16	12.45	1.18	18.94	1.36
24	7.98	1.10	15.22	1.26
32	5.92	1.04	13.56	1.22
48	3.88	1.04	12.13	1.11
64	2.92	0.98	11.46	1.08
96	1.96	0.98	10.82	1.06
128	1.45	1.06	10.52	1.00
256	0.73	0.98	10.03	1.16

First consider Mesh A. Since the orientation of this mesh matches the normal of the rank-1 connection between F_0 and RF_1, U^h can be made from layers with gradient matrices equal to them alternately such that the energy is close to zero in the interior elements of Ω^h. So the energy $E(U^h)$ can be estimated by the total area of all boundary elements multiplied by a constant C. This gives an upper bound for the energy density function $W(U^h)$ on these element, so that

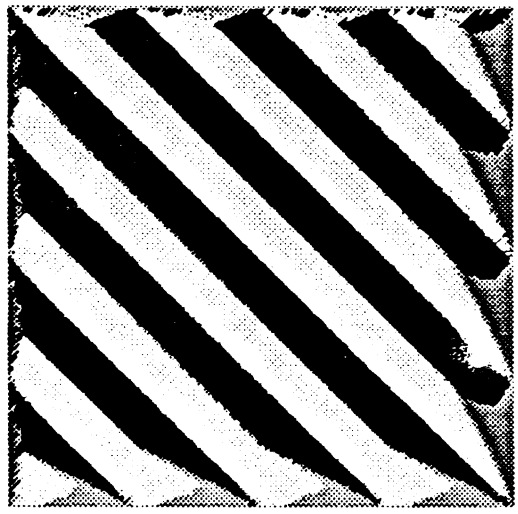
$$E(U^h) \approx 4n \times (h^2/2) \times C = O(h).$$

The second and third columns of the table show that the minimized energy of U^h indeed approaches zero approximately linearly in h. Thus U^h forms a minimizing sequence as h approaches zero in this case. These functions all display the kind of microstructures shown in Figure 12.

On the other hand when Mesh B is used, the minimized energy $E(U^h)$ is not found to converge to zero as the mesh is refined. Hence the functions U^h computed in Mesh B in the table do not seem to be forming a minimizing sequence, although the gradient distributions of these functions all display the fine structures similar to the one in Figure 13. In these structures the gradients of the functions equal G_0 and G_1 alternately in the fine bands. As stated earlier, G_0 and G_1 seem in practice to satisfy the boundary conditions in an average sense. Adopting this as a hypothesis, we are able to predict a limiting value for the energies appearing in the table. This calculation is in Appendix A and gives the limiting value 9.624747 which is reasonably close to the extrapolation of those energies in the fourth column of the table. The convergence rates in the last column are calculated from the differences between the energies and the above limiting value and they appear to be linear in h. This supports our hypothesis and may suggest computational techniques to eliminate these solutions from coarse mesh calculations.

The coarseness and orientation of Mesh B are the main reasons for occurrence of the incorrect gradient structures. However, a minimizing sequence can still be constructed in this kind of mesh provided it is sufficiently fine. To illustrate this, Figure 15 plots the

distribution of ∇U^h on a 256x256 mesh where $E(U^h) = 8.08$. The significance of this value is that it is lower than the minimum energy to which the function with gradients G_0 and G_1 in the alternating bands converges in Mesh B.



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10$. Mesh size: 256x256
$E(U^h) = 8.08$ Mesh B
Figure 15

In this kind of structure, the interface between any two bands contributes relatively large amounts of energy because of the orientation of the mesh. Since narrower bands generate more interfaces but with less energy in the boundary strip and wider bands involve more energy in the boundary strip but use fewer interfaces, there exists a optimal width of bands in this structure. In Appendix B, it is shown that if the width of the bands is $O(\sqrt{h})$, then the energy approaches zero like $\sqrt{h}$.

An alternative way to solve the orientation problem is to use the more symmetric triangular mesh illustrated in Figure 16. This mesh permits more freedom of orientation than the two meshes in Figure 11. Using this mesh, we can get four kinds of microstructures with four different normals. Figure 17 shows four possible microstructures in this mesh. Some computed microstructures are shown in Figure 18. In fact a simple layered microstructure of the types shown here with arbitrary orientation can be computed by a suitable choice of the ratio between the mesh sizes of the x-axis and y-axis.

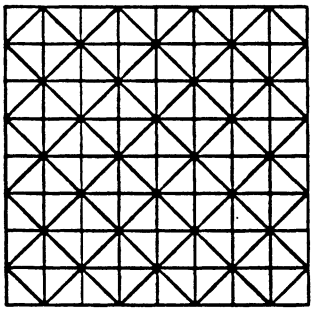


Figure 16

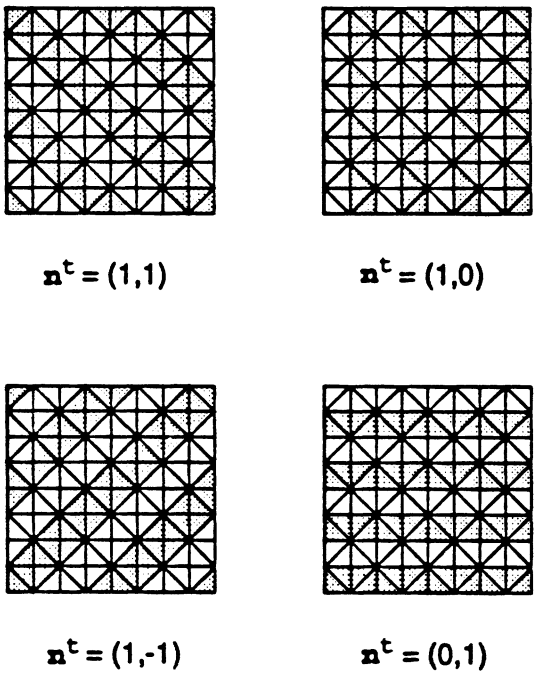
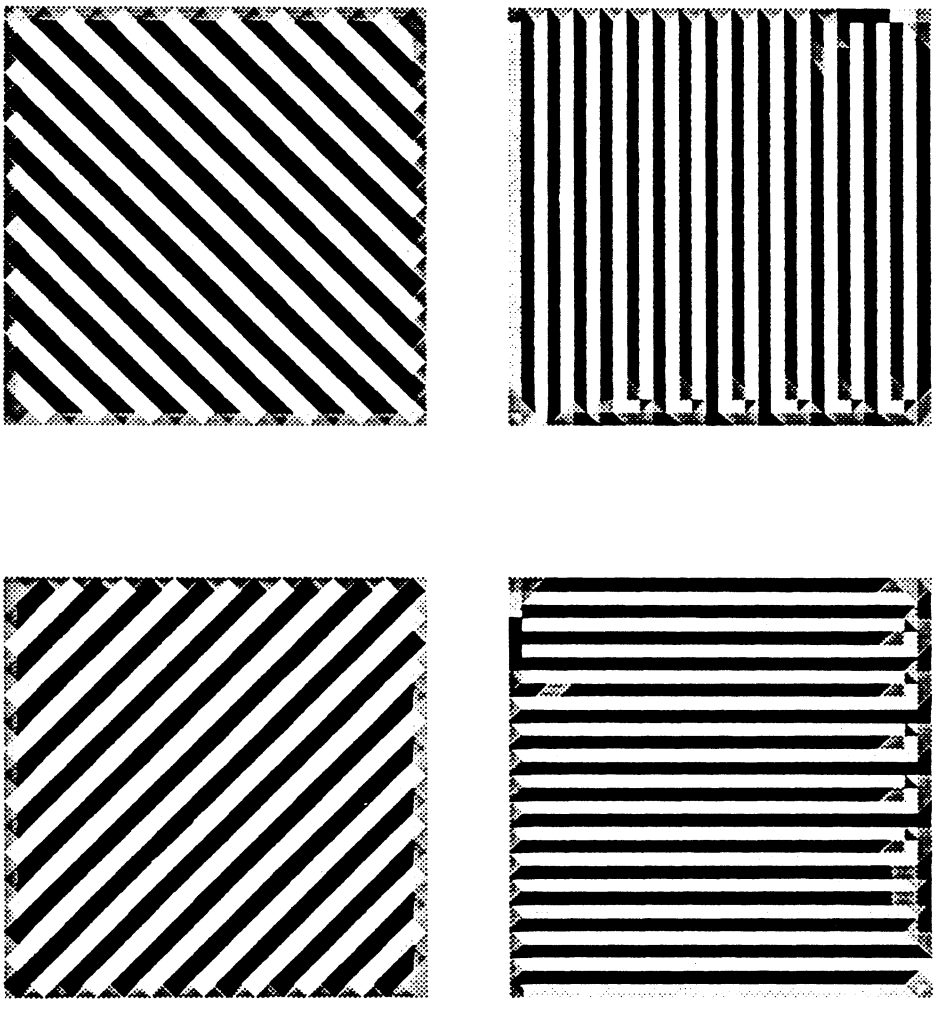


Figure 17



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10, h = 1/32.$

Figure 18

5. MULTI VARIANTS

In this section we will use the symmetric type of mesh discussed in section 4 to compute microstructures which involve more than two variants from two energy wells.

We first consider a three variants case. The two matrices F_0 and RF_1 defined in section 2 are chosen as the first two variants. Then we look for a rotation $Q \in SO_2$ such that QF_0 and $(F_0 + RF_1)/2$ are rank-1 connected:

$$QF_0 - \frac{1}{2}(F_0 + RF_1) = \mathbf{b} \otimes \mathbf{m}. \tag{5.1}$$

QF_0 is then chosen as the third variant. See Figure 19.

Solving the equation (5.1) gives :

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \mathbf{m} = \begin{pmatrix} \frac{1}{2}\alpha + \frac{3}{2}\beta \\ -\frac{3}{2}\alpha - \frac{1}{2}\beta \end{pmatrix},$$

where

$$\sin\theta = -\frac{8\sqrt{\alpha\beta}(\alpha-\beta)}{(\alpha-\beta)^2 + 16\alpha\beta}, \qquad \cos\theta = -\frac{(\alpha-\beta)^2 - 16\alpha\beta}{(\alpha-\beta)^2 + 16\alpha\beta}.$$

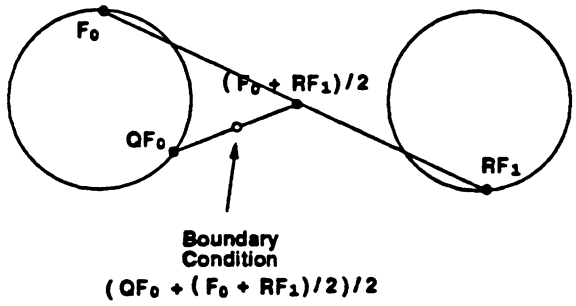


Figure 19

If the boundary condition is defined by

$$U(\mathbf{x}) = \frac{1}{2}(QF_0 + \frac{1}{2}(F_0 + RF_1))\mathbf{x} \qquad \mathbf{x} \in \partial\Omega,$$

we expect a microstructure of the form shown in Figure 20.

Using $\alpha = 5$ and $\beta = 1$, we get $\mathbf{n} = (1,1)^t$ and $\mathbf{m} = (1,-2)^t$, where $\mathbf{n}$ is the normal between F_0 and RF_1, and $\mathbf{m}$ is the normal between QF_0 and $(F_0 + RF_1)/2$. For convenience we rotated the domain by 45° to let $\mathbf{n} = (1,0)^t$ and $\mathbf{m} = (1,3)^t$. The mesh will fit these normals if we use $h_x = 3h_y$, where h_x and h_y are mesh sizes in the x and y directions. Our computed solution is shown in Figure 21.

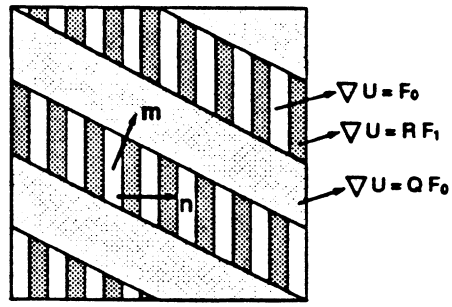
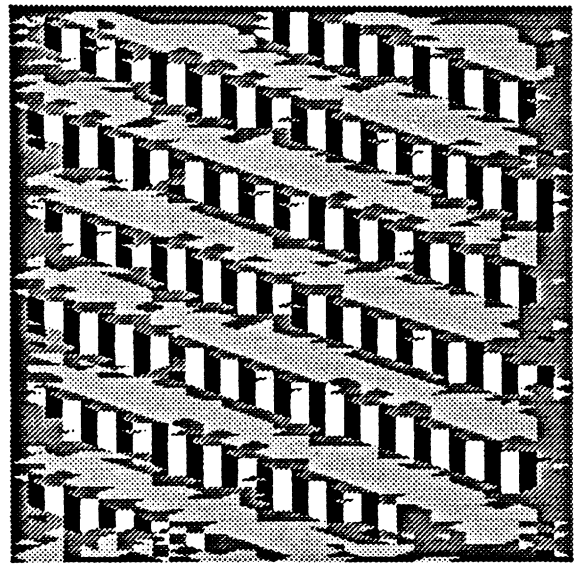


Figure 20



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10$. Mesh size: 32x96
$E(U^h) = 9.41$

Figure 21

In the four variants case, we make use of the following rank-1 relations:

$$RF_1 - F_0 = \mathbf{a}_1 \otimes \mathbf{n}_1,$$

$$RF_0 - F_1 = \mathbf{a}_2 \otimes \mathbf{n}_2,$$

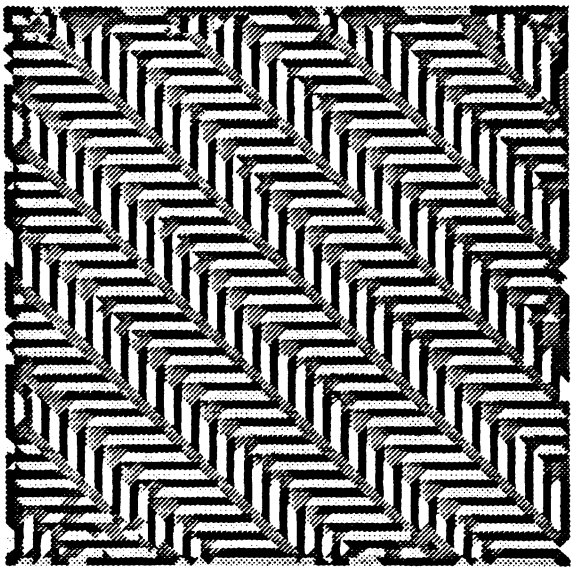
where F_0, F_1 and R are as in the section 2, and

$$\mathbf{n_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{n_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Then we use F_0, RF_1, F_1, and RF_0 as four variants. It follows that F_0 and RF_1 are rank-1 compatible with normal $\mathbf{n_1}$ while F_1 and RF_0 are rank-1 compatible with normal $\mathbf{n_2}$. Now we define the boundary condition by

$$U(\mathbf{x}) = \frac{1}{4}(F_0 + RF_1 + F_1 + RF_0)\mathbf{x} \qquad \mathbf{x} \in \partial\Omega.$$

We rotate the domain by 45° so that four variants appear along boundary evenly. Using a suitable initial approximation we obtained a function with the microstructure shown in Figure 22.



$\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3, k_3 = 10$. Mesh size: 64x64
$E(U^h) = 5.69$

Figure 22

In conclusion, while a triangular mesh avoids the skewing problems caused by square mesh, its orientation can cause problems if it is not reasonably aligned with the desired microstructure. A symmetric triangular mesh is a suitable mesh for microstructures with

several orientations. An alternative way to overcome the mesh orientation problem is to allow wide band microstructures since they can be constructed on any kind of triangular mesh. But the mesh needs to be very fine to insure a such structure, and the fineness is not known a priori.

Appendix A

In this appendix we calculate matrices G_0 and G_1 which are discussed in section 4. The same result can also be derived by the asymptotic expansions [14].

Let B be a given 2x2 constant matrix, which represents the prescribed boundary condition in the context of section 4. Consider two 2x2 constant matrices G_0 and G_1 which are rank-1 connected with normal $\mathbf{m} = (-1,1)^t$ and satisfy $(G_0 + G_1)/2 = B$. Let $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$, and $G_0 = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$. Then it follows that

$$G_1 = \begin{pmatrix} 2b_{11} - u_x & 2b_{12} - u_y \\ 2b_{21} - v_x & 2b_{22} - v_y \end{pmatrix}.$$

Because of the rank-1 connection between G_0 and G_1, we also have

$$G_1 - G_0 = \mathbf{a} \otimes \mathbf{m} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \begin{pmatrix} -1 & 1 \end{pmatrix} = \begin{pmatrix} -a_1 & a_1 \\ -a_2 & a_2 \end{pmatrix},$$

for some vector $\mathbf{a} = (a_1, a_2)^t$. This implies

$$\begin{aligned} (2b_{11} - u_x) - u_x &= -((2b_{12} - u_y) - u_y) \\ (2b_{21} - v_x) - v_x &= -((2b_{22} - v_y) - v_y) \end{aligned}$$

or

$$\begin{aligned} u_y &= b_{11} + b_{12} - u_x \\ v_y &= b_{21} + b_{22} - v_x \end{aligned}$$

Therefore G_0 and G_1 depend only on two independent variables u_x and v_x:

$$G_0 = \begin{pmatrix} u_x & b_{11} + b_{12} - u_x \\ v_x & b_{21} + b_{22} - v_x \end{pmatrix}, \qquad G_1 = \begin{pmatrix} 2b_{11} - u_x & b_{12} - b_{11} + u_x \\ 2b_{21} - v_x & b_{22} - b_{21} + v_x \end{pmatrix}.$$

A continuous piecewise linear vector valued function U^h can be constructed in Mesh B to have the uniform gradient distribution illustrated in Figure 23, where ∇U^h equals G_0 and G_1 in alternate bands. This can be done because G_0 and G_1 are rank-1 connected with the compatible normal and satisfy the boundary condition in an average sense.

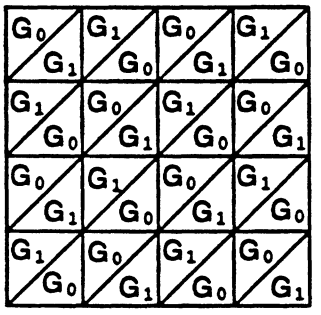


Figure 23

Let $T = W(G_0) + W(G_1)$, where $W(.)$ is defined in section 2. Since G_0 and G_1 are constant matrices, the energy $E(.)$ for U^h in each square is T times the area of each triangle, $h^2/2$. Thus the total energy $E(U^h)$ in the entire domain Ω equals $T/2$. Clearly T also depends on u_x and v_x. So the energy $E(U^h)$ can be minimized by solving the system

$$\frac{\partial T}{\partial u_x} = 0$$
$$\frac{\partial T}{\partial v_x} = 0.$$

Using $\alpha = 5, \beta = 1, k_1 = 10, k_2 = 3$, and $k_3 = 10$ in $W(.)$, we found that the minimum value of $E(U^h)$ equals 9.624747 and it is achieved when

$$G_0 = \begin{pmatrix} 0.4726 & 1.1940 \\ -1.1303 & 1.8756 \end{pmatrix}, \qquad G_1 = \begin{pmatrix} 2.194 & -0.5274 \\ -0.3604 & 1.1058 \end{pmatrix}.$$

Numerically G_0 and G_1 appear to correspond to RF_0 and F_1 in Figure 14.

$$RF_0 = \begin{pmatrix} 0.7454 & 1.4907 \\ -0.6667 & 1.6667 \end{pmatrix}, \qquad F_1 = \begin{pmatrix} 2.2360 & 0.0000 \\ 0.0000 & 1.0000 \end{pmatrix}.$$

Appendix B

This appendix gives an estimate of the optimal width of the bands in the wide band microstructure shown in Figure 15. For more general results see [13].

We want to construct a function in Mesh B (see Figure 11 in section 4) whose gradient distribution has the wide band structure illustrated in Figure 24. Since F_0 and RF_1 are not rank-1 connected with normal $\mathbf{n} = (1, -1)^t$, interfaces with non zero energy must be formed to separate them. The shaded area in Figure 24 indicates a relatively thick boundary strip in

which the gradients are not expected to be on either energy well. This is because a function with one constant gradient matrix in one wide band can not locally match the boundary condition in a way that some components of the gradient of the function approach infinity as the mesh size h goes to zero. So we expect the boundary strip to form in order to reduce the high energy caused by this mismatch.

The construction of this function is as follows. First we make a function with gradients being F_0 and RF_1 alternating in the wide bands. Then the boundary layer of the function within a given distance of the boundary is replaced by a piecewise linear function extended smoothly to the boundary. The energy of this function concentrates in two parts: interfaces and boundary layer. Let b be the thickness of the boundary layer, let w be the width of bands in the horizontal direction and define $n = 1/h$.

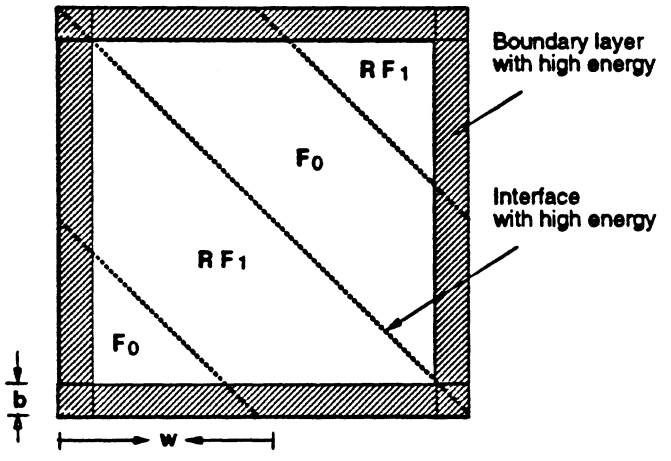


Figure 24

The energy accounting is as follows:

Interface Energy Interfaces are elements separating F_0 and RF_1 bands, hence the gradients of the function in the interface elements are constant. So the energy on each interface element only depends on its area $h^2/2$. The total number of elements in all interfaces is $O(n/w)$. Therefore the total energy of all interfaces E_I is:

$$E_I = O(h^2 \times \frac{n}{w}) = O(\frac{h}{w}).$$

Boundary Layer Energy Let us look at one typical piece of the boundary layer which connects a F_0 band to the bottom boundary. See Figure 25.

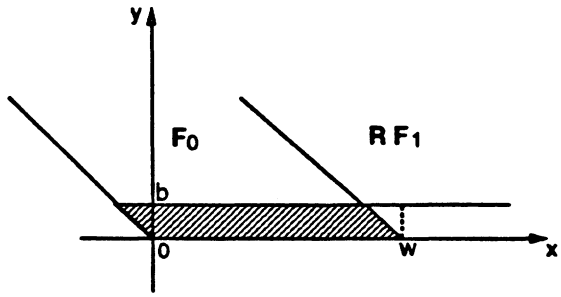


Figure 25

Assume the gradient of the function equals $\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$. Clearly u_x and v_x are bounded, but u_y and v_y depend on the variable x. By construction, $u_y = O(x/b)$ and $v_y = O(x/b)$, where b is the thickness of the boundary layer. Using the energy function defined in section 2, we have

$$E_{piece} = \int_0^b \int_0^w (C_1 \frac{x^4}{b^4} + C_2 \frac{x^2}{b^2} + C_3) dx dy = C_1' \frac{w^5}{b^3} + C_2' \frac{w^3}{b} + C_3' wb.$$

Altogether there are $O(1/w)$ such pieces. So the total contribution from the boundary layer E_B is :

$$E_B = E_{piece} \times O(\frac{1}{w}) = O(\frac{w^4}{b^3} + \frac{w^2}{b} + b).$$

The optimal choice of w and b is that $w = O(b)$, so that the interface energy and boundary energy are balanced. Then

$$E_I = E_B \Rightarrow O(\frac{h}{w}) = O(w) \Rightarrow w = O(\sqrt{h}).$$

In this case, the total energy is $O(\sqrt{h})$. The conclusion is that if both the width of bands and the thickness of the boundary layer are $O(\sqrt{h})$, then the total energy converges to zero like $\sqrt{h}$.

Reference

[1] J.M. Ball and R. James. *Fine phase mixtures as minimizers of energy.* Arch. Rat. Mech. Anal., 100, 15-52, 1987.

[2] J.M. Ball and R. James. *Proposed experimental tests of a theory of fine microstructure and the two well problem.* Phil. Trans. Roy. Soc. Lond., 1991, (to appear).

[3] J. M. Boland and R. A. Nicolaides. *Stable and semistable low over elements for viscous flow.* SIAM J. Num. Anal. 22, p474, 1985.

[4] Collins C. *Computation and Analysis of Twinning in Crystalline Solids.* Ph.D. Thesis, University of Minnesota, Minneapolis, MN, 1990.

[5] C. Collins and M Luskin. *Numerical modeling of the microstructure of crystal with symmetry-related variants.* Proc. ARO US-Japan Workshop on Smart/Intelligent Materials and systems, Techomic.

[6] C. Collins and M Luskin. *The computation of the austenitic-martensitic phase transition.* Lecture Notes in Physics 344, Springer-Verlag,34-50, 1989.

[7] J. L. Ericksen. *Some phase transitions in crystals.* Arch. Rat. Mech. Anal. 73, 99-124, 1980.

[8] J. L. Ericksen. *Twinning of crystals I, Metastability and Incompletely Posed Problem.* IMA Vol. Math. Appl. 3, Springer, 77-96, 1987.

[9] R. D. James. *Microstructure and weak convergence.* Proc. Symp. Material Instabilities in Continuum Mechanics, Heriot-Watt, Oxford, 175-196, 1988.

[10] R. D. James. *Relation between microscopic and macroscopic properties of crystals undergoing phase transformation.* in Proc. 7th Army Conf. on Applied Mathematics and Computing, 1989.

[11] R. D. James and D. Kinderlehrer. *Theory of diffusionless phase transitions.* PDE's and continuum models of phase transitions, Lecture Notes in Physics, 344, Springer, 51-84, 1989.

[12] D. Kinderlehrer. *Remarks about the equilibrium configurations of crystals.* Proc. Symp. Material instabilities in continuum mechanics, Heriot-Watt, Oxford, 217-242, 1988.

[13] Chipot M. and Collins C. *Numerical Approximation in Variational Problems with Potential Wells.* SIAM, J. Number. Anal., 29, 1002-1019, 1992.

[14] L. Ma and N. Walkington. *On Algorithms for Non-Convex Optimization.* Carnegie Mellon University, preprint, 1993.

[15] R. A. Nicolaides and N. Walkington. *Computation of Microstructure Utilizing Young Measure Representations.* Proc. Recent Adv. Adaptive Sensory Materials and Their Appl., Blacksburg, VA., 1992.

[16] James R. and D. Kinderlehrer. *Mathematical Approaches to the Study of Smart Materials.* Proc. of 1993 N. American Conference on Smart Structures and Materials, SPIE, 1993.