

# 1

## Singularities and computation of minimizers for variational problems

J.M. Ball

*Mathematical Institute  
University of Oxford  
24-29 St. Giles'  
Oxford OX1 3LB  
U.K.*

*Email ball@maths.ox.ac.uk*

### Abstract

Various issues are addressed related to the computation of minimizers for variational problems. Special attention is paid (i) to problems with singular minimizers, which natural numerical schemes may fail to detect, and the role of the choice of function space for such problems, and (ii) to problems for which there is no minimizer, which lead to difficult numerical questions such as the computation of microstructure for elastic materials that undergo phase transformations involving a change of shape.

### 1 Introduction

In this article I give a brief tour of some issues related to the computation of minimizers for integrals of the calculus of variations. In this I take the point of view not of a numerical analyst, which I am not, but of an applied mathematician for whom questions of computation have arisen not just because of the need to understand phenomena inaccessible to contemporary analysis, but also because they are naturally motivated by attempts to apply analysis to variational problems.

I will concentrate on two specific issues. The first is that minimizers of variational problems may have singularities, *but natural numerical schemes may fail to detect them*. Connected with this is the surprising *Laurentiev phenomenon*, according to which minimizers in different function spaces may be different. The second issue is that minimizers may not exist, in which case the question naturally arises as to what the behaviour of numerical schemes designed to compute such minimizers will be. In this case the predictive power of the variational problem may still

be retained, for example as an explanatory mechanism for the formation of *microstructure* in materials. A key tool here is the elusive concept of quasiconvexity, which helps to describe the passage from microscales to macroscales.

As a motivating example, consider (nonlinear) elasticity theory. For a homogeneous elastic body the total elastic energy is given by the integral

$$I(y) = \int_{\Omega} W(Dy) dx,$$

where  $W$  is the stored-energy function of the material. Here  $\Omega$  is a bounded open subset of  $\mathbf{R}^3$ , with Lipschitz boundary  $\partial\Omega$ , that the body occupies in a reference configuration, and  $y : \Omega \rightarrow \mathbf{R}^3$  denotes a typical deformation with gradient

$$Dy(x) = \left( \frac{\partial y_i}{\partial x_j} \right).$$

Thus for each  $x$ ,  $Dy(x) \in M^{3 \times 3}$ , where  $M^{m \times n} = \{\text{real } m \times n \text{ matrices}\}$ . In this case the singularities of minimizers could potentially be related to various kinds of fracture or its onset, dislocations, or phase boundaries, while microstructure arises in materials undergoing phase transformations, for which the minimum of  $I$  subject to suitable boundary conditions may not be attained.

## 2 Singular minimizers and the Lavrentiev phenomenon

### 2.1 The Lavrentiev phenomenon and repulsion property

Consider first the simple problem due to Manià [45] of minimizing the integral

$$I(u) = \int_0^1 (u^3 - x)^2 u_x^6 dx \tag{2.1}$$

among absolutely continuous functions  $u$  satisfying the end conditions

$$u(0) = 0, u(1) = 1. \tag{2.2}$$

The unique minimizer of this problem is easily seen to be

$$u^*(x) = x^{\frac{1}{3}}.$$

In fact  $I(u^*) = 0$ , and if  $\bar{u}$  were any other function satisfying the end conditions (2.2) with  $I(\bar{u}) = 0$  then  $\bar{u}_x(x) = 0$  for  $x \in E$  and  $\bar{u}(x) =$

$u^*(x)$  for  $x \in (0, 1) \setminus E$ , where  $E$  has positive one-dimensional Lebesgue measure. Thus

$$0 = \int_0^1 (u_x^* - \bar{u}_x) dx = \int_E \frac{1}{3} x^{-\frac{2}{3}} dx > 0,$$

a contradiction.

Consider now a very natural finite-element scheme for computing the minimizer. Take a uniform mesh subdividing  $[0, 1]$  into  $N$  subintervals of length  $h = 1/N$  and minimize  $I$  among continuous functions satisfying the end conditions (2.2) which are affine on each element  $(i/N, (i+1)/N)$ . For any such function  $v_h$  the integral  $I(v_h)$  can be computed exactly (due to the explicit form of the integrand), so that questions of quadrature can in the first instance be ignored. For each  $h$  there is at least one minimizer  $u_h^*$  to this discrete problem. What is the behaviour of  $u_h^*$  as  $h \rightarrow 0$ ? Remarkably,  $u_h^*$  converges as  $h \rightarrow 0$ , but not to the minimizer  $u^*$ ! In fact the limit  $u_0$  is a monotone increasing function that is smooth in  $(0, 1)$  but has infinite slope at the end-points  $x = 0, 1$ .

This behaviour is hard to credit at a first glance. An illuminating initial calculation is to compute  $I(u_h)$  for the function

$$u_h(x) = \begin{cases} h^{-\frac{2}{3}} x & \text{if } x \in (0, h) \\ u^*(x) & \text{if } x \in (h, 1) \end{cases}$$

in which  $u^*$  is altered only on the first element. Surely  $\lim_{h \rightarrow 0} I(u_h) = I(u^*) = 0$ ? But no,

$$\int_0^1 (u_h^3 - x)^2 u_{hx}^6 dx = \frac{8}{105} h^{-1},$$

which tends to  $+\infty$  as  $h \rightarrow 0$ !

In fact it can be shown (see Ball & Knowles [12]) that if  $1 \leq p \leq \infty$  and

$$\mathcal{A}_p = \{v \in W^{1,p}(0, 1) : v(0) = 0, v(1) = 1\}$$

(so that  $\mathcal{A}_1$  is the admissible class of functions considered above) then

$$\inf_{\mathcal{A}_\infty} I = \min_{\mathcal{A}_{3/2}} I > \min_{\mathcal{A}_1} I = 0.$$

The fact that the infimum of  $I$  in different function spaces can be different is known as the *Lavrentiev phenomenon* (see Lavrentiev [37] for the original example). The initial calculation above has the following generalization, let us call it the *repulsion property*, that if  $u^{(j)} \in \mathcal{A}_{3/2}$  and  $u^{(j)} \rightarrow u^*$  a.e. in  $(0, 1)$  then  $I(u^{(j)}) \rightarrow \infty$ .

In the Manià example the integrand  $f(x, u, p) = (u^3 - x)^2 p^6$  is convex in  $p$ , but not strictly convex. However, as was shown by Ball & Mizel [13, 14], the Lavrentiev phenomenon and the repulsion property hold for elliptic integrands, i.e. those for which  $f_{pp}(x, u, p) \geq \mu > 0$  for all  $x, u, p$ . Such an example is given by the problem [14] of minimizing

$$I(u) = \int_{-1}^1 [(x^4 - u^6)^2 u_x^{28} + \varepsilon u_x^2] dx \quad (2.3)$$

in  $\mathcal{A}_p = \{v \in W^{1,p}(-1, 1) : v(-1) = -1, v(1) = 1\}$ . Note that the integrand  $f(x, u, p) = (x^4 - u^6)^2 p^{28} + \varepsilon p^2$  satisfies  $f_{pp} \geq 2\varepsilon > 0$ . Here, for sufficiently small  $\varepsilon > 0$ , there is an absolute minimizer  $u^*$  of  $I$  in  $\mathcal{A}_1$  that is a smooth solution of the Euler-Lagrange equation in  $[-1, 0) \cup (0, 1]$  but has derivative  $+\infty$  at  $x = 0$ , where  $u^*(x) \sim |x|^{\frac{2}{3}} \text{sign } x$ . The Lavrentiev phenomenon holds in the form

$$\inf_{\mathcal{A}_\infty} I = \min_{\mathcal{A}_3} I > \inf_{\mathcal{A}_1} I = I(u^*). \quad (2.4)$$

As indicated in (2.4),  $I$  attains its infimum in  $\mathcal{A}_3$ , and every such minimizer is a smooth solution of the Euler-Lagrange equation in the whole interval  $[-1, 1]$ . The repulsion property also holds in the form that if  $u^{(j)} \in \mathcal{A}_3$  with  $u^{(j)} \rightarrow u^*$  a.e. then  $I(u^{(j)}) \rightarrow \infty$ .

## 2.2 Computation of singular minimizers

What are possible numerical methods for detecting such singular minimizers? Consider the problem of minimizing

$$I(u) = \int_a^b f(x, u, u_x) dx$$

in

$$\mathcal{A}_1 = \{u \in W^{1,1}(a, b) : u(a) = \alpha, u(b) = \beta\},$$

where  $\alpha, \beta$  are given constants.

A first method proposed by Ball & Knowles [12] consists in decoupling  $u$  from its derivative. Thus given  $\varepsilon > 0$  we minimize

$$I(u, v) = \int_a^b f(x, u, v) dx$$

among piecewise affine functions  $u$  in  $\mathcal{A}_1$  on a uniform mesh of size  $h$ ,

and functions  $v$  that are piecewise constant on the same grid, subject to the constraint

$$\int_a^b \varphi(u_x - v) dx \leq \varepsilon,$$

where  $\varphi \geq 0$  is a suitable even continuous function satisfying (i)  $\varphi(p) \geq |v|^s$  for all  $v \in \mathbf{R}$ , where  $1 \leq s < \infty$ , (ii)  $\varphi(p_1 + p_2) \leq C(\varphi(p_1) + \varphi(p_2))$  for all  $p_1, p_2 \in \mathbf{R}$ . Then under suitable growth and convexity hypotheses on  $f$  (in particular guaranteeing that the infimum of  $I(u)$  in  $\mathcal{A}_1$  is attained) it can be shown that minimizers  $\{u_{h,\varepsilon}, v_{h,\varepsilon}\}$ , with  $h < \gamma(\varepsilon)$  for a suitable function  $\gamma$ , converge to minimizers  $\{u^*, u_x^*\}$  of  $I$  in  $\mathcal{A}_1$ , possibly after extraction of a subsequence, and that

$$\lim_{h,\varepsilon \rightarrow 0, 0 < h < \gamma(\varepsilon)} I(u_{h,\varepsilon}, v_{h,\varepsilon}) = \inf_{u \in \mathcal{A}_1} I(u) = I(u^*).$$

The effect of numerical quadrature is also studied in [12]; in fact for the Manià example a direct numerical minimization of (2.1) among piecewise affine functions in  $\mathcal{A}_1$  using the trapezoidal rule succeeds in finding  $u^*$ , but this is a freak resulting from the special form of the integrand and in particular its degeneracy when  $u^3 = x$ .

A second idea (see Li [43]) is the *truncation method*, in which  $f = f(x, u, p)$  is replaced by a truncated integrand  $f_M(x, u, p)$  satisfying  $f_M(x, u, p) = f(x, u, p)$  whenever  $|p| \leq M$ , and  $f_M \rightarrow f$  monotonically as  $M \rightarrow \infty$ . If  $f_M$  has suitable mild growth properties as  $|p| \rightarrow \infty$ , ensuring in particular that the truncated integral

$$I_M(u) = \int_a^b f_M(x, u, u_x) dx$$

does not have the Lavrentiev phenomenon, then we can first minimize  $I_M$ , and then let  $M \rightarrow \infty$ . Although this method works theoretically, it has practical drawbacks in that it may not be easy to find an absolute minimizer of  $I_M$  (for example, in the case when  $I$  attains a minimum among smooth functions, as for the integral (2.3), there is the danger of finding this minimizer instead, since it is a local minimizer of  $I_M$  for all sufficiently large  $M$ ).

A more interesting method (see Li [41]) is that of *element removal*. Here we use piecewise affine approximations, but for each  $h$  minimize

$$\int_{[a,b] \setminus E} f(x, u, u_x) dx$$

where  $E$  consists of a (controlled) small number of elements (in practice

these turn out to be elements where  $u_x^*$  is large). As for the method in [12] the number of unknowns is increased, in this case by variables tracking which elements are removed.

A potentially promising method, which as far as I am aware has not been studied, is that of using nonconforming elements. It seems possible that this could lead to ways of detecting minimizers in a continuous scale of Sobolev spaces.

There is a growing literature on singular minimizers of one-dimensional variational problems. These singular minimizers *do not in general satisfy the Euler-Lagrange equation in weak or integrated form on the whole interval*  $[a, b]$ . For elliptic integrands the Euler-Lagrange equation is satisfied on the complement of a closed set of Lebesgue measure zero. This is part of the content of the Tonelli partial regularity theorem [60]II p. 359. This theorem is shown to be optimal in [14] (where a slightly improved version can be found), and in Davie [29]. The singularities of minimizers can also be studied in the  $(x, u)$  plane; it was shown by Ball & Nadirashvili [16] that under natural hypotheses on  $f$  there is a *universal singular set*  $\mathcal{D}_f$  to which all points  $(x, u(x))$  with  $|u_x(x)| = \infty$  for minimizers  $u$  of  $I$  in  $\mathcal{A}_1$  for any  $a, b, \alpha, \beta$  must belong, and that  $\mathcal{D}_f$  is of first category in  $\mathbf{R}^2$ . Later Sychev [58] proved that  $\mathcal{D}_f$  has two-dimensional Lebesgue measure zero.

There is an important philosophical consequence of the above discussion. Suppose for a moment that one of the variational integrals (2.1), (2.3) represented the energy of some physical system (of course this is not the case, but I will give a physical example later). Since minimizers in different function spaces can be different, in other words different function spaces lead to different predictions, it follows that *the function space is part of the model*. This conclusion seems to me inescapable, but is an uncomfortable one in the sense that little attention is traditionally paid to function spaces when deriving mathematical models (an interesting exception is quantum mechanics, in which the underlying Hilbert space is introduced at the foundational level). If we accept it, then the next question is where the function space (for example, for a model of continuum physics) should come from? To the extent that it is made explicit, common practice is to adopt a pragmatic attitude to this question, making a partly phenomenological choice based on the experimentally observed singularities and the form of natural expressions such as energy that appear in the theory. A more satisfactory approach would be to derive the function space (as well as the governing equations) from a more detailed (e.g. atomistic) model. The only example that I am familiar

with where this is done is in the paper of Braides, Dal Maso & Garroni [18], where a one-dimensional model for softening phenomena in fracture mechanics is derived from a (primitive) atomistic model complete with a function space (the space BV of functions of bounded variation).

As was pointed out to me by J.F. Traub at the Oxford FoCM conference, the issue of the choice of a function space arises in continuous complexity theory, for example in analyzing the complexity of the problem of computing the integral of a function (see, for example, Traub & Werschütz [61]), where some *a priori* hypothesis has to be made about the regularity of the function to be integrated. It would be interesting to analyze the complexity of computation of integrals of the calculus of variations, and of the problem of minimization of such integrals, in the light of the Lavrentiev phenomenon.

### 3 Nonlinear elasticity

As described in the introduction, the total elastic energy of a homogeneous elastic body has the form

$$I(y) = \int_{\Omega} W(Dy) dx. \quad (3.1)$$

Consider the problem of minimizing  $I$  among deformations  $y$  satisfying the boundary condition

$$y|_{\partial\Omega_1} = \bar{y}, \quad (3.2)$$

where  $\partial\Omega_1 \subset \partial\Omega$  has positive area and where  $\bar{y} : \partial\Omega_1 \rightarrow \mathbf{R}^3$  is a given measurable mapping. No condition is specified on the remainder of the boundary  $\partial\Omega \setminus \partial\Omega_1$ , where minimization leads to a natural boundary condition corresponding to zero applied traction.

To be physically meaningful, the deformation  $y$  should be invertible on  $\Omega$ . (This is another requirement on the function space that we could ask to be the consequence of a derivation of (3.1) from a more detailed model.) In particular this leads to the requirement that

$$\det Dy(x) > 0 \quad \text{for a.e. } x \in \Omega. \quad (3.3)$$

To guarantee (3.3) for deformations of finite energy, suppose that  $W : M_+^{3 \times 3} \rightarrow \mathbf{R}$ , where  $M_+^{3 \times 3} = \{A \in M^{3 \times 3} : \det A > 0\}$ , with

$$W(A) \rightarrow \infty \quad \text{as } \det A \rightarrow 0^+. \quad (3.4)$$

If  $\partial\Omega_1 \neq \partial\Omega$  there is the possibility of self-contact of the boundary (for

a treatment of this see Ciarlet & Necas [24]), and for this reason  $y$  need not be invertible on the closure  $\bar{\Omega}$  of  $\Omega$ . However, even if  $\partial\Omega_1 = \partial\Omega$  the requirement of invertibility leads to difficulties. Consider for example the mapping  $u : D \rightarrow \mathbf{R}^2$ , where  $D$  is the unit disk of  $\mathbf{R}^2$ , given in plane polar coordinates by  $(r, \theta) \mapsto (\frac{1}{\sqrt{2}}r, 2\theta)$ . It is easily seen that  $u$  is Lipschitz with  $\det Du(x) = 1$  a.e., but  $u$  is not locally invertible at 0. Thus for mappings in Sobolev spaces local invertibility does not follow from (3.3). This difficulty can be overcome by an appropriate use of degree theory (see [4], Šverák [55], Fonseca & Gangbo [33]).

The problem of numerical minimization of  $I$  via finite elements leads naturally to the

*Open question.* If  $y \in W^{1,p}$  is invertible, can  $y$  be approximated in  $W^{1,p}$  by piecewise affine invertible mappings?

Here  $W^{1,p} = W^{1,p}(\Omega; \mathbf{R}^n)$ ,  $\Omega \subset \mathbf{R}^n$ ,  $n \geq 2$ . This question is also of considerable theoretical interest, and I first heard of it from L.C. Evans [31] in the context of his attempts to prove a version of his partial regularity theorem [32] for quasiconvex integrals that would be valid for elastic energies satisfying (3.4). He remarked to me that the existing literature on simplicial approximation (see e.g. [46]) did not cover the case of mappings in Sobolev spaces, since the techniques used relied on composition of mappings, and mappings in Sobolev spaces are not closed under composition.

Consider the simplest case  $n = 2$ . For a continuous  $y \in W^{1,p}$  with  $\det Dy(x) > 0$  a.e. a natural algorithm is to triangulate  $\Omega$  with a regular mesh of size  $h$  and define the approximating mapping  $y^h$  to be that piecewise affine mapping coinciding with  $y$  at the mesh points. Unfortunately this fails because even if  $y \in W^{1,\infty}$  with  $\det Dy(x) \geq \sigma > 0$  there can be, for arbitrarily small  $h$ , triangles on which  $\det Dy^h$  is negative. An algorithm is needed for choosing a sequence of finer and finer meshes without this undesirable behaviour.

The existence theory for minimizers of (3.1) has been reviewed in many places (see, for example, [7, 8], Ciarlet [23], Dacorogna [28], Šilhavý [63], Pedregal [51]). For the existence of minimizers it is necessary to impose growth and convexity conditions on the stored-energy function  $W$ . The natural convexity condition is that of *quasiconvexity*. In fact for the variational problem of minimizing

$$I(y) = \int_{\Omega} f(Dy) dx$$



subject to

$$y|_{\partial\Omega_1} = \bar{y},$$

where  $\Omega \subset \mathbf{R}^n$  is a bounded Lipschitz domain,  $y : \Omega \rightarrow \mathbf{R}^m$  and  $f : M^{m \times n} \rightarrow \mathbf{R}$  is continuous and satisfies suitable growth conditions, it is now understood (Morrey [47], Acerbi & Fusco [1]) that for the existence of minimizers it is sufficient (and, up to the addition of lower-order terms, necessary, see Ball & Murat [15]) that  $f$  be quasiconvex.

**Definition 3.1** *Let  $f : M^{m \times n} \rightarrow \mathbf{R} \cup \{+\infty\}$  be continuous. Then  $f$  is quasiconvex if*

$$\int_{\Omega} f(Dv) \, dx \geq \int_{\Omega} f(A) \, dx$$

for all  $A \in M^{m \times n}$  and all  $v \in Ax + C_0^\infty(\Omega; \mathbf{R}^m)$ .

(This condition seems to, but does not, depend on  $\Omega$ .) No tractable necessary and sufficient conditions are known for a function  $f$  to be quasiconvex. If  $f$  is quasiconvex then  $f$  is *rank-one convex*, that is the mapping  $t \mapsto W(A + t\lambda \otimes \mu)$  is convex for all  $A \in M^{m \times n}, \lambda \in \mathbf{R}^m, \mu \in \mathbf{R}^n$ , but it was shown by Šverák [56] that the converse is false for  $n \geq 2, m \geq 3$ . Based on Šverák's example, Kristensen [36] proved the striking result that for the same dimensions there is no local necessary and sufficient condition for quasiconvexity.

Unfortunately, the known existence theorems for quasiconvex integrands do not apply to elasticity, because they assume growth conditions incompatible with (3.4). For this reason it is at present necessary (see the references cited above), to make the stronger convexity hypothesis that  $W$  be *polyconvex*, namely

$$W(A) = g(A, \operatorname{cof} A, \det A)$$

for some convex  $g$ , where  $\operatorname{cof} A$  denotes the matrix of  $2 \times 2$  subdeterminants of  $A$ . Then the existence of an absolute minimizer is assured, provided the growth condition

$$W(A) \geq c_0(|A|^2 + |\operatorname{cof} A|^{\frac{3}{2}}) - c_1 \tag{3.5}$$

holds, where  $c_0 > 0, c_1$  are constants (see Müller, Qi & Yan [49] for this improved version of a result of [3]). However essentially nothing is known about the smoothness of absolute minimizers  $y^*$  for strictly polyconvex or strictly quasiconvex  $W$ . It is not even known if the usual weak form

of the Euler-Lagrange equation is satisfied (though certain weak forms may be obtained, see [5, 2], Bauman, Owen & Phillips [17]), or if

$$\det Dy^*(x) \geq \sigma > 0 \text{ for a.e. } x \in \Omega. \quad (3.6)$$

Of course a proof of smoothness and of (3.6) would lead to a justification of standard finite-element minimization schemes for (3.1), (3.2). Otherwise the above Open Problem makes the construction of a scheme generating invertible approximate minimizers problematic. Another approach is to tolerate a small set on which the approximate minimizers fail to satisfy (3.3). This approach is taken in Li [42], who applies the element removal method to find a sequence of such (possibly noninvertible) approximate minimizers converging, at least theoretically, to a minimizer of (3.1),(3.2).

Under the hypotheses of the existence theorem it is not known whether the Lavrentiev phenomenon can hold. However if the growth condition (3.5) is slightly weakened then there is a physically interesting example involving *cavitation*.

As an illustrative example of cavitation consider the problem of minimizing

$$I(y) = \int_{B(0,1)} W(Dy) dx$$

subject to the pure displacement boundary condition

$$y|_{\partial B(0,1)} = \lambda x, \quad \lambda > 0,$$

where  $B(0,1)$  denotes the unit ball in  $\mathbf{R}^3$ , and where

$$W(A) = |A|^2 + h(\det A),$$

with  $h : (0, \infty) \rightarrow \mathbf{R}$  smooth and satisfying  $h'' > 0$ ,  $\lim_{\delta \rightarrow \infty} \frac{h(\delta)}{\delta} = \lim_{\delta \rightarrow 0^+} h(\delta) = \infty$ . Note that (3.5) does not hold, but that  $W$  is polyconvex. Hence, since polyconvexity implies quasiconvexity, the minimizer of  $I$  among smooth (or even  $W^{1,3}$ )  $y$  is given by

$$\tilde{y}_\lambda(x) \equiv \lambda x.$$

But among radial maps

$$y(x) = r_\lambda(|x|) \frac{x}{|x|}$$

we have nontrivial minimizers for  $\lambda > \lambda_{cr}$  for some critical value  $\lambda_{cr}$ .

These radial minimizers satisfy  $r_\lambda(0) > 0$ . Thus a hole is formed at the origin. Furthermore we have the Lavrentiev phenomenon in the form

$$\inf_{\mathcal{A}_1} I < \inf_{\mathcal{A}_3} I = I(\tilde{y}_\lambda),$$

where  $\mathcal{A}_p = \{y \in W^{1,p}(B(0,1); \mathbf{R}^3) : y|_{\partial B(0,1)} = \lambda x\}$ . In fact cavitation is a common failure mechanism in polymers. See Lazzeri & Bucknall [38] for some striking images of almost radial cavitation of roughly spherical rubber particles imbedded in a matrix of nylon-6; such rubber-toughened plastics are used, for example, in car bumpers. See [7] for further remarks about cavitation and function spaces.

## 4 Computation of microstructure

### 4.1 Nonattainment of minimum energy and microstructure

Consider a single crystal of a material (for example, some metallic alloy) that can undergo a phase transformation involving a change of shape at some critical temperature  $\theta = \theta_c$  from a higher symmetry *austenite* phase to a lower symmetry *martensite* phase. The crystal is assumed to be elastic with stored-energy function  $W_\theta(Dy)$  that depends on the temperature  $\theta$ . If  $W$  attains a finite minimum, then by adding a suitable function of  $\theta$  there is no loss of generality in assuming that

$$\min_A W_\theta(A) = 0.$$

Consider the corresponding set of energy-minimizing gradients

$$K_\theta = \{A \in M^{3 \times 3} : W_\theta(A) = 0\}.$$

Since the stored-energy function must satisfy the frame-indifference condition

$$W_\theta(QA) = W_\theta(A) \quad \text{for all } Q \in SO(3)$$

it follows that  $K_\theta = SO(3)K_\theta$ . Let  $U = U^T > 0$  be the linear transformation describing the change of shape at  $\theta = \theta_c$  relative to undistorted austenite. Then at  $\theta = \theta_c$

$$K_{\theta_c} = SO(3) \cup \bigcup_{i=1}^N SO(3)U_i,$$

where the  $U_i$  are the distinct matrices  $RUR^T$  for  $R$  belonging to the symmetry group  $\mathcal{S}$  of the material (assumed to be a subgroup of  $SO(3)$ ). The energy well  $SO(3)$  corresponds to the austenite, while each energy

well  $SO(3)U_i$  corresponds to one of the  $N$  variants of the martensite. If, as is often the case, the austenite is stable for temperatures  $\theta > \theta_c$  then for these temperatures  $K_\theta = \alpha(\theta)SO(3)$ , where  $\alpha(\cdot)$  accounts for thermal expansion, while, for  $\theta < \theta_c$ ,  $K_\theta$  is given by the  $N$  martensitic variants, with  $U = U(\theta)$ .

An important example is provided by a cubic to tetragonal transformation. Here  $N = 3$ , and the three variants correspond to the matrices

$$\begin{aligned} U_1 &= \text{diag}(\eta_2, \eta_1, \eta_1), \\ U_2 &= \text{diag}(\eta_1, \eta_2, \eta_1), \\ U_3 &= \text{diag}(\eta_1, \eta_1, \eta_2), \end{aligned}$$

where  $\eta_1, \eta_2$  are lattice parameters.

Interfaces between variants are described by rank-one connections between the corresponding energy wells  $SO(3)U_i, SO(3)U_j$ , i.e. by pairs of matrices  $R_iU_i, R_jU_j$  with

$$R_iU_i - R_jU_j = \lambda \otimes \mu,$$

where  $i \neq j$  and  $R_i, R_j \in SO(3)$ , and where  $\mu$  is the interface normal. Such rank-one connections exist between any pair of the three tetragonal wells in a cubic to tetragonal transformation. Given such a rank-one connection, the function  $t \mapsto W_\theta(R_iU_i + t\lambda \otimes \mu)$  has a double-well form, and therefore is not convex. Hence  $W_\theta$  is not rank-one convex, and so not quasiconvex either, leading to the expectation that the minimum of

$$I_\theta(y) = \int_{\Omega} W_\theta(Dy) dx \quad (4.1)$$

in  $W^{1,1}(\Omega; \mathbf{R}^3)$  subject to the boundary condition

$$y|_{\partial\Omega_1} = \bar{y}, \quad (4.2)$$

is in general *not attained*. In fact this has been proved in certain cases (see Ball & James [11]Theorem 7.1, Ball & Carstensen [9]Theorems 3.1, 3.2). Minimizing sequences  $y^{(j)}$  typically have gradients  $Dy^{(j)}$  that oscillate more and more finely as  $j$  increases, generating in the limit an infinitely fine *microstructure*.

Real microstructures are not of course infinitely fine, and their limited fineness can be modelled by introducing interfacial energy. For example, a crude way of doing this is to change the energy functional to

$$I_\theta^\varepsilon(y) = \int_{\Omega} [W_\theta(Dy) + \frac{1}{2}\varepsilon^2|D^2y|^2] dx, \quad (4.3)$$

for some small  $\varepsilon > 0$ .

I refer the reader to Ball & James [10, 11], Luskin [44], Müller [48], Pedregal [51] for further details concerning the physical model and its analysis.

#### 4.2 What should we compute and how?

According to the elasticity model described above, minimizing sequences for the total elastic energy may develop infinitely fine microstructure. Thus, however fine a finite-element mesh is used, numerical minimization of the energy can be expected to yield oscillations in  $Dy$  at a length-scale comparable with the mesh size. For a model such as (4.3) incorporating interfacial energy a very fine mesh is still needed to capture the details of the microstructure, which in real materials can have a length-scale of as little as a few atomic spacings. Thus numerical minimization of the energy is computationally highly intensive.

A second difficulty is that, because oscillations can develop at the level of the mesh, and the problem has preferred crystallographic directions, the computations will in general be sensitive to mesh orientation.

A third difficulty is that the discretized energy has a huge number of local minimizers. Consider, for example, the one-dimensional problem of minimizing

$$I(u) = \int_0^1 [(u_x^2 - 1)^2 + u^2] dx$$

in  $W^{1,1}(0,1)$  (so that there are no end conditions). Then

$$\inf_{W^{1,1}} I = 0$$

but the infimum is not attained. Consider the discretized problem of minimizing  $I$  among piecewise affine functions on a mesh of size  $h = \frac{1}{N^2}$ . Let  $u_h$  be a minimizer and set  $I(u_h) = E_h$ . Then (Carstensen [19]) there exists a family  $\mathcal{K}$  consisting of  $N^N$  local minimizers of the same discretized problem, such that

- (i)  $\mathcal{K}$  is a subset of the ball in  $L^2(0,1)$  with centre  $u_h$ , radius  $5h$ ,
- (ii)  $I(v) < (1 + 24\sqrt{h})E_h$  for each  $v \in \mathcal{K}$ ,
- (iii) if  $v_0, v_1$  are distinct points of  $\mathcal{K}$  then  $\sup_{t \in [0,1]} I(v(t)) > \frac{13}{h}E_h$  for any continuous path  $v : [0,1] \rightarrow L^2(0,1)$  with  $v(0) = v_0, v(1) = v_1$ .

Because of the exponential number of local minimizers, and the relatively high energy thresholds between them, local descent methods will typically fail to detect a global minimizer of the discretized problem.

But what should we compute? For the problem (4.1),(4.2) (similar issues arise for problems incorporating small interfacial energy) one possible answer is the set of possible *Young measures*  $(\nu_x)_{x \in \Omega}$  corresponding to sequences of deformation gradients  $Dy^{(j)}$  for minimizing sequences  $y^{(j)}$  (for the definition and properties of these measures see, for example, [6, 59, 48, 50, 62]). For each  $x \in \Omega$ ,  $\nu_x$  is a probability measure on  $M^{3 \times 3}$ . The Young measure determines through the formula  $Dy(x) = \bar{\nu}_x$ , where  $\bar{\nu}_x = \int_{M^{3 \times 3}} A d\nu_x(A)$  denotes the centre of mass of  $\nu_x$ , the *macroscopic* deformation gradient  $Dy$ , where  $y$  is the weak limit of the minimizing sequence  $y^{(j)}$  in an appropriate Sobolev space determined by the growth of  $W$ .

In principle  $y$  can be computed by minimizing

$$I_\theta^{\text{qc}}(y) = \int_{\Omega} W_\theta^{\text{qc}}(Dy) dx, \quad (4.4)$$

subject to (4.2), where  $W_\theta^{\text{qc}}$  is the *quasiconvex envelope* of  $W_\theta$ , that is the supremum of all quasiconvex functions that are less than  $W_\theta$ . This is the content of the relaxation theorem of Dacorogna [27], though as for the existence theorems assuming quasiconvexity, the theorem does not strictly speaking apply to elasticity because the growth hypotheses are inconsistent with the property  $W_\theta(A) \rightarrow \infty$  as  $\det A \rightarrow 0^+$ . Thus quasiconvexification describes the passage from microscopic to macroscopic stored-energy functions for these materials.

The idea of minimizing (4.4) to compute  $y$  is attractive, but a very serious drawback is that the lack of a suitable characterization of quasiconvexity means that it is only in rare cases (see, for example, Kohn [35], Pipkin [52, 53], LeDret & Raoult [39]) that  $W_\theta^{\text{qc}}$  is known. Further, to compute  $W_\theta^{\text{qc}}$  numerically leads to a problem of similar difficulty to the original one.

Despite all these difficulties there have been a number of interesting computations of martensitic microstructure, though it is fair to say that there is a long way to go before the computer can be used as an effective predictive tool in these problems. Some key references are Carstensen and Plecháč [21, 20, 22], Collins, Kinderlehrer & Luskin [25], Collins & Luskin [26], Dolzmann [30], Killough [34], Li & Luskin [40] (these last two papers concerning computations of needle-like martensitic microstructures), and the review article of Luskin [44].

As an illustration of what can be achieved, in Figure 1 are shown the results of some computations due to P. Plecháč [54]. These computations were carried out for a two-dimensional version of  $I_\theta^\varepsilon$  (see (4.3)) with

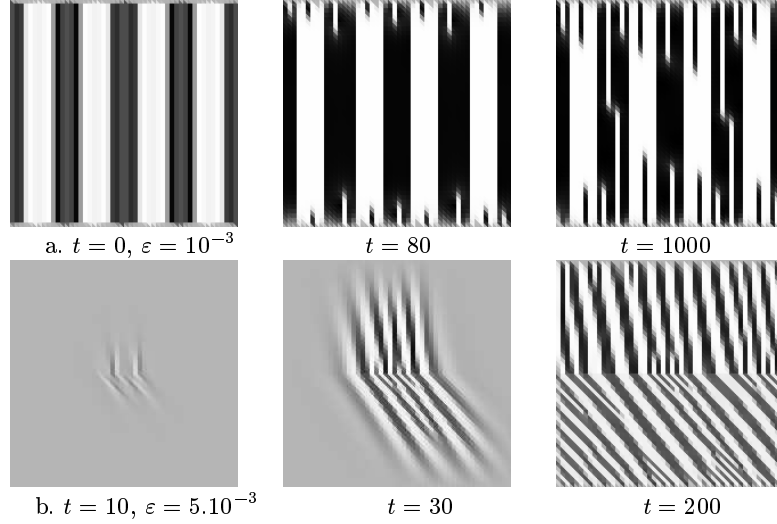


Fig. 1. Microstructure evolution a. for a single crystal b. for two adjacent grains.

$W = W_\theta : M^{2 \times 2} \rightarrow \mathbf{R}$  given by

$$W(A) = \kappa_1(C_{11} - (1 + \delta^2))^2 + \kappa_2(C_{22} - 1)^2 + \kappa_3(C_{12}^2 - \delta^2)^2, \quad (4.5)$$

where  $C = A^T A$ . In this case we have that

$$K = \{A \in M^{2 \times 2} : W(A) = 0\} = SO(2)U_1 \cup SO(2)U_2,$$

where

$$U_1 = \begin{pmatrix} 1 & 0 \\ -\delta & 1 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix}.$$

The two energy wells are rank-one connected with

$$U_2 - U_1 = \lambda \otimes \mu,$$

where  $\lambda = (0, 2\delta)$ ,  $\mu = (1, 0)$ . In the computations,  $\Omega = (0, 1)^2$ , the boundary condition was taken to be linear with gradient  $\frac{1}{2}(U_1 + U_2)$ , that is

$$y(x) = x \text{ for } x \in \partial\Omega,$$

and the constants in (4.5) had the values  $\kappa_i = 1, \delta = 0.5$ . For these boundary conditions with  $\varepsilon = 0$  the minimum of the energy (4.3) is not

attained, and the Young measure of any minimizing sequence is given by

$$\nu_x = \frac{1}{2}(\delta_{U_1} + \delta_{U_2}).$$

Similarly, for small  $\varepsilon$  we expect a finely layered configuration to evolve in which  $Dy$  mostly takes the values  $U_1, U_2$ . The minimization algorithm was to solve the evolution equation

$$\Delta y_t = -\operatorname{div} \sigma(Dy) + \varepsilon^2 \Delta^2 y,$$

where  $\sigma(A) = DW(A)$ . (This equation, used in the calculations as a numerical solver, can also be thought of as a viscoelastic model for the evolution of  $y$  that neglects inertia; however it should be noted that the damping term  $\Delta y_t$  is not frame-indifferent.) The time-discretization was a fully implicit scheme, approximating the  $H^{-1}$ -gradient flow for  $I_\theta^\varepsilon$  by a corresponding minimization problem at each time step. The mesh was uniform with  $500 \times 500$  elements. This algorithm is of course not guaranteed to tend as time  $t \rightarrow \infty$  to a global minimizer. In Figure 1a are shown stages in the evolution starting from sinusoidal initial data, with  $\varepsilon = 10^{-3}$ , the shading varying from white when  $Dy(x) \in SO(2)U_1$  to black when  $Dy(x) \in SO(2)U_2$ . Note how the refinement of the layering is via the initiation of new layers at the boundary; a similar effect was observed earlier in the computations of Swart (see Swart & Holmes [57]).

For the computation in Figure 1b the crystal consists of two adjacent grains, corresponding to the stored-energy function  $W = W(x, A)$  given by

$$W(x, A) = \begin{cases} W(A) & x \in \Omega_1 = (0, 1) \times (1/2, 1) \\ W(AR_{\pi/4}) & x \in \Omega_2 = (0, 1) \times (0, 1/2) \end{cases},$$

where  $R_{\pi/4}$  denotes a planar rotation through the angle  $\pi/4$ , and  $\varepsilon = 5 \cdot 10^{-3}$ . The initial data was taken to be the identity with a small smooth perturbation in a neighbourhood of the centre of the square. Two simple laminates are formed meeting at the grain boundary.

### Acknowledgments

I am grateful to Petr Plecháč for permission to include his computations and to him, Carsten Carstensen, Craig Evans, Matt Killough, Bob Kohn and Joe Traub for useful comments that helped improve the paper and lecture at the 1999 Oxford FoCM conference on which it was based. This



work was supported by EC TMR Contract ERBFMRX CT 98-229 on 'Phase transformations in crystalline solids'.

### Bibliography

- [1] E. Acerbi and N. Fusco. Semicontinuity problems in the calculus of variations. *Arch. Rat. Mech. Anal.*, 86:125–145, 1984.
- [2] J.M. Ball. Some open problems in elasticity. To appear.
- [3] J.M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rat. Mech. Anal.*, 63:337–403, 1977.
- [4] J.M. Ball. Global invertibility of Sobolev functions and the interpenetration of matter. *Proc. Royal Soc. Edinburgh*, 88A:315–328, 1981.
- [5] J.M. Ball. Minimizers and the Euler-Lagrange equations. In *Proc. ISIMM conference, Paris*. Springer-Verlag, 1983.
- [6] J.M. Ball. A version of the fundamental theorem for Young measures. In M. Rascle, D. Serre, and M. Slemrod, editors, *Proceedings of conference on 'Partial differential equations and continuum models of phase transitions'*, pages 3–16. Springer Lecture Notes in Physics. No. 359, 1989.
- [7] J.M. Ball. Nonlinear elasticity and materials science; a survey of some recent developments. In P.J. Aston, editor, *Nonlinear Mathematics and Its Applications*, pages 93–119. Cambridge University Press, 1996.
- [8] J.M. Ball. The calculus of variations and materials science. *Quart. Appl. Math.*, 56:719–740, 1998.
- [9] J.M. Ball and C. Carstensen. Compatibility conditions for microstructures and the austenite-martensite transition. *Materials Science & Engineering A*, 273-275:231–236, 1999.
- [10] J.M. Ball and R.D. James. Fine phase mixtures as minimizers of energy. *Arch. Rat. Mech. Anal.*, 100:13–52, 1987.
- [11] J.M. Ball and R.D. James. Proposed experimental tests of a theory of fine microstructure, and the two-well problem. *Phil. Trans. Roy. Soc. London A*, 338:389–450, 1992.
- [12] J.M. Ball and G. Knowles. A numerical method for detecting singular minimizers. *Numerische Math.*, 92:193–204, 1986.
- [13] J.M. Ball and V.J. Mizel. Singular minimizers for regular one-dimensional problems in the calculus of variations. *Bull. Amer. Math. Soc.*, 11:143–146, 1984.
- [14] J.M. Ball and V.J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equations. *Arch. Rat. Mech. Anal.*, 90:325–388, 1985.
- [15] J.M. Ball and F. Murat.  $W^{1,p}$ -quasiconvexity and variational problems for multiple integrals. *J. Functional Analysis*, 58:225–253, 1984.
- [16] J.M. Ball and N. Nadirashvili. Universal singular sets for one-dimensional variational problems. *Calculus of Variations and Partial Differential Equations*, 1:429–438, 1993.
- [17] P. Bauman., N.C. Owen, and D. Phillips. Maximum principles and a priori estimates for a class of problems from nonlinear elasticity. *Annales de l'Institut Henri Poincaré - Analyse non linéaire*, 8:119–157, 1991.

- [18] A. Braides, G. Dal Maso, and A. Garroni. Variational formulation for softening phenomena in fracture mechanics: the one-dimensional case. *Arch. Rat. Mech. Anal.*, 146:23–58, 1999.
- [19] C. Carstensen. Numerical analysis of nonconvex minimization problems allowing microstructures. *Zeitschrift für Angewandte Mathematik und Mechanik*, 76(S2):437–438, 1996.
- [20] C. Carstensen and P. Plecháč. Adaptive algorithms for scalar non-convex variational problems. *Appl. Numer. Math.*, 26:203–216, 1997.
- [21] C. Carstensen and P. Plecháč. Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.*, 6:997–1026, 1997.
- [22] C. Carstensen and P. Plecháč. Numerical analysis of compatible phase transitions in. *SIAM J. Numer. Anal.*, 37:2061–2081, 2000.
- [23] P.G. Ciarlet. *Mathematical Elasticity, Vol.I: Three-Dimensional Elasticity*. North-Holland, 1988.
- [24] P.G. Ciarlet and J. Nečas. Unilateral problems in nonlinear three-dimensional elasticity. *Arch. Rat. Mech. Anal.*, 87:319–338, 1985.
- [25] C. Collins, D. Kinderlehrer, and M. Luskin. Numerical approximation of the solution of a variational problem with a double well potential. *SIAM J. Numerical Analysis*, 28:321–332, 1991.
- [26] C. Collins and M. Luskin. Optimal-order error-estimates for the finite-element approximation of the solution of a nonconvex variational problem. *Mathematics of Computation*, 57:621–637, 1991.
- [27] B. Dacorogna. Quasiconvexity and relaxation of non convex variational problems. *J. Funct. Anal.*, 46:102–118, 1982.
- [28] B. Dacorogna. *Direct methods in the calculus of variations*. Springer, New York, 1989.
- [29] A.M. Davie. Singular minimizers in the calculus of variations in one dimension. *Arch. Rat. Mech. Anal.*, 101:161–177, 1988.
- [30] G. Dolzmann. Numerical computation of rank-one convex envelopes. *SIAM J. Numer. Anal.*, 36(5):1621–1635, 1999.
- [31] L.C. Evans. Private communication.
- [32] L.C. Evans. Quasiconvexity and partial regularity in the calculus of variations. *Arch. Rat. Mech. Anal.*, 95:227–268, 1986.
- [33] I. Fonseca and W. Gangbo. Local invertibility of Sobolev functions. *SIAM J. Math. Anal.*, 26:280–304, 1995.
- [34] M. Killough. *A diffuse interface approach to the development of microstructure in martensite*. PhD thesis, Courant Institute, New York University, 1998.
- [35] R.V. Kohn. The relaxation of a double-well energy. *Continuum Mechanics and Thermodynamics*, 3:193–236, 1991.
- [36] J. Kristensen. On the non-locality of quasiconvexity. *Ann. Inst. Henri Poincaré-Analyse Nonlinéaire*, 16:1–13, 1999.
- [37] M. Lavrentiev. Sur quelques problèmes du calcul des variations. *Ann. Mat. Pura Appl.*, 4:7–28, 1926.
- [38] A. Lazzeri and C.B. Bucknall. Applications of a dilatational yielding model to rubber-toughened polymers. *Polymer*, 36:2895–2902, 1995.
- [39] H. LeDret and A. Raoult. The quasiconvex envelope of the Saint Venant-Kirchhoff stored energy function. *Proc. Royal Soc. Edinburgh*, 125A:1179–1192, 1995.
- [40] B. Li and M. Luskin. Theory and computation for the microstructure

- near the interface between twinned layers and a pure variant of martensite. *Materials Science & Engineering A*, 273:237–240, 1999.
- [41] Z.-P. Li. Element removal method for singular minimizers in variational problems involving Lavrentiev phenomenon. *Proc. R. Soc. Lond. A*, 439:131–137, 1992.
- [42] Z.-P. Li. Element removal method for singular minimizers in problems of hyperelasticity. *Math. Models and Methods in Appl. Sci.*, 5:387–399, 1995.
- [43] Z.-P. Li. A numerical method for computing singular minimizers. *Numer. Math.*, 71:317–330, 1995.
- [44] M. Luskin. On the computation of crystalline microstructure. *Acta Numerica*, 5:191–258, 1996.
- [45] B. Manià. Sopra un esempio di Lavrentieff. *Bull. Un. Mat. Ital.*, 13:36–41, 1934.
- [46] E.E. Moise. *Geometric topology in dimensions 2 and 3*, volume 47 of *Graduate texts in Mathematics*. Springer-Verlag, New York, 1977.
- [47] C.B. Morrey. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.
- [48] S. Müller. Variational methods for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems*, volume 1713 of *Lecture Notes in Math.*, pages 85–210. Springer, Berlin, 1999.
- [49] S. Müller, T. Qi, and B.S. Yan. On a new class of elastic deformations not allowing for cavitation. *Ann. Inst. Henri Poincaré, Analyse Nonlinéaire*, 11:217–243, 1994.
- [50] P. Pedregal. *Parametrized measures and variational principles*, volume 30 of *Progress in nonlinear differential equations and their applications*. Birkhäuser, Basel, 1991.
- [51] P. Pedregal. *Variational methods in nonlinear elasticity*. SIAM, Philadelphia, 2000.
- [52] A.C. Pipkin. The relaxed energy density for isotropic elastic membranes. *IMA Journal of Applied Mathematics*, 36:85–99, 1986.
- [53] A.C. Pipkin. Elastic materials with two preferred states. *Quarterly J. Mech. Appl. Math.*, 44:1–15, 1991.
- [54] P. Plecháč. Computation of microstructure with interfacial energies. In *Proceedings ENUMATH 97, (Heidelberg)*, Singapore, 1998. World Scientific.
- [55] V. Šverák. Regularity properties of deformations with finite energy. *Arch. Rat. Mech. Anal.*, 100:105–127, 1988.
- [56] V. Šverák. Rank-one convexity does not imply quasiconvexity. *Proc. Royal Soc. Edinburgh*, 120A:185–189, 1992.
- [57] P.J. Swart and P.J. Holmes. Energy minimization and the formation of microstructure in dynamic anti-plane shear. *Arch. Rat. Mech. Anal.*, 121:37–85, 1992.
- [58] M.A. Sychev. Lebesgue measure of the universal singular set for the simplest problems in the calculus of variations. *Siberian Mathematical J.*, 35:1220–1233, 1994.
- [59] L. Tartar. The compensated compactness method applied to systems of conservation laws. In J.M. Ball, editor, *Systems of Nonlinear Partial Differential Equations*, pages 263–285. NATO ASI Series, Vol. C111, Reidel, 1982.

- [60] L. Tonelli. *Fondamenti di Calcolo delle Variazioni*, Volumes I, II. Zanichelli, 1921-23.
- [61] J.F. Traub and A.G. Werschulz. *Complexity and information*. Lezioni Lincei. Cambridge Univ. Press, 1998.
- [62] M Valadier. A course on Young measures. *Rend. Istit. Mat. Univ. Trieste*, 26:suppl., 349–394, 1994.
- [63] M. Šilhavý. *The mechanics and thermodynamics of continuous media*. Springer, 1997.