Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Variational models for microstructure and phase transitions

by

Stefan Müller

Lecture notes no.: 2 1998

Variational models for microstructure and phase transitions

Lectures at the C.I.M.E. summer school
'Calculus of variations and geometric evolution problems',
Cetraro, 1996

Stefan Müller
Max Planck Institute for
Mathematics in the Sciences
Inselstr. 22-26, 04103 Leipzig, Germany

Contents

1 Setting of the problem

1.1 What are microstructures?

For the purpose of these lectures, a microstructure is any structure on a scale between the macroscopic scale (on which we usually make observations) and the atomic scale. Such structures are abundant in nature: the fine hierarchical structures in a leaf and many other biological materials, the complex arrangements of fissures, cracks, voids and inclusions in rock or soil, fine scale mixing patterns in turbulent or multiphase flow, man-made layered or fibre-reinforced materials and fine phase mixtures in solid-solid phase transformations, to quote but a few examples. The microstructure influences in a crucial way the macroscopic behaviour of the material or system and is often chosen (or spontaneously generated) to optimize its performance (maximum strength at given weight, minimal energy, maximal entropy, maximal or minimal permeability, ...). Microstructures often develop on many different scales in space and time, and to understand the formation, interaction, and overall effect of these structures is a great scientific challenge, weather modelling providing an illustrative example. In the applied literature the passage from microscales to macroscales is frequently achieved by clever ad hoc "averaging" or "renormalization". A good mathematical framework in which these procedures could be justified and systematically improved is often lacking, and its development would be a difficult, but very rewarding, task.

The mathematical analysis of microstructures usually neglects the atomic scale by considering a continuum model from the outset. The issue is then to understand scales that are small (or converge to zero) compared to the fixed macroscopic scale. Research has mostly focused on three areas: homogenization, variational models of microstructure and optimal design which lies between the two first areas as the optimal structure often corresponds to a homogenization limit. The basic problem in homogenization is to determine the macroscopic behaviour (or at least bounds on it) induced by a given microstructure (given for example by a periodic mixture of two heat conductors in the limit of vanishing period, by a weakly convergent sequence of conductivity tensors or by statistical information). Variational models of microstructures try to model systems which spontaneously form internal microstructure by assuming that the structure formed has a certain optimality property. The reason for the formation of such microstructure is typically that no exact optimum exists and optimizing sequences have to develop finer and finer oscillations (which may only be limited by effects neglected in the model, such as the underlying atomic structure). An important task is to extract the relevant features of minimizing sequences. Young measures, which are discussed in Section 3 below, are one possibility to do this, but by no means the only one.

In these lectures I will focus on variational models for microstructures that arise from solid-solid phase transitions in certain elastic crystals (usually alloys, such as In-Th, Cu-Al-Ni, Ni-Ti). These materials display a fascinating variety of microstructures (see Fig. 1) which is closely linked to unusual and technologically interesting material behaviour (shape memory effect, pseudoelasticity). A mathematical model for elastic crystals will be introduced in Section 1.3 below. Before doing this let us briefly review the relation between microstructure and energy minimization in more detail in some simple examples.

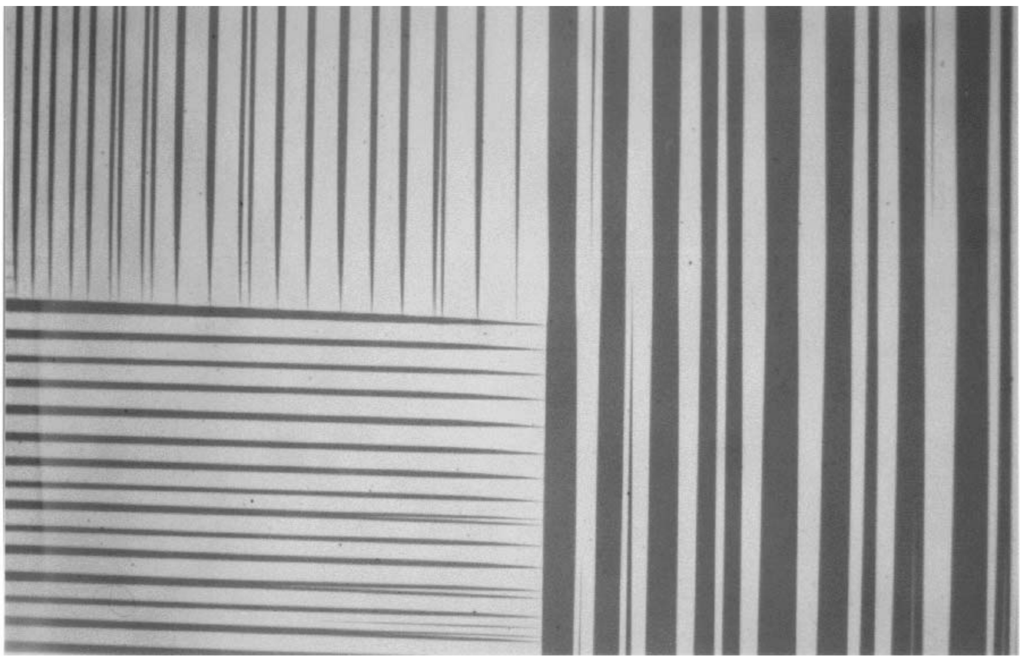

Figure 1: Microstructure in a Cu-Al-Ni single crystal; the imaged area is approximately 2 mm $\times$ 3 mm (courtesy of C. Chu and R.D. James, University of Minnesota)

1.2 Microstructures as energy minimizers

Example 1: Consider the problem:

$$\text{Minimize} \qquad \int_0^1 (u_x^2 - 1)^2 \, dx$$

subject to

$$u(0) = u(1) = 0.$$

The minimum is attained but the set of minimizers is highly degenerate. Every Lipschitz function whose slopes are ± 1 almost everywhere and that attains the boundary values is a minimizer. In particular the weak$*$ closure in $W^{1,\infty}$ of the set of minimizers consists of all functions with Lipschitz constant less then or equal to one that are bounded by $\pm \min(x, 1-x)$.

Example 2 (Bolza, L.C. Young): Consider the problem:

$$\text{Minimize} \qquad I(u) := \int_0^1 (u_x^2 - 1)^2 + u^2 \, dx$$

subject to

$$u(0) = u(1) = 0.$$

The infimum of the functional is zero since there exist rapidly oscillating functions with slope ± 1 whose supremum is arbitrarily small. Indeed if s denotes the periodic extension of the sawtooth function

$$s(x) = \begin{cases} x & \text{on } [0, 1/4) \\ 1/2 - x & \text{on } [1/4, 3/4) \\ x - 1 & \text{on } [3/4, 1) \end{cases} \tag{1.1}$$

then $u_j(x) := j^{-1} s(jx)$ satisfy $I(u_j) \to 0$ as $j \to \infty$. The infimum cannot be attained since there is no function that satisfies simultaneously $u \equiv 0$ and $u_x = \pm 1$ almost everywhere. Minimizing sequences must oscillate and converge weakly (in the Sobolev space $W^{1,4}(0,1)$), but not strongly, to zero.

This provides a first example how minimization can lead to fine scale oscillation or microstructure. The failure of classical minimization was investigated by L.C. Young in the 1930's in the context of optimal control. It led him to the introduction of generalized measure-valued solutions (see Section 3 below on Young measures). His book [Yo 69] describes various interesting situations where generalized solutions naturally arise, including applications to sailing and the construction of railway tracks.

Example 3: Let $\Omega = [0, L] \times [0, 1]$ be a rectangle and consider the problem:

$$\text{Minimize} \qquad J(u) = \int_\Omega u_x^2 + (u_y^2 - 1)^2 \, dx \, dy$$

subject to

$$u = 0 \quad \text{on} \quad \partial\Omega.$$

Clearly $J(u) > 0$ since otherwise $u_x = 0$ almost everywhere, whence $u \equiv 0$ on Ω and $(u_y^2 - 1)^2 = 1$. On the other hand the infimum of J is zero. One way to see this is to consider the sawtooth function s given by (1.1), to define

$$u(x, y) = j^{-1} s(jy) \ \text{ for } \delta < x < L - \delta,$$

and to use linear interpolation to achieve the boundary values at $x = 0$ and $x = L$. Considering first the limit $j \to \infty$ and then $\delta \to 0$ one obtains $\inf J = 0$. As in Example 2 no (classical) minimizers exist and minimizing sequences must develop rapid oscillations.

Two questions arise from the consideration of these examples.

Question 1: Are there special minimizers or minimizing sequences? Are, for example, the maximal solutions $\pm \min(x, 1-x)$ in Example 1 in a certain way preferred minimizers?

Question 2: Are there certain common features of all minimizing sequences?

1.3 Variational models for elastic crystals

The basic idea is to model the elastic crystal as a nonlinearly elastic continuum. The crystalline structure enters in this approach through the symmetry properties of the stored-energy function. The (usually stress-free) reference configuration of the crystal is identified with a bounded domain $\Omega \subset \mathbf{R}^3$. A deformation $u : \Omega \to \mathbf{R}^3$ of the crystal requires an elastic energy

$$I(u) = \int_\Omega W(Du)\, dx, \tag{1.2}$$

where $W : M^{m\times m} \to \mathbf{R}$ is the stored-energy density function that describes the properties of the material. Under the Cauchy-Born rule $W(F)$ is given by the (free) energy per unit volume that is required for an affine deformation $x \mapsto Fx$ of the crystal lattice.

The stored energy is invariant under rotations in the ambient space and under the action of the isotropy group $\mathcal{P}$ of the crystal lattice which usually is a discrete subgroup of $SO(3)$. Thus

$$W(QF) = W(F) \qquad \forall\, Q \in SO(3), \tag{1.3}$$

$$W(FP) = W(F) \qquad \forall\, P \in \mathcal{P} \subset SO(3). \tag{1.4}$$

Instead of the compact group $\mathcal{P}$ one could also consider the larger noncompact group of all lattice invariant transformation which is conjugate to the group $GL(3, \mathbf{Z})$. This leads to a highly degenerate situation and in particular such an invariance implies (in connection with the consideration of global rather than local mimimizers) that the material has no macroscopic shear resistance. We will thus use the point group and refer to [BJ 87, BJ 92], [CK 88], [Er 77, Er 79, Er 80, Er 84, Er 89], [Fo 87], [Pa 77, Pa 81], [Pi 84], [Za 92] for further discussion of this point.

The stored-energy also depends on temperature but we will always assume that the temperature is constant throughout the crystal and thus suppress this dependence.

The basic assumption of the variational approach to microstructure is:

The observed microstructures correspond to minimizers or almost minimizers of the elastic energy I.

It is convenient to normalize W so that $\min W = 0$. The set $K = W^{-1}(0)$ then corresponds to the zero energy affine deformations of the crystal lattices. Experimentally it is often observed that microstructures do not only minimize the integral I (subject to suitable boundary conditions) but in fact minimize the integrand pointwise. We are thus led to consider the simpler problem:

Determine (Lipschitz) maps that satisfy exactly
or approximately $Du \in K$.

The difference in behaviour of different materials is thus closely related to the set K which depends on the material and temperature. For ordinary materials K is (conjugate to) $SO(3)$ (the smallest set compatible with the rotation invariance) while for materials forming microstructures K consists of several copies of $SO(3)$. The Cu-Al-Ni alloy for which the microstructures in Fig.1 were observed undergoes a solid-solid phase transition at a critical temperature T_c, i.e. the preferred crystal structure, and hence the set K changes at T_c.

	$T > T_c$		$T < T_c$
crystal structure	cubic	phase transition $\longrightarrow$	orthorhombic
K	$SO(3)$		$SO(3)U_1 \cup \ldots SO(3)U_6$
			$U_1 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & \gamma \\ 0 & \gamma & \beta \end{pmatrix}$
micro-structure	none (see Section 2.3)		large variety observed (see Section 2.2)

Figure 2: A cubic to orthorhombic transition

In view of (1.4) the matrices U_i are related by conjugation under the cubic group.

1.4 The basic problems

We slightly generalize the setting of the previous section and consider maps $u: \Omega \subset \mathbf{R}^n \to \mathbf{R}^m$ on a bounded domain Ω (with Lipschitz boundary if needed). In particular the Sobolev space $W^{1,\infty}$ agrees with the class of Lipschitz maps. Let $K \subset M^{m\times n}$ be a compact set in the space $M^{m\times n}$ of $m \times n$ matrices.

Problem 1 (exact solutions): Characterize all Lipschitz maps u that satisfy

$$Du \in K \quad \text{a.e. in } \Omega.$$

Problem 2 (approximate solutions): Characterize all sequences u_j of Lipschitz functions with uniformly bounded Lipschitz constant such that

$$\text{dist }(Du_j, K) \to 0 \quad \text{a.e. in } \Omega.$$

Problem 3 (relaxation of K): Determine the sets K^{ex} and $K^{app} \subset M^{m\times n}$ of all affine maps $x \mapsto Fx$ such that Problem 1 and 2 have a solution that satisfies

$$\begin{aligned} u(x) &= Fx \text{ on } \partial\Omega, \\ u_j(x) &= Fx \text{ on } \partial\Omega, \end{aligned}$$

respectively.

Problems 1-3 also arise in many other contexts, e.g. in the theory of isometric immersions. An important technical difference is that in geometric problems one is often interested in connected sets K (and hence C^1 solutions u) while we will usually consider sets with more than one component. For further information we refer to Gromov's treatise [Gr 86] and to Šverák's ICM lecture [Sv 95].

In the context of crystal microstructure discussed in the previous section the sets K^{ex} and K^{app} in Problem 3 have an important interpretation. They consist of the affine *macroscopic* deformations of the crystal with (almost) zero energy. They trivially contain the set K of microscopic zero energy deformations but can be much larger. For the set $K = SO(2)A \cup SO(2)B$ one obtains (see Section 4.5) that under suitable conditions on A and B the sets K^{app} and K^{ex} contain an open set (relative to the constraint $\det F = 1$), leading to fluid-like behaviour.

Problem 4: Find an efficient description of approximating sequences that eliminates nonuniqueness due to trivial modifications while keeping the relevant "macroscopic" features.

We saw in Section 1.2 how failure of minimization can lead to "infinitely fine" microstructure. In practice crystal microstructures always arise on some finite scale (albeit on a wide range from a few atomic distances to $10 - 100\ \mu m$). Minimization of elastic energy alone may not be enough to explain this since there is no natural scale in the theory.

Problem 5: Explain the length scale and the fine geometry of the microstructure, possibly by including other contributions to the energy, such as interfacial energy.

Another possible explanation for limited fineness is that infinitely fine mixtures are (generalized) energy minimizers but not accessible by the natural dynamics of the system. This is a very important issue, but we can only touch briefly on it in these notes and refer to Section 7.2 and the references quoted there.

2 Examples

It is instructive to look at some examples before studying a general theory related to Problems 1-3. These simple examples already show a rich variety of phenomena and interesting connections with (nonlinear) elliptic regularity, functional analytic properties of minors and quasiconformal geometry. In the following K always denotes a subset of the space $M^{m\times n}$ of $m \times n$ matrices, $m, n \geq 2$, and Ω is a domain, i.e. an open and connected set, in $\mathbf{R}^n$.

2.1 The two-gradient problem

Exact solutions: Let $K = \{A, B\}$. The simplest solutions of the relation

$$Du \in K$$

are so called simple laminates, i.e. maps for which Du is constant in alternating bands that are bounded by hyperplanes $x \cdot n = \text{const}$ (see Fig. 3). Tangential continuity of u at these interfaces enforces that $A\tau = B\tau$ for vec-

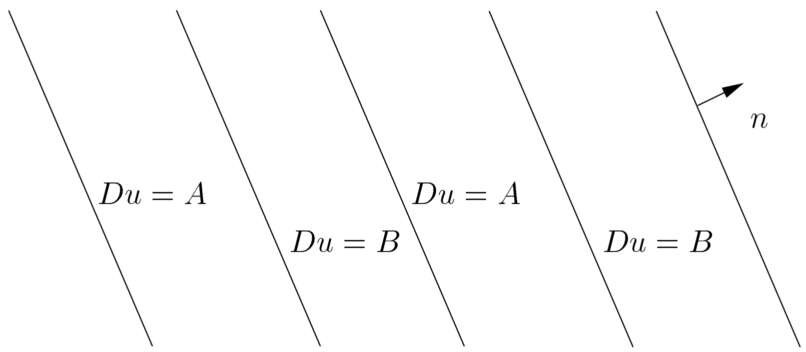

Figure 3: A simple laminate

tors τ perpendicular to n and thus $A - B$ has rank one and can be written as

$$B - A = a \otimes n.$$

In this case we say that A and B are rank-1 connected. We recall that the matrix $a \otimes n$ has entries $(a \otimes n)_{ij} = a_i n_j$. If one assumes that the interfaces between the regions $\{Du = A\}$ and $\{Du = B\}$ are smooth then a similar argument shows that they must be hyperplanes with fixed normal n. Moreover no such smooth arrangement is possible if $\mathrm{rk}(B - A) \geq 2$. The following proposition gives a much stronger statement because it shows that also among possibly very irregular maps there are no other solutions.

Proposition 2.1 *([BJ 87], Prop.1) Let Ω be a domain in $\mathbf{R}^n$ and let $u : \Omega \to \mathbf{R}^m$ be a Lipschitz map with $Du \in \{A, B\}$ a.e.*

(i) *If $\mathrm{rk}(B - A) \geq 2$, then $Du = A$ a.e. or $Du = B$ a.e.;*

(ii) *if $B - A = a \otimes n$ then u can locally be written in the form*

$$u(x) = Ax + ah(x \cdot n) + \text{const}$$

where h is Lipschitz and $h' \in \{0, 1\}$ a.e. If Ω is convex this representation holds globally.

In particular, Du is constant if u satisfies an affine boundary condition $u(x) = Fx$ on $\partial\Omega$.

Proof. The key idea is that the curl of a gradient vanishes. By translation we may assume $A = 0$ and thus $Du = B\chi_E$, for some measurable set $E \subset \Omega$. For part (i) we may assume in addition, after an affine change of the dependent and independent variables, that the first two rows of the matrix B are given by the standard basis vectors e_1 and e_2 and thus

$$Du^1 = e_1\chi_E, \; Du^2 = e_2\chi_E.$$

Symmetry of the second distributional derivatives and the first equation imply that $\partial_j\chi_E = 0$ for $j \neq 1$ while the second equation yields $\partial_k\chi_E = 0$ for $k \neq 2$. Hence $D\chi_E = 0$ in the sense of distributions and therefore $\chi_E = 1$ a.e. or $\chi_E = 0$ a.e. since Ω is connected.

To prove part (ii) we may assume $A = 0$, $a = n = e_1$ and thus $Du^1 = e_1\chi_B$, $Du^k = 0$, $k = 2, \ldots, m$. Hence $u^2, \ldots, u^m$ are constant and $\partial_k u^1 = 0$, for $k = 2, \ldots, m$. Therefore u^1 is locally only a function of x^1 as claimed. If Ω is convex then u^1 is constant on the hyperplanes $x^1 = \text{const}$ that intersect Ω and thus globally of the desired form.

Finally if $u = Fx$ on $\partial\Omega$, then $F = (1 - \lambda)B$, $\lambda \in [0, 1]$ since by the Gauss-Green theorem

$$|E| \, B = \int_\Omega Du \, dx = \int_{\partial\Omega} u \otimes n \, d\mathcal{H}^{n-1} = \int_\Omega F \, dx,$$

where n is the outer normal of Ω. Extending u by Fx on $\mathbf{R}^n \setminus \Omega$ we can argue as in the proof of (ii) to deduce $u(x) = Ax + a\,\tilde{h}(x \cdot n) + b$ on $\mathbf{R}^n$, where $\tilde{h}' \in \{0, 1-\lambda, 1\}$. Hence $u(x) \equiv Fx$ since each plane $x \cdot n = \text{const}$ intersects the set where $u = Fx$. $\square$

Approximate solutions: Consider again $K = \{A, B\}$ and suppose

$$B - A = a \otimes n, \quad F = \lambda A + (1-\lambda)B, \; \lambda \in [0,1].$$

We show that there exist sequences u_j with uniformly bounded Lipschitz constant such that in Ω

$$\operatorname{dist}(Du_j, \{A, B\}) \to 0 \quad \text{in measure,} \tag{2.1}$$

and

$$u_j(x) = Fx \quad \partial\Omega. \tag{2.2}$$

Note that (2.1) and the bound on the Lipschitz constant imply that convergence also holds in L^p, $\forall p < \infty$. After translation we may assume

$$F = 0, A = -(1-\lambda)a \otimes n, \quad B = \lambda\, a \otimes n.$$

Let h be the periodic extension of the function given by

$$h(t) = \begin{cases} -(1-\lambda)t & t \in [0, \lambda), \\ \lambda(t-1) & t \in [\lambda, 1], \end{cases}$$

and consider

$$v_j(x) = \frac{1}{j}\, a\, h(jx \cdot n).$$

Then $Dv_j \in \{A, B\}$ a.e. and $v_j \to 0$. To achieve the boundary conditions consider a cut-off function $\varphi \in C^\infty([0,\infty))$, $0 \le \varphi \le 1, \varphi = 0$ on $[0, 1/2]$, $\varphi = 1$ on $[1, \infty)$ and let

$$u_j(x) = \varphi(j \operatorname{dist}(x, \partial\Omega)) v_j(x).$$

Then $u_j = 0$ on $\partial\Omega$, Du_j is uniformly bounded and $Du_j = Dv_j$ except in a strip of thickness $1/j$ around $\partial\Omega$. If follows that u_j satisfies (2.1) and (2.2). Various modifications of this construction are possible, and we return in Section 3.2 to the question whether all approximating sequences are in a certain sense equivalent.

Note that due to the assumption $B - A = a \otimes n$, the problem (2.1), (2.2) essentially reduces to the scalar problem discussed in Example 3 of Section 1.2.

We now consider the case $\mathrm{rk}(B-A) \geq 2$. We have shown that in this case there are no nontrivial exact solutions. The argument used strongly the fact that Du only takes two values and that the curl of a gradient vanishes. It does not apply to approximating sequences. Nonetheless we have

Lemma 2.2 *([BJ 87], Prop.2) Suppose that* $\mathrm{rk}(B-A) \geq 2$ *and that* u_j *is a sequence with uniformly bounded Lipschitz constant such that*

$$\mathrm{dist}(Du_j, \{A, B\}) \to 0 \quad \textit{in measure in } \; \Omega.$$

Then

$$Du_j \to A \textit{ in measure} \quad \textit{or} \quad Du_j \to B \textit{ in measure.}$$

In particular the problem (2.1), (2.2) has only the trivial solution, $F \in \{A, B\}$ *and* $Du_j \to F$ *in measure.*

The proof uses the following fundamental properties of minors. We recall that the semiarrow $\rightharpoonup$ denotes weak convergence.

Theorem 2.3 *[Ba 77, Mo 66, Re 67] Let* M *be an* $r \times r$ *minor (subdeterminant).*

(i) *If* $p \geq r$ *and* $u, v \in W^{1,p}(\Omega), u - v \in W_0^{1,p}(\Omega)$ *then*

$$\int_\Omega M(Du) = \int_\Omega M(Dv). \tag{2.3}$$

In particular

$$\int_\Omega M(Du) = \int_\Omega M(F) \quad \textit{if } u = Fx \textit{ on } \partial\Omega.$$

(ii) *If* $p > r$ *and if the sequence* u_j *satisfies*

$$u_j \rightharpoonup u \text{ in } W^{1,p}(\Omega; \mathbf{R}^m),$$

then

$$M(Du_j) \rightharpoonup M(Du) \text{ in } L^{p/r}(\Omega).$$

Remark. Integrands f for which the integral $\int f(Du)$ only depends on the boundary values of u are called null Lagrangians, since the Euler-Lagrange equations are automatically satisfied for all functions u. Affine combinations of minors are the only null Lagrangians and the only functions that have the weak continuity property expressed in (ii) (see also Section 4.3).

Proof of Theorem 2.3. The main point is that minors can be written as divergences. For $n = m = 2$ one has

$$\det Du = \partial_1(u^1 \partial_2 u^2) - \partial_2(u^1 \partial_1 u^2), \tag{2.4}$$

for all $u \in C^2$ and hence for all $u \in W^{1,2}$ if the identity is understood in the sense of distributions. More generally, for $n = m \geq 2$ the cofactor matrix that consists of the $(n-1) \times (n-1)$ minors of Du satisfies

$$\text{div cof } Du = 0, \quad \text{i.e. } \partial_j(\text{cof} Du)_{ij} = 0 \tag{2.5}$$

and thus

$$\det\ Du = \frac{1}{n} \partial_j (u^i (\text{cof} Du)_{ij}),$$

since $F(\text{cof } F)^T = \text{Id} \det F$. Similar formulae hold for general $r \times r$ minors, see [Mo 66, Da 89, GMS 96] for the detailed calculations. The multilinear algebra involved in these calculations can be expressed very concisely through the use of differential forms. In this setting one has for $n = m = 2$

$$\det\ Du\, dx^1 \wedge dx^2 = du^1 \wedge du^2 = d\,(u^1 \wedge du^2),$$

while for the $r \times r$ minor $M(Du)$ that involves the rows $1, \ldots, r$ and the columns $1, \ldots, r$ one has

$$\begin{aligned} M(Du) dx^1 \wedge \ldots dx^n &= du^1 \wedge \ldots \wedge du^r \wedge dx^{r+1} \wedge \ldots \wedge dx^n \\ &= d\,(u^1 \wedge du^2 \wedge \ldots du^r \wedge\ dx^{r+1} \wedge \ldots \wedge dx^n). \end{aligned}$$

In either formulation (i) follows from the Gauss-Green (or Stokes) theorem (and approximation by smooth functions) while (ii) follows from induction over the order r of minors and the fact that u_j converges strongly in L^p. $\square$

Proof of Lemma 2.2. We may assume $A = 0$ and that there exists a 2×2 minor M such that $M(B) = 1$. By assumption there thus exist sets E_j such that

$$Du_j - B\chi_{E_j} \to 0 \text{ in measure }, \tag{2.6}$$

and hence in L^p for all $p < \infty$. Moreover there exists a subsequence (not relabelled) such that

$$\chi_{E_j} \overset{*}{\rightharpoonup} \theta \text{ in } L^\infty(\Omega), \quad u_j \overset{*}{\rightharpoonup} u \text{ in } W^{1,\infty}(\Omega; \mathbf{R}^m). \tag{2.7}$$

It follows from Theorem 2.3 and (2.6)

$$B\chi_{E_j} \overset{*}{\rightharpoonup} Du = B\theta,$$

$$M(B)\chi_{E_j} \overset{*}{\rightharpoonup} M(Du) = M(B)\theta^2. \tag{2.8}$$

Combining the first convergence in (2.7) and (2.8) we see that $\theta = \theta^2$ a.e. Thus θ must be a characteristic function χ_E. Hence (2.7) implies that (use e.g. the fact $\|\chi_{E_j}\|_{L^2} \to \|\chi_E\|_{L^2}$)

$$\chi_{E_j} \to \theta = \chi_E \quad \text{in measure.}$$

Therefore by (2.6)

$$Du_j \to Du = B\chi_E \quad \text{in measure.}$$

Finally Lemma 2.1 (i) implies that $Du = B$ a.e. or $Du = A = 0$ a.e. $\square$

2.2 Applications to crystal microstructures

Before proceeding with the mathematical discussion of the problem $Du \in K$ let us briefly review what can be learned about crystal microstructure from the considerations so far. Which microstructures can form and why are they so fine?

First let us consider again the rôle of rank-1 connections. In the continuum theory discussed in the previous section they were related to continuity of the tangential derivatives or to the fact that the curl of a gradient vanishes (in Section 2.6 we still study the connections with the Fourier transform). The condition can also be understood in the discrete setting of crystal lattices. Two homogeneous lattices, obtained by affine deformations A and B of the same reference lattice can meet at a common plane S only if the deformations differ by a shear that leaves S invariant. Analytically we recover the condition $B - A = a \otimes n$, where n is the normal of S (see Figure 4).

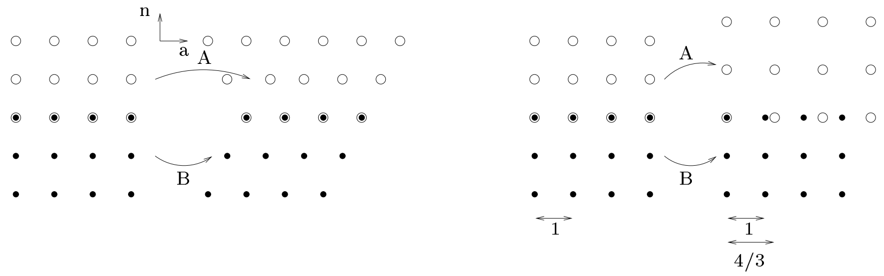

Figure 4: Compatible and incompatible lattice deformations. On the left the condition $B - A = a \otimes n$ is satisfied, on the right $B = \mathrm{Id}$, $A = 4/3\,\mathrm{Id}$, so the condition is violated. After deformation there is no interface on which the two lattices meet.

Under certain additional conditions the two sublattices are referred to as twins. There are different definitions what precisely constitutes a twin; a common requirement is that $B = QAH$, where $Q \in SO(3) \setminus \{\mathrm{Id}\}$, $Q^2 = \mathrm{Id}$ and where H belongs to the point group of the crystal, see [Ja 81] and [Za 92] for further discussion. Compatible lattice deformations can be arranged in alternating bands of different deformations, see Figure 5 (cf. also Fig. 3).

If the set $K \in M^{m\times n}$ of minimizing affine deformations contains more rank-1 connections then more complicated patterns such as the double laminates (or 'twin crossings') in Figure 6 are possible.

In this way one can explain the observation of a number of microstructures through an analysis of rank-1 connections. The constructions based on rank-1 connections, however, involve no length scale. Why, then, are the observed structures often so fine?

For the situation of just two deformations A and B Proposition 2.1 (ii) and the discussion of approximate solutions provide an explanation. As soon as one imposes a nontrivial affine boundary condition $F = \lambda A + (1 - \lambda)B$ there are no exact solutions, and approximate solutions become the better

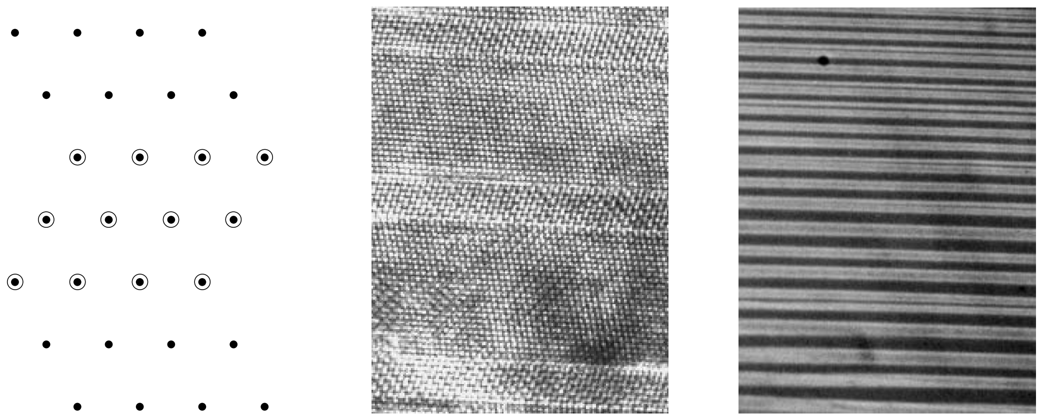

Figure 5: Compatible lattice deformations can be arranged in laminar patterns. Schematic drawing (left), atomic resolution micrograph of fine twinning in Ni-Al (middle; courtesy of D. Schryvers, RUCA, Antwerp), twinning in Cu-Al-Ni (right; courtesy of C. Chu and R. D. James), grey and black represent two different lattice deformations.

the finer A and B are mixed (in a real crystal, additional contribution to the energy may eventually limit the fineness, see Section 6). In practice boundary conditions are often not so much imposed globally but by contact with other parts of the crystal where other deformation gradients prevail (e.g. because the phase transformation has not yet taken place there).

A typical example is the frequently observed austenite/finely-twinned martensite interface (see Figure 7). In an idealized situation this corresponds to a homogeneous affine deformation C on one side of the interface and a fine mixture of A and B on the other side. Neither A nor B are rank-1 connected to C but a suitable convex combination $\lambda A + (1-\lambda)B$ is.

There is no deformation that uses all three gradients A, B and C and only these (see the end of the proof of Proposition 2.1). However, the volume fraction of gradients other than A, B and C can be made arbitrarily small by matching C to a fine mixture of layers of A and B in volume fractions λ and $1-\lambda$.

The analysis of the rank-1 connections determines the volume fraction λ as well as the interface normals n and m, in very good agreement with experiment; see [BJ 87], Theorem 3 and [JK 89], Section 5 for a detailed discussion.

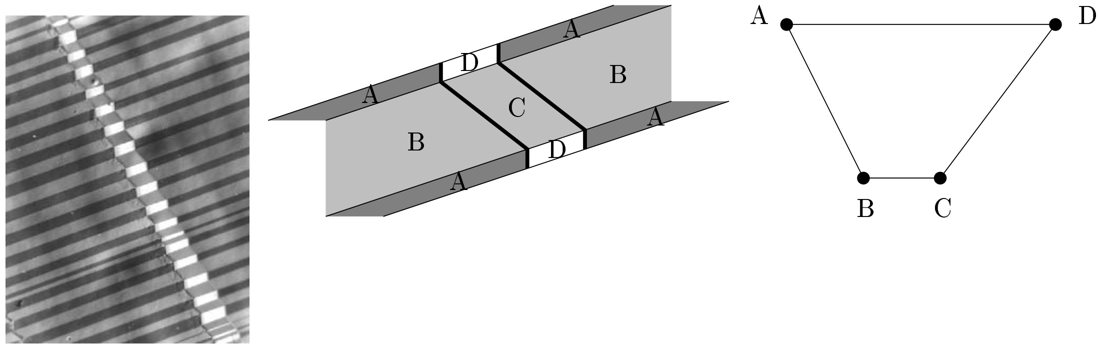

Figure 6: Twin crossings in Cu-Al-Ni (courtesy of C. Chu and R. D. James) and schematic drawings of the different deformation gradients and their rank-1 connections (indicated by solid lines).

More complex patterns like the wedge microstructure in Figure 8 can be understood in a similar vein. In this particular case so many rank-1 connections are required that the microstructure can only arise if the transformation strain satisfies a special relation; see [Bh 91], [Bh 92] for a comparison of theory and experiment.

The considerations in this subsection focused on constructions of microstructures based on rank-1 connections. Do these constructions cover (in a suitable sense) all possible microstructures? We return to this fundamental question in the remainder of this Section and in particular in Sections 4.3, 4.6 and 4.7.

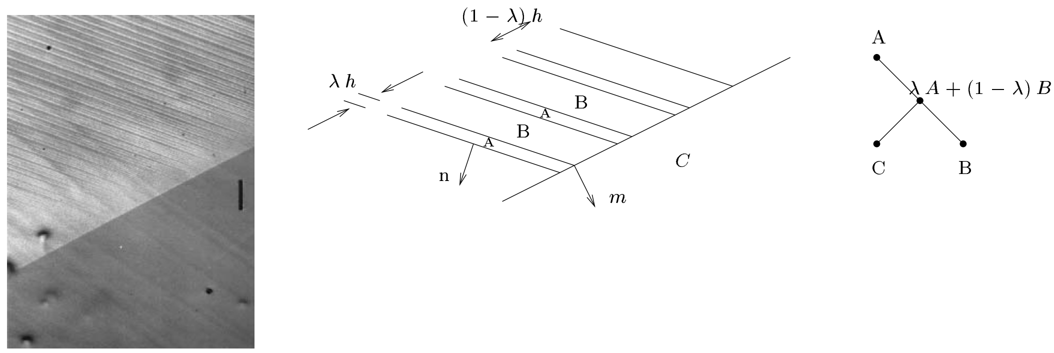

Figure 7: Austenite/finely twinned martensite interface in Cu-Al-Ni (courtesy of C. Chu and R. D. James), schematic distribution of deformation gradients and rank-1 connections; a simple model for the refinement (branching) of the A/B twins towards the interface with C is discussed in section 6.2.

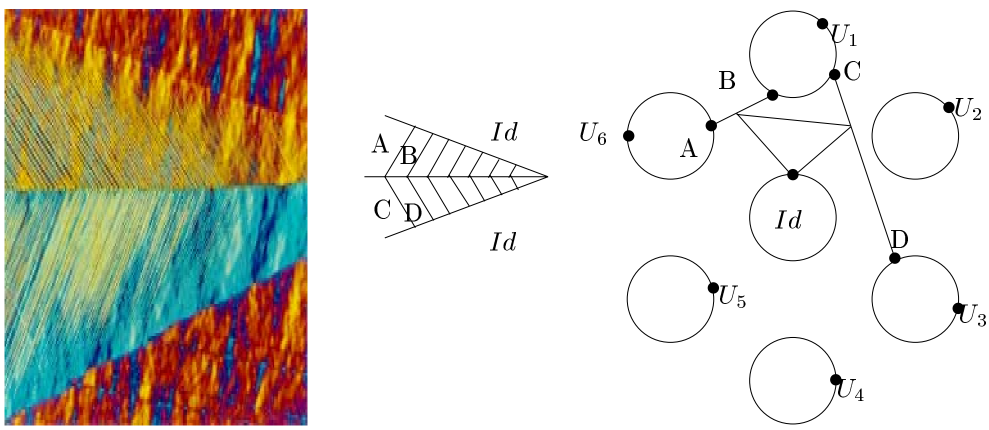

Figure 8: Wedge microstructure in Cu-Al-Ni (courtesy of C. Chu and R. D. James). The necessary rank-1 connections between the six orthorhombic wells $SO(3)U_i$ and the untransformed phase only exist for special transformation strains U_1.

2.3 The one-well problem

The simplest set K that is compatible with the symmetry requirements (1.3) and (1.4) is $K = SO(n)$. In this case approximating sequences must converge strongly.

Theorem 2.4 *([Ki 88], p.231) Suppose that*

$$Du \in SO(n) \text{ a.e. in } \Omega.$$

Then Du is constant and $u(x) = Qx + b$, $Q \in SO(n)$. If u_j is a sequence of functions with uniformly bounded Lipschitz constant such that

$$\operatorname{dist}(Du_j, SO(n)) \to 0 \quad \textit{in measure,} \tag{2.9}$$

then

$$Du_j \to \text{const} \quad \textit{in measure.}$$

Proof. To prove the first statement recall from (2.5) that

$$\operatorname{div} \ \operatorname{cof} Du = 0$$

for any Lipschitz map. Now $\operatorname{cof} F = F$ for all $F \in SO(n)$ and thus u is harmonic and therefore smooth. Moreover $|Du|^2 = n$, where $|F|^2 = \operatorname{tr} F^T F = \sum_{i,j} F_{ij}^2$, and therefore

$$2|D^2 u|^2 = \Delta |Du|^2 - 2\, Du \cdot D\Delta u = 0.$$

Thus Du is constant.

To prove the second assertion of the theorem we may assume that $u_j \overset{*}{\rightharpoonup} u$ in $W^{1,\infty}(\Omega; \mathbf{R}^m)$. Consider the function

$$f(F) = |F|^n - c_n \det F, \quad c_n = n^{n/2}.$$

One easily checks that $f \geq 0$ and that f vanishes exactly on matrices of the form λQ, $\lambda \geq 0$, $Q \in SO(n)$ (use polar decomposition, diagonalize and apply the arithmetic-geometric mean inequality). Hence (2.9), the weak continuity of minors (Theorem 2.3) and the weak lower semicontinuity of the L^n norm imply that

$$\begin{aligned} 0 &= \liminf_{n\to\infty} \int_\Omega f(Du_j)\,dx \\ &= \liminf_{n\to\infty} (\int_\Omega |Du_j|^n\,dx - c_n \int_\Omega \det Du_j\,dx) \\ &\geq \int_\Omega |Du|^n\,dx - c_n \int_\Omega \det Du\,dx = \int_\Omega f(Du) \geq 0. \end{aligned}$$

Therefore all the inequalities must be equalities and in particular

$$f(Du) = 0 \text{ a.e.}, \quad \|Du_j\|_{L^n} \to \|Du\|_{L^n}.$$

It follows that

$$\begin{aligned} Du_j &\to Du \quad \text{in} \quad L^n(\Omega; M^{m\times n}) \text{ (hence in measure)}, \\ Du(x) &= \lambda(x)Q(x), \quad \lambda \geq 0, \quad Q(x) \in SO(n) \text{ a.e.} \end{aligned}$$

Moreover $|Du_j|^2 = n$ a.e., whence $|Du|^2 = n$ a.e. Thus $Du \in SO(n)$ a.e. and, by the first part of the theorem $Du = \text{const.}$ □

The case $n = 2$ of the above result shows some interesting connections with the Cauchy-Riemann equations. Identify $\mathbf{C} \simeq \mathbf{R}^2$ as usual via $z = x+iy$ and let $\partial_z = 1/2(\partial_x - i\partial_y)$, $\partial_{\bar{z}} = 1/2(\partial_x + i\partial_y)$. Suppose that $1 < p < \infty$ and

$$\text{dist}(Du_j, SO(2)) \to 0 \quad \text{in } L^p(\Omega). \tag{2.10}$$

Then in particular $|\partial_z u_j| \to 1$ and

$$\partial_{\bar{z}} u_j \to 0 \quad \text{in } L^p(\Omega; \mathbf{C}),$$

and regularity for the Cauchy-Riemann operator implies that there exists a function u s.t.

$$u_j \to u \text{ in } W^{1,p}(\Omega; \mathbf{C}), \quad \partial_{\bar{z}} u = 0.$$

Thus u is (weakly) holomorphic and $|\partial_z u| = \lim_{j\to\infty} |\partial_z u_j| = 1$. Hence $\partial_z u = \text{const.}$

2.4 The three-gradient problem

Theorem 2.5 *([Sv 91b]). Let* $K = \{A_1, A_2, A_3\}$ *and suppose that* $\mathrm{rk}(A_i - A_j) \neq 1$.

(i) *If* $Du \in K$ a.e. *then* Du *is constant (a.e.).*

(ii) *If* u_j *is a sequence with uniformly bounded Lipschitz constant such that*

$$\mathrm{dist}(Du_j, K) \to 0 \quad \textit{in measure}$$

then

$$Du_j \to \mathrm{const} \quad \textit{in measure.}$$

Proof of part (i). For simplicity we only consider the case $n = m = 2$. The general case can be reduced to this if one considers separately the cases that the span E of $A_2 - A_1$ and $A_3 - A_1$ contains two, one or no rank-1 lines and uses Lemma 2.7 below, see also [Sv 91b].

We may assume that $A_1 = 0$ and thus $\det A_2 \neq 0, \det A_3 \neq 0$. Multiplying by A_2^{-1} we may further assume $A_2 = \mathrm{Id}$. Using the Jordan normal form we see that after a change of variables we have either

$$A_3 = \begin{pmatrix} \lambda & -\mu \\ \mu & \lambda \end{pmatrix}, \quad \lambda^2 + \mu^2 \neq 0$$

or

$$A_3 = \begin{pmatrix} \lambda & a \\ 0 & \mu \end{pmatrix}, \quad \lambda \neq 0, \mu \notin \{0, 1\}.$$

In the first case u satisfies the Cauchy-Riemann equations and is holomorphic and therefore smooth. Thus $Du \equiv A_i$ since K is discrete. In the second case $Du \in K$ implies that

$$\partial_1 u^2 = 0.$$

Hence $u^2(x) = h(x^2)$ (locally) and $\partial_2 u^2(x) = h'(x^2)$. Since $\mu \notin \{0, 1\}$ the value of $\partial_2 u^2$ uniquely determines one of the matrices A_i. Thus $Du(x) = g(x^2)$. In particular

$$\partial_1 \partial_1 u = 0, \quad \partial_2 \partial_1 u = \partial_1 \partial_2 u = 0$$

in the sense of distributions. Thus $\partial_1 u = \text{const}$ and $Du = \text{const} \otimes e_1 + \tilde{g}(x^2) \otimes e_2$. Therefore $\text{rk}(Du(x) - Du(\tilde{x})) \leq 1$ and thus $Du \equiv A_i$. □

An alternative proof that features an interesting connection with the theory of quasiconformal (or more precisely quasiregular) maps proceeds as follows. After possible renumbering we may assume that $\det(A_2 - A_1)$ and $\det(A_3 - A_1)$ have the same sign. Taking $A_1 = 0$ and multiplying by $\text{diag}(1, -1)$ if needed we have $\det A_2 > 0, \det A_3 > 0$. Thus $Du \in K$ implies that

$$|Du|^2 \geq k \, \det Du$$

for a suitable constant k. Hence u is quasiregular and a deep result of Reshetnyak says that either $u = \text{const}$ or u is a local homeomorphism up to a discrete set B_u of branch points and that the (local) inverse u^{-1} preserves sets of measure zero (see [Ah 66], [Bo 57], [Re 89]). Hence either $Du = 0$ a.e. or $Du \neq 0$ a.e. In view of the results for the two-gradient problem this implies the assertion.

The proof of (ii) requires more subtle arguments (see [Sv 91b], [Sv 92b]). Šverák first shows that after suitable transformations (and elimination of some simpler special cases) one may assume

$$A_i = A_i^T, \quad \det A_i = 1.$$

Now a gradient Du is symmetric if and only if u is itself a gradient Dv. Thus assertion (ii) is essentially reduced to a study of approximate solutions of the Monge-Ampère equation

$$\det D^2 v_j \to 1, \quad v_j : \Omega \in \mathbf{R}^2 \to \mathbf{R}.$$

The difficulty is that, different from the usual literature on the Monge-Ampère equation, one cannot assume that $D^2 v_j$ is positive (semi-)definite. Indeed a crucial step in the proof that uses ideas from the theory of quasiregular maps is to show that $\det D^2 v > 0$ a.e. implies that v is locally convex or concave.

2.5 The four-gradient problem

The following example which was found independently by a number of authors (I am aware of [AH 86], [CT 93] and [Ta 93]; see [BFJK 94] for the adaptation of Tartar's construction for separately convex functions to diagonal matrices) shows that the absence of rank-1 connections does not guarantee absence of microstructures (i.e. strong convergence of approximating sequences).

Lemma 2.6 *Consider the* 2×2 *diagonal matrices* $A_1 = -A_3 = \mathrm{diag}(-1,-3)$, $A_2 = -A_4 = \mathrm{diag}(-3,1)$ *and let* $K = \{A_1, A_2, A_3, A_4\}$. *Then* $\mathrm{rk}(A_i - A_j) \neq 1$ *but there exists a sequence* u_j

$$\mathrm{dist}(Du_j, K) \to 0 \quad \textit{in measure,}$$

and Du_j *does not converge in measure.*

Exercise: Show that there is no nontrivial solution of $Du \in K$ for the above choice of K. Hint: consult the previous subsection.

It is not known whether there is another choice of four matrices with $\mathrm{rk}(A_i - A_j) \neq 1$ for which nontrivial solutions exist. It is known, but not trivial, that for each $\epsilon > 0$ there exist nontrivial maps such that $\mathrm{dist}(Du, K) < \epsilon$ (see the discussion after Theorem 5.4). Note that for small ϵ the set of admissible gradients still contains no rank-1 connections.

Proof. Since K contains no rank-1 connections the key idea is to 'borrow' four additional matrices J_i (see Fig. 9) and to successively remove the regions where Du assumes J_i. We will construct a sequence v_k that satisfies the affine boundary condition

$$v_k(x) = J_4 x \quad \text{on } \partial Q = \partial(0,1)^2.$$

As a first approximation we may take $v^{(0)}(x) = J_4 x$. To increase the measure of the set where the gradients lie in K we observe that J_4 is a rank-1 convex combination of A_1 and J_1,

$$J_4 = \frac{1}{2}A_1 + \frac{1}{2}J_1.$$

As in Section 2.1 we can thus construct a map $v^{(1)}$ that agrees with $v^{(0)}$on ∂Q and uses only gradients A_1 and J_1 (in layers of thickness 1/2k) except for a boundary layer of thickness c/k where the gradient remains uniformly bounded. In the next step we replace the stripes where $Dv^{(1)} = J_1$ by fine

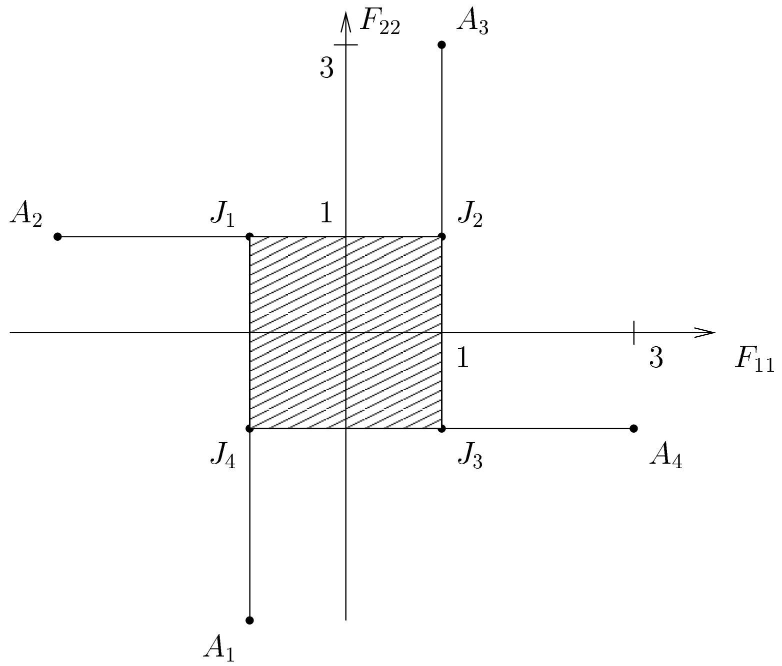

Figure 9: Four incompatible matrices that support a nontrivial minimizing sequence

layers of A_2 and J_2 and k new boundary layers of thickness c/k^2. This yields $v^{(2)}$ (see Fig. 10). The volume fraction of the J_i phases has been decreased to $\left(\frac{1}{2}\right)^2$ (up to small corrections due to the boundary layers). If we replace J_2 by fine layers of A_3 and J_3 (with k^2 boundary layers of thickness c/k^3) we obtain $v^{(3)}$ and replacing J_3 by A_4 and J_4 we obtain $v^{(4)}$. Up to the boundary layers $Dv^{(4)}$ only uses the values A_i and J_4. Compared to $v^{(0)}$ the volume fraction of the set where J_4 is taken has been reduced from one to (slightly less than) $(1/2)^4$. The volume fraction of the boundary layers is bounded by

$$\frac{c}{k} + k\frac{c}{k^2} + k^2\frac{c}{k^3} + k^3\frac{c}{k^4} = 4\frac{c}{k}.$$

Hence we have

$$|\{Dv^{(4)} \notin K\}| \leq \frac{4c}{k} + \frac{1}{16}.$$

To further reduce the volume fraction of the set $Dv \notin K$ we can now apply the same procedure to each of the small rectangles where $Dv^{(4)} = J_4$.

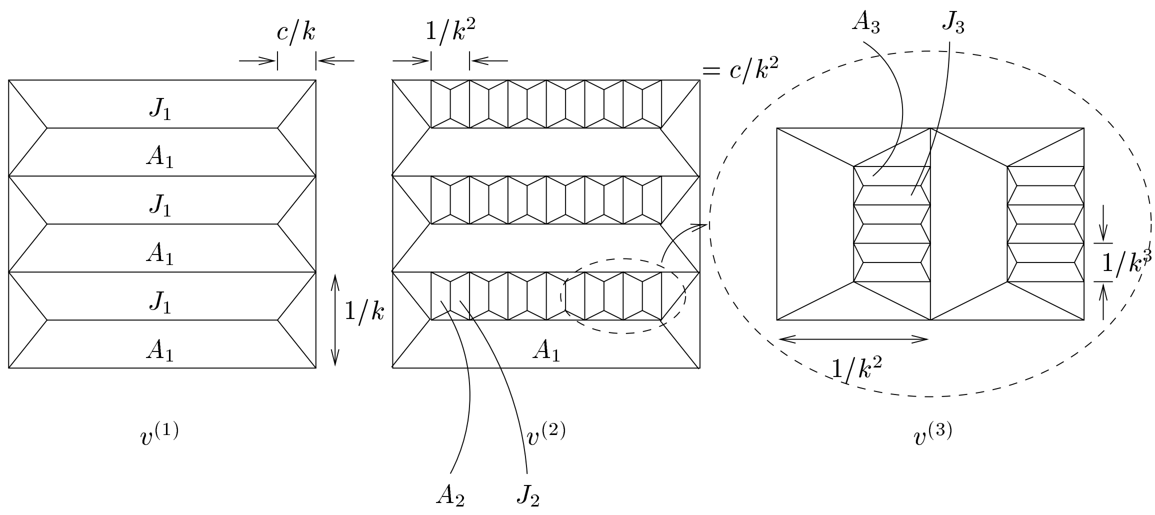

Figure 10: The first three stages in the construction of v_k.

After l iterations we obtain

$$\begin{aligned} |\{Dv^{(4l)} \notin K\}| &\leq \sum_{m=0}^{l-1} \left(\frac{1}{16}\right)^m \frac{4c}{k} + \left(\frac{1}{16}\right)^l \\ &\leq \frac{C}{k} + \left(\frac{1}{16}\right)^l . \end{aligned}$$

With a suitable choice of l we thus find maps $v_k : Q \to \mathbf{R}^2$ such that $|Dv_k| \leq L$ and

$$\begin{aligned} |\{Dv_k \notin K\}| &\leq \frac{C}{k} \to 0, \\ v_k(x) &= J_4 x \text{ on } \partial Q. \end{aligned}$$

In particular $\operatorname{dist}(Dv_k, K) \to 0$ in measure.

We finally show that no subsequence of Dv_k can converge in measure. Indeed if $Dv_{k_j} \to Dv$ in measure then $Dv \in K$ in Q and $v = J_4 x$ on $\partial\Omega$. This is impossible since $Dv \in K$ implies $Dv \equiv \text{const}$ (see Exercise). Alternatively one can easily verify that

$$Dv_k \overset{*}{\rightharpoonup} J_4 \quad \text{in } L^\infty(Q; \mathbf{R}^2).$$

This contradicts convergence in measure since $J_4 \notin K$. □

Exercise. Show that for all F in the shaded region in Figure 4 there exists a sequence such that

$$\begin{array}{rcl} Dv_k & \overset{*}{\rightharpoonup} & F \quad L^\infty(Q; \mathbf{R}^m), \\ \operatorname{dist}(Dv_k, K) & \to & 0 \quad \text{in measure.} \end{array}$$

In fact the shaded region together with the rank-1 lines between the A_i and the J_i contains all such F. One possible proof uses the nontrivial fact that the function

$$f(F) = \begin{cases} \det F & F \quad \text{symmetric } \geq 0 \\ 0 & F \quad \text{symmetric } \ngeq 0 \\ +\infty & F \quad \text{not symmetric} \end{cases}$$

is quasiconvex (see [Sv 92b]). See Sections 4.3 and 4.4 for further information about the classification of weak limits.

2.6 Linear subspaces and elliptic systems

Lemma 2.7 *Let L be a linear subspace of $M^{m\times n}$ that contains no rank-1 line.*

(i) *If u is Lipschitz and $Du \in L$* a.e. *then u is smooth*

(ii) *If u_j is a sequence that satisfies*

$$\begin{array}{rcl} u_j & \overset{*}{\rightharpoonup} & u \quad \textit{in } W^{1,\infty}(\Omega; \mathbf{R}^m), \\ \operatorname{dist}(Du_j, L) & \to & 0 \quad \textit{in measure,} \end{array}$$

then $Du \in L$ a.e. *and*

$$Du_j \to Du \quad \textit{in measure.}$$

Remark. In (i) it suffices to assume that $u \in W^{1,1}$, in (ii) it suffices that $u_j \to u$ in L^1_{loc} and that Du_j is bounded in L^1_{loc}.

Proof. Let $A : M^{m\times n} \to M^{m\times n}$ denote the projection onto the orthogonal complement of L. Then $Du \in L$ is equivalent to

$$A\ Du = 0. \tag{2.11}$$

The assumption that L contains no rank-1 lines essentially assures that (2.11) is a linear elliptic system, and the assertions follow easily from the general theory of such systems. We sketch the proof for the convenience of the reader.

Suppose that v has compact support in Ω, f belongs to the Sobolev space $W^{k,2}(\Omega; \mathbf{R}^m)$ (i.e. all distributional derivatives up to order k belong to L^2) and v satisfies

$$A\ Dv = f. \tag{2.12}$$

We claim that $v \in W^{k+1,2}(\Omega; \mathbf{R}^m)$ and

$$\|D^{k+1}v\|_{L^2} \le C\|D^k f\|_{L^2}. \tag{2.13}$$

To prove this consider the Fourier transform

$$iA\ \hat{v}(\xi) \otimes \xi = \hat{f}(\xi)$$

of (2.11). Since L contains no rank-1 connections we have $A(a \otimes \xi) \neq 0$ if $a \neq 0$, $\xi \neq 0$, and by homogeneity

$$|A(a \otimes \xi)| \ge c|a|\,|\xi|$$

for some constant $c > 0$. The claim follows now from Plancherel's Theorem.

To prove (i) let $\varphi \in C_0^\infty(\Omega)$. Then

$$A\ D(\varphi u) \quad = \quad A(u \otimes D\varphi).$$

In view of (2.13) we have the implication $u \in W^{k,2}_{\text{loc}} \Rightarrow u \in W^{k+1,2}_{\text{loc}}$, and this yields (i).

To prove (ii) observe that the hypothesis and the linearity of (2.11) imply that $u_j \to u$ in L^2_{loc}, $ADu_j \to 0$ in L^2_{loc}, $ADu = 0$.

Application of (2.13) with $v = \varphi(u_j - u)$ yields the assertion. □

To establish (i) for $u \in W^{1,1}$ it suffices to mollify u and to pass to the limit. To prove (ii) under the hypothesis in the remark one can use the weak L^1 estimates for elliptic systems (or, more precisely, for good Fourier multipliers).

Examples.

1. $L = \left\{ F \in M^{2\times 2} : F = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \right\}$; this corresponds to the Cauchy-Riemann equations

$$\partial_1 u^1 - \partial_2 u^2 = \partial_2 u^1 + \partial_1 u^2 = 0.$$

2. $L = \left\{ F \in M^{n\times n} : F^T = F, \mathrm{tr} F = 0 \right\}$; this corresponds to the Laplace equation $\Delta v = 0$, since Du symmetric implies $u = Dv$ (locally).

3. $L = \{ F \in M^{n\times n} : \mathrm{tr} F = 0, F_{ij}\xi_k - F_{ik}\xi_j = 0 \ , \forall \xi \in \mathbf{R}^n \setminus \{0\} \}$; this corresponds to the system div $u = 0$, curl $u = 0$.

Problem. What is the largest dimension $d(m, n)$ of a subspace of $M^{m\times n}$ that contains no rank-1 line?

This is closely related to questions in algebraic geometry and K-theory, e.g. to the number of linearly independent vector fields on S^{n-1}. For $m = n$ Example 2 provides the lower bound $d(n, n) \geq \frac{n(n+1)}{2}$. The upper bound $d(n, n) = n^2 - n$ is sharp exactly in dimension $n = 2, 4$ and 8. See [BFJK 94] for further information.

3 Efficient description of minimizing sequences - Young measures

3.1 The fundamental theorem on Young measures

We have seen in the examples in Sections 1.2 and 2.1 that there are usually many minimizing sequences for a variational problem. We return now to the question whether all these sequences have some common features and whether one can describe the 'macroscopic' features of a sequence without paying attention to unnecessary details. Closely related is the issue of defining a notion of generalized solution for variational problems that do not admit classical solutions.

A reasonable condition for an object that describes the macroscopic behaviour of a sequence $z_j : E \to \mathbf{R}^d$ is that it should determine the limits of

$$\int_U f(z_j)$$

for continuous functions f (such as energy-, stress- or entropy density) and for all measurable subsets U of E. Such an object exists and was first introduced by L.C. Young in connection with generalized solutions of optimal control problems. By $C_0(\mathbf{R}^d)$ we denote the closure of continuous functions on $\mathbf{R}^d$ with compact support. The dual of $C_0(\mathbf{R}^d)$ can be identified with the space $\mathcal{M}(\mathbf{R}^d)$ of signed Radon measures with finite mass via the pairing

$$\langle \mu, f \rangle = \int_{\mathbf{R}^d} f d\mu.$$

A map $\mu : E \to \mathcal{M}(\mathbf{R}^d)$ is called weak$*$ measurable if the functions $x \mapsto \langle \mu(x), f \rangle$ are measurable for all $f \in C_0(\mathbf{R}^d)$. We often write μ_x instead of $\mu(x)$.

Theorem 3.1 *(Fundamental theorem on Young measures)*

Let $E \subset \mathbf{R}^n$ be a measurable set of finite measure and let $z_j : E \to \mathbf{R}^d$ be a sequence of measurable functions. Then there exists a subsequence z_{j_k} and a weak$*$ measurable map $\nu : E \to \mathcal{M}(\mathbf{R}^d)$ such that the following holds.

(i) $\nu_x \geq 0, \quad \|\nu_x\|_{\mathcal{M}(\mathbf{R}^d)} = \int_{\mathbf{R}^d} d\nu_x \leq 1, \quad$ *for* a.e. $x \in E$.

(ii) *For all* $f \in C_0(\mathbf{R}^d)$

$$f(z_{j_k}) \overset{*}{\rightharpoonup} \bar{f} \quad \textit{in} \;\; L^\infty(E),$$

where

$$\bar{f}(x) = \langle \nu_x, f \rangle = \int\limits_{\mathbf{R}^d} f d\nu_x.$$

(iii) *Let* $K \subset \mathbf{R}^d$ *be compact. Then*

$$\mathrm{supp}\nu_x \subset K \quad \textit{if} \;\; \mathrm{dist}(z_{j_k}, K) \to 0 \;\textit{in measure.}$$

(iv) *Furthermore one has*

$$(i') \qquad \|\nu_x\|_{\mathcal{M}} = 1 \quad \textit{for}\text{ a.e. } x \in E$$

if and only if the sequence does not escape to infinity, i.e. if

$$\lim_{M \to \infty} \sup_k |\{|z_{j_k}| \geq M\}| = 0. \tag{3.1}$$

(v) *If (i') holds, if* $A \subset E$ *is measurable, if* $f \in C(\mathbf{R}^d)$ *and if*

$$f(z_{j_k}) \textit{ is relatively weakly compact in } L^1(A),$$

then

$$f(z_{j_k}) \rightharpoonup \bar{f} \text{ in } L^1(A), \quad \bar{f}(x) = \langle \nu_x, f \rangle.$$

(vi) *If (i') holds, then in (iii) one can replace 'if' by 'if and only if'.*

Remarks. 1. The map $\nu : E \to \mathcal{M}(\mathbf{R}^d)$ is called the Young measure generated by (or: associated to) the sequence z_{j_k}. Every (weakly$*$ measurable) map $\nu : E \to \mathcal{M}(\mathbf{R}^d)$ that satisfies (i) is generated by some sequence z_k.

2. The assumption $|E| < \infty$ was only introduced for notational convenience, cf. [Ba 89]. In fact $\mathbf{R}^d$ with Lebesgue measure can be replaced by a more general measure space $(\mathcal{S}, \Sigma, \mu)$, e.g. a locally compact space with a Radon measure. The converse statement in Remark 1 requires that μ be non-atomic.

3. The target $\mathbf{R}^d$ can be replaced e.g. by a compact metric space K. In this case one always has $\|\nu_x\|= 1$ a.e. The condition (3.1) has a simple interpretation if we replace $\mathbf{R}^d$ by its one-point compactification $K = \mathbf{R}^d \cup \{\infty\} \simeq S^d$ and consider the corresponding family of measures $\tilde{\nu}_x$ on K. Then $\|\tilde{\nu}_x\|=$ 1 a.e., and (3.1) ensures that $\tilde{\nu}_x$ does not charge the point ∞.

4. If, for some $s > 0$ (!) and all $j \in \mathbf{N}$

$$\int_E |z_j|^s \leq C$$

then (3.1) holds.

5. Here is a typical application of (v): if $\{z_j\}$ is bounded in L^p and $|f(s)| \leq C(1+|s|^q), q < p$, then $f(z_{j_k}) \rightharpoonup \bar{f}$ in $L^{p/q}$. In particular, for $p > 1$ the choice $f = \mathrm{id}$ yields

$$z_{j_k} \rightharpoonup z, \quad z(x) = \langle \nu_x, \mathrm{id} \rangle. \tag{3.2}$$

Proof. The point is to pass from the functions z_j which take values in $\mathbf{R}^d$ to maps which take values in the space of $\mathcal{M}(\mathbf{R}^d)$ of measures in $\mathbf{R}^d$. Thus we allow new limiting objects which do not take a precise function value at every point but a probability distribution of values.

Let

$$Z_j(x) = \delta_{z_j(x)}.$$

Then $\|Z_j(x)\|_{\mathcal{M}(\mathbf{R}^d)}= 1$ and $\langle Z_j(x), f\rangle = f(z_j(x))$. Thus Z^j belongs to the space $L^\infty_w(E; \mathcal{M}(\mathbf{R}^d))$ of weak$*$ measurable maps $\mu : E \to \mathcal{M}(\mathbf{R}^d)$ that are (essentially) bounded. Now it turns out $L^\infty_w(E; \mathcal{M}(\mathbf{R}^d))$ is the dual of the separable space $L^1(E; C_0(\mathbf{R}^d))$ (see e.g. [Ed 65, p.588], [IT 69, p.93], [Me 66, p.244]), where the duality pairing is given by

$$\langle \mu, g\rangle = \int_E \langle \mu(x), g(x)\rangle dx.$$

Hence the Banach-Alaoglu theorem yields a subsequence such that

$$Z_{j_k} = \delta_{z_{j_k}(\cdot)} \overset{*}{\rightharpoonup} \nu \quad \text{in } L^\infty_w(E; \mathcal{M}(\mathbf{R}^d)). \tag{3.3}$$

Lower semicontinuity of the norm implies that $\| \nu_x \| \leq 1$ for a.e. x. For $\varphi \in L^1(E)$ and $f \in C_0(\mathbf{R}^d)$ we denote by $\varphi \otimes f$ the element of $L^1(E; C_0(\mathbf{R}^d))$

given by $x \mapsto \varphi(x)f$. The definition of Z_j and (3.3) thus imply

$$\int_E \varphi(x) f(z_{j_k}(x))dx = \langle Z_{j_k}, \varphi \otimes f \rangle \to \int_E \varphi(x) \langle \nu_x, f \rangle dx.$$

Hence (ii) follows, and considering all functions $f \geq 0, \varphi \geq 0$ we also deduce $\nu_x \geq 0$.

To prove (iii) it suffices to show that

$$\langle \nu_x, f \rangle = 0 \qquad \forall\, f \in C_0(\mathbf{R}^d \setminus K). \tag{3.4}$$

Let $f \in C_0(\mathbf{R}^d \setminus K)$. Then for every $\epsilon > 0$ there exist C_ϵ such that $|f(y)| \leq \epsilon + C_\epsilon \operatorname{dist}(y, K)$. Hence the hypothesis $\operatorname{dist}(z_{j_k}, K) \to 0$ in measure implies that $(|f| - \epsilon)^+(z_{j_k}) \to 0$ in measure, and in view of (ii) we conclude that

$$\langle \nu_x, (|f| - \epsilon)^+ \rangle = 0 \qquad \text{for a.e. } x.$$

Now (3.4) follows since $\epsilon > 0$ was arbitrary.

The proof of (iv) and (v) is easily achieved by a careful truncation argument and the characterization of weakly compact sets in L^1 [Me 66], see [Ba 89] for the details. Finally the proof of (vi) follows by an application of (v) to the bounded function $f = \min(\operatorname{dist}(\cdot, K), 1)$. □

Remark. Since the span of tensor products $\varphi \otimes f, \varphi \in L^1(\Omega), f \in C_0(\mathbf{R}^d)$, is dense in $L^1(\Omega; C_0(\mathbf{R}^d))$ assertion (ii) of the theorem is equivalent to $Z_{jk} \stackrel{*}{\rightharpoonup} \nu$.

The measure ν_{x_0} describes the probability of finding a certain value in the sequence $z_{j_k}(x)$ for x in a small neighbourhood $B_r(x_0)$ in the limits $j \to \infty$ and $r \to 0$. The following useful fact reflects this probabilistic interpretation.

Corollary 3.2 *Suppose that a sequence z_j of measurable functions from E to $\mathbf{R}^d$ generates the Young measure $\nu : E \to \mathcal{M}(\mathbf{R}^d)$. Then*

$$z_j \to z \text{ in measure} \quad \text{if and only if} \quad \nu_x = \delta_{z(x)} \text{ a.e.}$$

Proof. If $z_j \to z$ in measure then $f(z_j) \to f(z)$ in measure for all $f \in C_0(\mathbf{R}^d)$. Hence by Theorem 3.1 (ii) one has $\langle \nu_x, f \rangle = f(z(x))$ for all $f \in C_0(\mathbf{R}^d)$ and thus $\nu_x = \delta_{z(x)}$. If conversely $\nu_x = \delta_{z(x)}$ a.e. we claim that

$$\limsup_{j \to \infty} |\{|z_j - w| > \epsilon\}| \leq |\{|z - w| > \epsilon/2\}|,$$

for all piecewise constant measurable functions $w : E \to \mathbf{R}^d$. To see this it suffices to consider constant functions $w \equiv a$ and to apply (v) with $f(y) = \varphi(|y-a|)$ where φ is continuous $0 \le \varphi \le 1, \varphi = 1$ on $[\epsilon, \infty), \varphi = 0$ on $[0, \epsilon/2]$. Thus

$$\begin{aligned}\limsup_{j\to\infty} |\{|z_j - z| > \epsilon\}| &\le \limsup_{j\to\infty} |\{|z_j - w| > \epsilon/2\}| + |\{|w - z| > \epsilon/2\}| \\ &\le 2|\{|z - w| > \epsilon/4\}|.\end{aligned}$$

The last term can be made arbitrarily small since measurable functions can be approximated by piecewise constant functions, and the assertion follows (note that z is measurable since $\{\nu_x\}_{x\in E}$ is weak$*$ measurable). □

An alternative approach to the 'if' part of the corollary is to apply Corollary 3.3 below to the Carathéodory function $f(x, y) = \min(|y - z(x)|, 1)$.

3.2 Examples

a) Let $h : \mathbf{R} \to \mathbf{R}$ be the periodic extension of the function given by

$$h(x) = \begin{cases} a & \text{if } 0 \le x < \lambda, \\ b & \text{if } \lambda \le x < 1, \end{cases}$$

and define $z_j : [0, 1] \to \mathbf{R}$ by

$$z_j(x) = h(jx). \tag{3.5}$$

Using the periodicity of h one easily checks that (see e.g. [Da 81], p.8),

$$z_j \overset{*}{\rightharpoonup} \int_0^1 h(y)dy = \lambda a + (1 - \lambda)b$$

and similarly

$$f(z_j) \overset{*}{\rightharpoonup} \lambda f(a) + (1 - \lambda)f(b).$$

Hence z_j generates a Young measure ν given by

$$\nu_x = \lambda\delta_a + (1 - \lambda)\delta_b.$$

In particular ν_x is independent of x. Such Young measures are called *homogeneous Young measures.*

More generally (see e.g. [BM 84]) if $h : \mathbf{R}^n \to \mathbf{R}$ is locally integrable and periodic with unit cell $[0,1]^n$ and z_j is defined by (3.5), then z_j generates a homogeneous Young measure ν given by

$$\int_{\mathbf{R}} g d\nu = \int_{[0,1]^n} g(h(y))dy.$$

For a Borel set $B \subset \mathbf{R}$ one has

$$\nu(B) = |(0,1)^n \cap h^{-1}(B)|.$$

b) Let

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

let u_j be a sequence such that

$$I(u_j) \to 0, \qquad u_j(0) = u_j(1) = 0, \tag{3.6}$$

and let $z_j = (u_j)_x$ (cf. Example 2 in Section 1.2). Then z_j is bounded in L^4, a subsequence generates a Young measure ν and $\|\nu_x\| = 1$ a.e. If we let $g(p) = \min((p^2 - 1)^2, 1)$ we deduce from (3.6) that

$$\langle \nu_x, g \rangle = 0 \qquad \text{for a.e. } x.$$

Hence $\operatorname{supp} \nu_x \subset \{-1, 1\}$ and $\nu_x = \lambda(x)\delta_{-1} + (1 - \lambda(x))\delta_1$ a.e. By Remark 5 after Theorem 3.1

$$z_{j_k} \overset{*}{\rightharpoonup} \langle \nu_x, \mathrm{id} \rangle = 1 - 2\lambda(x) \tag{3.7}$$

and

$$u_{j_k}(a) = \int_0^a z_{j_k} dx \to \int_0^a (1 - 2\lambda(x))dx. \tag{3.8}$$

By (3.6) $u_j \to 0$ in L^2 and thus $\lambda(x) = 1/2$ a.e. Hence z_{j_k} generates the unique (homogeneous) Young measure

$$\nu_x = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1.$$

By uniqueness the whole sequence z_j generates this Young measure.

Although there are many different minimizing sequences for I they all generate the same Young measure. The Young measure captures the essential feature of minimizing sequences. They have to use slopes (close to) ± 1 in equal proportion in a finer and finer mixture.

One may view the pair (u, ν) as a generalized solution of the problem $I \to \min$. The derivative u_x is replaced by a probability measure and the coupling between u and ν occurs through the centre of mass of ν (cf. (3.7), (3.8) and Theorem 4.9):

$$u_x = \langle \nu_x, \mathrm{id} \rangle.$$

c) (Approximate solutions of the two-well problem)

Let $A, B \in M^{m \times n}, B - A = a \otimes n, F = \lambda A + (1-\lambda)B, \lambda \in (0,1)$. Consider a sequence of maps $u_j : \Omega \subset \mathbf{R}^n \to \mathbf{R}^m$ with uniformly bounded Lipschitz constant that satisfies

$$\mathrm{dist}(Du_j, \{A, B\}) \to 0 \quad \text{in measure in } \Omega,$$

$$u_j(x) = Fx \quad \text{in } \partial\Omega.$$

Let ν be the Young measure generated by (a subsequence of) Du_j. Then $\|\nu_x\| = 1$ and Theorem 3.1 (iii) yields $\mathrm{supp}\nu_x \subset \{A, B\}$, i.e. $\nu_x = \mu(x)\delta_A + (1 - \mu(x))\delta_B$. Passing to a further subsequence we may assume $u_j \stackrel{*}{\rightharpoonup} u$ in $W^{1,\infty}(\Omega, \mathbf{R}^m)$, and in view of (3.2) we have

$$Du(x) = \mu(x)A + (1 - \mu(x))B = A + (1 - \mu(x))a \otimes n.$$

Extending u_j and u by Fx outside Ω we deduce that $v(x) = u(x) - Ax$ is constant on the planes $x \cdot n = \text{const}$. Hence $u(x) = Fx$ and $\mu(x) = \lambda$. Thus $\{Du_j\}$ generates the unique (homogeneous) Young measure

$$\nu_x = \lambda\delta_A + (1 - \lambda)\delta_B.$$

d) (Four-gradient example)

The sequence Du^j constructed in Section 2.5 generates the unique homogeneous Young measure

$$\nu_x = \frac{8}{15}\delta_{A_1} + \frac{4}{15}\delta_{A_2} + \frac{2}{15}\delta_{A_3} + \frac{1}{15}\delta_{A_4}.$$

Proof. Exercise.

3.3 What the Young measure cannot detect

The Young measure describes the local phase proportions in an infinitesimally fine mixture (modelled mathematically by a sequence that develops finer and finer oscillations). This is exactly what is needed to compute limits of integrals $\int_U f(z_j)$. There are, however, other natural quantities that cannot be computed from the Young measure.

Example 1 (micromagnetism).
The energy of a large rigid magnetic body represented by a domain $\Omega \subset \mathbf{R}^3$, is given by

$$I(m) = \int_{\Omega} \varphi(m) + \int_{\mathbf{R}^3} |h_m|^2.$$

Here $m : \Omega \to \mathbf{R}^3$ is the magnetization and h_m is the Helmholtz projection of $-m$ (extended by zero outside Ω), i.e. the unique gradient field that satisfies $\mathrm{div} h_m = -\mathrm{div} m$ in the sense of distributions. In suitable units m satisfies the saturation condition $|m| = 1$. For simplicity we have neglected exchange energy (this is a good approximation for large bodies, see [DS 93]).

Let $m_j : \Omega \to S^2 \subset \mathbf{R}^3$ be a sequence of magnetizations that generates a Young measure ν. Then

$$\int_{\Omega} \varphi(m_j) dx \to \int_{\Omega} \langle \nu_x, \varphi \rangle dx.$$

The limit of $\int_{\mathbf{R}^3} |h_{m_j}|^2$, however, is in general not determined by the Young measure (see Fig. 11). Indeed let f be the periodic extension of the sign function on $[-1/2, 1/2]$, let $\Omega = [0, 1]^3$ and let

$$m_j = f(jx^1) e^1 \chi_\Omega \, ; \quad \tilde{m}_j = f(jx^2) e^1 \chi_\Omega.$$

Both sequences generate the same (homogeneous) Young measure $\nu_x = \frac{1}{2}\delta_{e^1} + \frac{1}{2}\delta_{-e^1}$. On the other hand it is not difficult to verify that $\|h_{m_j}\|_2 \to 1$ while $\| h_{\tilde{m}^j} \|_2 \to 0$. First replace χ_Ω by a smooth function φ and show that the resulting fields M^j and $\tilde{M}^j$ satisfy curl $M_j \to 0$, $\mathrm{div}\tilde{M}_j \to 0$ in H^{-1}; then use the estimate $\|h_{m_j} - h_{M_j}\|_2 \leq \|m_j - M_j\|_2$ which holds since the map $m \mapsto -h_m$ is an orthogonal projection. Alternatively one may use the representation of h_m in Fourier space.

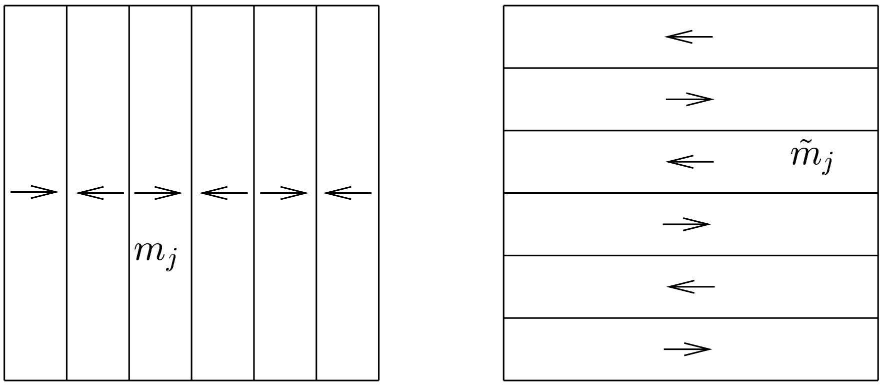

Figure 11: Both sequences generate the same Young measure but m_j is almost a gradient while $\tilde{m}_j$ is almost divergence free.

Example 2 (correlations).

The limit of

$$I_j(u_j) := \int_0^1 u_j(x) u_j(x + \frac{1}{j}) dx$$

is not determined by the Young measure of $\{u_j\}$. Indeed consider

$$\begin{aligned} u_j(x) &= \sin j\pi x, \\ v_j(x) &= \sin 2j^2\pi x. \end{aligned}$$

Both sequences generate the same (homogeneous) Young measure $\nu_x = \frac{2}{\pi}(\sin^{-1})'(y)dy$ (cf. Section 3.2 a)), but

$$\begin{aligned} I_j(u_j) &= \int_0^1 \sin(j\pi x)\ \sin(j\pi x + \pi) \to -1/2, \\ I_j(v_j) &= \int_0^1 \sin(j^2\pi x)\ \sin(j^2\pi x + 2j\pi) \to 1/2. \end{aligned}$$

3.4 More about Young measures and lower semicontinuity

We have seen that Young measures are useful as a concept since they give a precise meaning to the idea of 'infinitesimally fine phase mixture' and provide a framework for generalized solutions where no classical minimizers exist.

In this section, which may be omitted on first reading, we briefly discuss the advantages of Young measures as a technical tool. The following two results allow one, among other things, to extend lower semicontinuity results for integrals $\int f(Du(x))dx$ to integrals $\int f(x, u(x), Du(x))dx$ without additional effort. More generally, Young measures are a rather efficient tool to eliminate all dependence on 'lower order' terms by soft general arguments. The first result shows that the Young measure suffices to compute limits of Carathéodory functions, the second extends the characterization of strong convergence in Corollary 3.2.

Corollary 3.3 *Suppose that the sequence of maps $z_k : E \to \mathbf{R}^d$ generates the Young measure ν. Let $f : E \times \mathbf{R}^d \to \mathbf{R}$ be a Carathéodory function (measurable in the first argument and continuous in the second) and assume that the negative part $f^-(x, z_k(x))$ is weakly relatively compact in $L^1(E)$. Then*

$$\liminf_{k\to\infty} \int_E f(x, z_k(x))dx \geq \int_E \int_{\mathbf{R}^d} f(x, \lambda)d\nu_x(\lambda)dx. \tag{3.9}$$

If, in addition, the sequence of functions $x \mapsto |f|(x, z_k(x))$ is weakly relatively compact in $L^1(E)$ then

$$f(\cdot, z_k(\cdot)) \rightharpoonup \bar{f} \text{ in } L^1(E), \quad \bar{f}(x) = \int_{\mathbf{R}^d} f(x, \lambda)d\nu_x(\lambda)dx. \tag{3.10}$$

Remarks. 1. Assertion (3.9) still holds if f is (Borel) measurable on $E \times \mathbf{R}^d$ and lower semicontinuous in the second argument rather than a Carathéodory function (see [BL 73]).

2. The choice $f(x, p) = \min(|p - z(x)|, 1)$ in (3.10) can be used to prove the 'if' statement in Corollary 3.2.

Proof. It suffices to prove (3.9). The second assertion follows by application of this inequality to $\tilde{f}(x, p) = \pm\varphi(x)f(x, p)$ for all $\varphi \in L^\infty(E), \varphi \geq 0$.

To prove (3.9) first consider the case $f \geq 0$. Assume temporarily that, in addition,

$$f(x, \lambda) = 0 \quad \text{if } |\lambda| \geq R. \tag{3.11}$$

By the Scorza-Dragoni theorem there exists an increasing sequence of compact sets E_j such that $|E \setminus E_j| \to 0$ and $f_{|E_j \times \mathbf{R}^d}$ is continuous. Define

$F_j : E \to C_0(\mathbf{R}^d)$ by $F_j(x) = \chi_{E_j}(x)f(x,\cdot)$. Then $F_j \in L^1(E; C_0(\mathbf{R}^d))$ and the convergence of $\delta_{z_k(\cdot)}$ to ν in the dual space yields

$$\begin{array}{rll} & \int\limits_E f(x, z_k(x))dx & \geq \int\limits_E \langle \delta_{z_k(x)}, F_j(x) \rangle \\ \to & \int\limits_E \langle \nu_x, F_j(x) \rangle dx & = \int\limits_{E_j} f(x,\lambda) d\nu_x(\lambda). \end{array}$$

Letting $j \to \infty$ we obtain the assertion by the monotone convergence theorem. To remove the assumption (3.11) consider an increasing sequence $\{\eta_l\} \subset C_0^\infty(\mathbf{R}^d)$, that converges to 1, use the estimate for $f_l(x,\lambda) = f(x,\lambda)\eta_l(\lambda)$ and apply again the monotone convergence theorem. This finishes the proof if $f \geq 0$ or more generally if f is bounded from below.

For general f let

$$\begin{array}{ll} h_k(x) & = f(x, z_k(x)) = h_k^+(x) - h_k^-(x), \\ f_M(x,\lambda) & = \max(f(x,\lambda), -M). \end{array}$$

By the equivalent characterizations of equiintegrability (see e.g. [Me 66]) for each $\epsilon > 0$ there exists an $M > 0$ such that

$$\sup_k \int\limits_{h_k^- \geq M} h_k^-(x) dx < \epsilon.$$

Hence

$$\begin{aligned} \liminf_{k\to\infty} \int\limits_E f(x, z_k(x))dx + \epsilon &\geq \liminf_{k\to\infty} \int\limits_E f_M(x, z_k(x))dx \\ \geq \int\limits_E \int\limits_{\mathbf{R}^d} f_M(x,\lambda) d\nu_x(\lambda) dx &\geq \int\limits_E f(x,\lambda) d\nu_x(\lambda) dx. \end{aligned}$$

Since $\epsilon > 0$ was arbitrary the proof is finished. □

Corollary 3.4 *Let $u_j : E \to \mathbf{R}^d, v_j : E \to \mathbf{R}^{d'}$ be measurable and suppose that $u_j \to u$* a.e. *while v_j generates the Young measure ν. Then the sequence of pairs $(u_j, v_j) : E \to \mathbf{R}^{d+d'}$ generates the Young measure $x \mapsto \delta_{u(x)} \otimes \nu_x$.*

Proof. Let $\varphi \in C_0(\mathbf{R}^d), \psi \in C_0(\mathbf{R}^{d'}), \eta \in L^1(E)$. Then $\varphi(u_j) \to \varphi(u)$a.e. and $\eta\varphi(u_j) \to \eta\varphi(u)$ in $L^1(E)$ by the dominated convergence theorem. Moreover by assumption

$$\psi(v_j) \stackrel{*}{\rightharpoonup} \bar{\psi} \quad \text{in } L^\infty, \qquad \bar{\psi}(x) = \langle \nu_x, \psi \rangle.$$

Hence

$$\int_E \eta(\varphi \otimes \psi)(u_j, v_j)dx = \int_E \eta\varphi(u_j)\psi(v_j)dx \to \int_E \eta\varphi(u)\langle \nu_x, \psi\rangle dx$$

or

$$(\varphi \otimes \psi)(u_j, v_j) \overset{*}{\rightharpoonup} \langle \delta_{u(\cdot)} \otimes \nu_\cdot, \varphi \otimes \psi\rangle \text{ in } L^\infty(E).$$

The assertion follows since linear combinations of tensor products $\varphi \otimes \psi$ are dense in $C_0(\mathbf{R}^{d+d'})$. □

A typical application of the corollaries is as follows. Let $f : \Omega \times (\mathbf{R}^m \times M^{m\times n}) \to \mathbf{R}$ be a Carathéodory function and suppose that $f \geq 0$. Suppose that $u_j \rightharpoonup u$ in $W^{1,p}(\Omega; \mathbf{R}^m)$ and that Du_j generates the Young measure ν. Taking $v_j = Du_j, z_j = (u_j, v_j)$ we obtain.

$$\begin{aligned}
&\liminf_{j\to\infty} \int_\Omega f(x, u_j(x), Du_j(x))dx \\
&\geq \int_\Omega \int_{\mathbf{R}^m \times M^{m\times n}} f(x, \lambda, \mu)d\delta_{u(x)}(\lambda) \otimes d\nu_x(\mu)dx \\
&= \int_\Omega \int_{M^{m\times n}} f(x, u(x), \lambda)d\nu_x(\lambda)dx.
\end{aligned}$$

The proof of the lower semicontinuity is thus reduced to the verification of the inequality

$$\int_{M^{m\times n}} g(\lambda)\nu_x(\lambda) \geq g(Du(x)) = g(\langle \nu_x, \mathrm{id}\rangle) \tag{3.12}$$

for the function

$$g(\lambda) = f(x, u(x), \lambda)$$

with 'frozen' first and second argument. To see when (3.12) holds we need to understand which Young measures are generated by gradients. This is the topic of the next section.

4 Which Young measures arise from gradients?

To employ Young measures in the study of crystal microstructure we need to understand which Young measures arise from sequences of gradients $\{Du_j\}$. As before $\Omega \subset \mathbf{R}^n$ denotes a bounded domain with Lipschitz boundary.

Definition 4.1 *A (weakly∗ measurable) map* $\nu : \Omega \to \mathcal{M}(M^{m\times n})$ *is a* $W^{1,p}$ gradient Young measure *if there exists a sequence of maps* $u_j : \Omega \to \mathbf{R}^m$ *such that*

$$u_j \rightharpoonup u \quad in \quad W^{1,p}(\Omega; \mathbf{R}^m) \quad (\overset{*}{\rightharpoonup} \ \ if\, p = \infty),$$
$$\delta_{Du(\cdot)} \overset{*}{\rightharpoonup} \nu \quad L^\infty_w(\Omega; \mathcal{M}(M^{m\times n})).$$

Using this notion we may reformulate Problem 2 (approximate solutions) as follows.

Problem 2' Given a set $K \subset M^{m\times n}$, characterize all $W^{1,\infty}$- gradient Young measures ν such that

$$\mathrm{supp}\nu_x \subset K \qquad \text{for} \quad \text{a.e. } x.$$

An abstract characterization of gradient Young measures due to Kinderlehrer and Pedregal will be derived in Section 4.3 below. It involves the notion of quasiconvexity. Quasiconvexity, first introduced by Morrey in 1952, is clearly the natural notion of convexity for vector-valued problems (see Section 4.2) but still remains largely mysterious since it is very hard to determine whether a given function is quasiconvex. Therefore further notions of convexity were introduced to obtain necessary or sufficient conditions for quasiconvexity. We begin by reviewing these notions and their relationship.

4.1 Notions of convexity

For a matrix $F \in M^{m\times n}$ let $\boldsymbol{M}(F)$ denote the vector that consists of all minors of F and let $d(n,m) = \sum_{r=1}^{\min(n,m)} \binom{n}{r}\binom{m}{r}$ denote its length.

Definition 4.2 *A function* $f : M^{m\times n} \to \mathbf{R} \cup \{+\infty\} = (-\infty, \infty]$ *is*

(i) *convex if*

$$f(\lambda A + (1-\lambda)B) \le \lambda f(A) + (1-\lambda) f(B)$$
$$\forall\, A, B \in M^{m\times n}, \lambda \in (0,1);$$

(ii) *polyconvex if there exists a convex function* $g : \mathbf{R}^{d(n,m)} \to \mathbf{R} \cup \{+\infty\}$ *such that*

$$f(F) = g(\boldsymbol{M}(F));$$

(iii) *quasiconvex if for every open and bounded set* U *with* $|\partial U| = 0$ *one has*

$$\int_U f(F + D\varphi)dx \geq \int_U f(F)dx = |U|f(F) \quad \forall \varphi \in W_0^{1,\infty}(U; \mathbf{R}^m), \quad (4.1)$$

whenever the integral on the left hand side exists;

(iv) *rank-1 convex, if* f *is convex along rank-1 lines, i.e. if*

$$f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda)f(B)$$

$$\forall\, A, B \in M^{m\times n} \ \textit{with} \ \mathrm{rk}(B-A) = 1, \quad \forall\, \lambda \in (0,1).$$

Remarks. 1. If $f \in C^2$ then rank-1 convexity is equivalent to the Legendre-Hadamard condition

$$\frac{\partial^2 f}{\partial F^2}(F)(a \otimes b, a \otimes b) = \frac{\partial^2 f}{\partial F^i_\alpha \partial F^j_\beta}(F) a^i b_\alpha a^j b_\beta \geq 0.$$

2. Quasiconvexity is independent of the set U, i.e. if (4.1) holds for one open and bounded set with $|\partial U| = 0$ then it holds for all such sets. If f takes values in $\mathbf{R}$ it suffices to extend φ by zero outside U and to translate and scale U. For general f one can use the Vitali covering theorem.

3. If f takes values in $\mathbf{R}$ and is quasiconvex then it is rank-1 convex (see Lemma 4.3 below) and thus locally Lipschitz continuous (use that f is convex and thus locally Lipschitz in each coordinate direction in $M^{m\times n}$; see [Da 89], Chapter 2, Thm. 2.3, or [MP 98], Observation 2.3 for the details). In this case the integral on the left hand side of (4.1) always exists.

It is sometimes convenient to consider quasiconvex functions that take values in $[-\infty, \infty)$. The argument below shows that such functions are rank-1 convex and thus either take values in $\mathbf{R}$ or are identically $-\infty$.

If $n = 1$ or $m = 1$ then convexity, polyconvexity and rank-1 convexity are equivalent and they are equivalent to quasiconvexity if, in addition, f takes values in $\mathbf{R}$.

Lemma 4.3 *If $n \geq 2, m \geq 2$ then the following implications hold.*

$$\begin{array}{lll} f & \textit{convex} & \\ & \Downarrow & \not\Uparrow \\ f & \textit{polyconvex} & \\ & \Downarrow & \not\Uparrow \\ f & \textit{quasiconvex} & \\ & \Downarrow f < \infty & \not\Uparrow \textit{ if } m \geq 3 \\ f & \textit{rank-1 convex} & \end{array}$$

The most difficult question is whether rank-1 convexity implies quasiconvexity. Šverák's [Sv 92a] ingenious counterexample solved this long standing problem in the negative if $m \geq 3$; the case $m = 2, n \geq 2$ is completely open.

Proof. The first implication is obvious, the second follows from the fact that minors are null Lagrangians (see Theorem 2.3) and Jensen's inequality. To prove the last implication let f be quasiconvex, consider $A, B \in M^{m\times n}$, with $\mathrm{rk}(B - A) = 1$, and a convex combination $F = \lambda A + (1-\lambda)B$. After translation and rotation we may assume that $F = 0, A = (1-\lambda)a \otimes e_1$, $B = -\lambda a \otimes e_1$. Let h be a 1-periodic sawtooth function which satisfies $h(0) = 0$, $h' = (1-\lambda)$ on $(0, \lambda)$ and $h' = -\lambda$ on $(\lambda, 1)$. Define for $x \in Q = (0,1)^n$

$$\begin{aligned} u_k(x) &= ak^{-1}h(kx^1), \\ v_k(x) &= a \min\{k^{-1}h(kx^1), \mathrm{dist}_\infty(x, \partial Q)\}, \end{aligned}$$

where

$$\begin{aligned} \mathrm{dist}_\infty(x, \partial Q) &= \inf\{\|x - y\|_\infty \colon y \in \partial Q\}, \\ \|x\|_\infty &= \sup\{|x^i|, i = 1, \ldots, n\}. \end{aligned}$$

Then $Dv_k \in \{A, B\} \cup \{\pm a \otimes e_i\}, v_k = 0$ on ∂Q, and $|\{Dv_k \neq Du_k\}| \to 0$ as $k \to 0$ (see Fig. 12).

It follows from the definition of quasiconvexity that

$$\lambda f(A) + (1-\lambda)f(B) = \lim_{k\to\infty} \int_Q f(Du_k)dx = \lim_{k\to\infty} \int_Q f(Dv_k)dx \geq f(0),$$

as desired. Note that the inequality $\lambda f(A) + (1-\lambda)f(B) \geq f(0)$ still holds if f takes values in $[-\infty, \infty)$.

As for the reverse implications, the minors (subdeterminants) of order greater than one are trivially polyconvex but not convex. An example of a

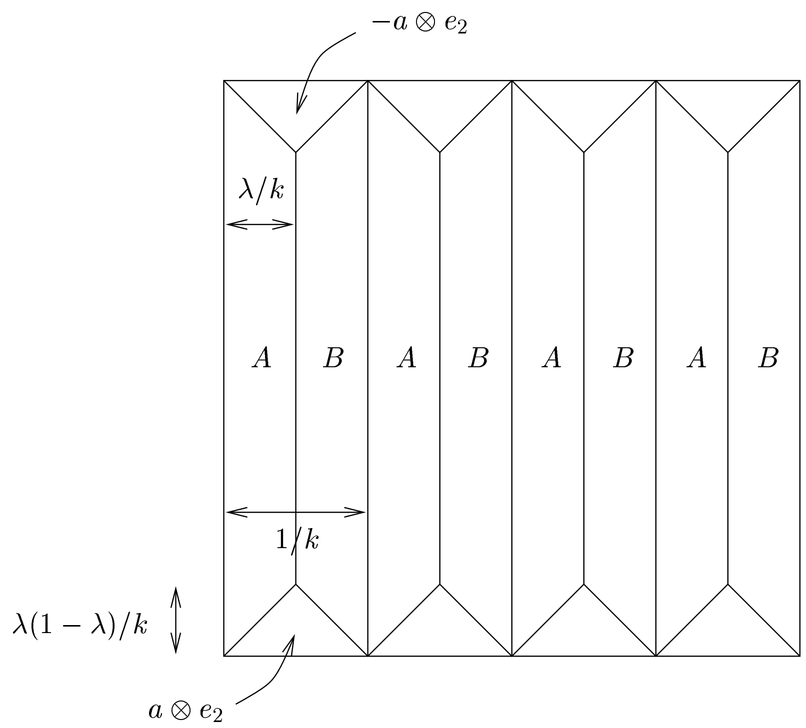

Figure 12: The gradients of v_k, for $n = 2$.

quasiconvex but not polyconvex function is given below. Šverák's counterexample of a rank-1 convex function that is not quasiconvex will be discussed in Section 4.7. □

Remark. The proof that quasiconvexity implies rank-1 convexity is similar to Fonseca's ([Fo 88], Theorem 2.4). In fact her method yields a slightly stronger result: if $f : M^{m\times n} \to [-\infty, \infty]$ is finite in a neighbourhood of F and quasiconvex then f does not take the value $-\infty$ on any rank–1 line through F and f is rank-1 convex at F, i.e. $f(F) \leq \lambda f(F - (1-\lambda)a \otimes b)$ $+(1-\lambda)f(F + \lambda a \otimes b)$, $\forall\ a \in \mathbf{R}^n, b \in \mathbf{R}^m, \lambda \in (0,1)$. To obtain this refinement it suffices to replace $\text{dist}_\infty(x, Q)$ in the definition of v_k by $\epsilon\ \text{dist}_\infty(x, Q)$ for small enough $\epsilon > 0$.

The following example, due to Dacorogna and Marcellini [AD 92], [DM 88], [Da 89], may serve as a simple illustration of the different notions of convexity. Let $n = m = 2$ and consider

$$f(F) = |F|^4 - \gamma |F|^2 \det F. \tag{4.2}$$

Then

$$\begin{array}{llll} f & \text{convex} & \Longleftrightarrow & |\gamma| \le \frac{4}{3}\sqrt{2}, \\ f & \text{polyconvex} & \Longleftrightarrow & |\gamma| \le 2, \\ f & \text{quasiconvex} & \Longleftrightarrow & |\gamma| \le 2 + \epsilon, \\ f & \text{rank-1 convex} & \Longleftrightarrow & |\gamma| \le \frac{4}{\sqrt{3}}. \end{array}$$

It is known that $\epsilon > 0$; whether or not $2 + \epsilon = \frac{4}{\sqrt{3}}$ is open.

Alberti raised the following interesting question which shows how little we know about quasiconvexity. Let $2 \le n \le m$ and let $g : M^{m\times n} \to \mathbf{R}$, $\tilde{g} : M^{n\times m} \to \mathbf{R}, \tilde{g}(F) = g(F^T)$.

Question (Alberti): g quasiconvex $\overset{?}{\Longleftrightarrow}$ $\tilde{g}$ quasiconvex.

Obviously equivalence holds for the other three notions of convexity. Kružik recently answered Alberti's question in the negative if g is allowed to take the value $+\infty$ and $m \ge 3$. Refining his argument one can show that Šverák's quartic polynomial provides a finite-valued counterexample (see the end of section 4.7).

Ball, Kirchheim and Kristensen [BKK 98] recently solved a long-standing problem by proving that the quasiconvex hull of a C^1 function f (i.e. the largest quasiconvex function below f) is again C^1, provided that f satisfies polynomial growth conditions. The representation of the quasiconvex hull through gradient Young measures (see Section 4.3) plays a crucial rôle in their argument.

4.2 Properties of quasiconvexity

Quasiconvexity is the fundamental notion of convexity for vector-valued variational problems. It is closely related to lower semicontinuity of integral functionals, existence and regularity of minimizers and the passage from microscopic and macroscopic energies. Quasiconvex functions are the natural dual objects to gradient Young measures (see Section 4.3).

In the following Ω always denotes a bounded (Lipschitz) domain in $\mathbf{R}^n$ and we consider maps $u : \Omega \to \mathbf{R}^m$ and the functional

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

In this section we merely summarize the results. Some of the proofs for $p = \infty$ are given in Sections 4.8 and 4.9 below. Further comments and references can be found at the end of these notes.

Theorem 4.4 *Suppose that $f : M^{m\times n} \to \mathbf{R}$ is continuous.*

(i) *The functional I is weak$*$ sequentially lower semicontinuous (w^*slsc) on $W^{1,\infty}(\Omega; \mathbf{R}^m)$ if and only if f is quasiconvex.*

(ii) *Suppose, in addition, that*

$$0 \le f(F) \le C(|F|^p + 1) \tag{4.3}$$

for some $p \in [1, \infty)$. If f is quasiconvex then I is wslsc on $W^{1,p}(\Omega; \mathbf{R}^m)$.

Remarks. If $f \ge 0$ it can be shown that I is finite and wslsc on $W^{1,p}$ if and only if f satisfies (4.3) and is quasiconvex (see [Kr 94]). Part (i) is an essential ingredient in the classification of gradient Young measures. Using this classification and simple general facts about Young measures (see Section 3.4) one easily obtains similar lower semicontinuity results for integrands $f(x, u(x), Du(x))$.

Theorem 4.5 *(existence and relaxation)*
Suppose that $p \in (1, \infty), c > 0$ and that f satisfies

$$c|F|^p \le f(F) \le C(|F|^p + 1).$$

(i) *If f is quasiconvex and $v \in W^{1,p}(\Omega; \mathbf{R}^m)$ then I attains its minimum in the class*

$$W^{1,p}_v(\Omega; \mathbf{R}^m) := \{u \in W^{1,p}(\Omega; \mathbf{R}^m) : u - v \in W^{1,p}_0(\Omega; \mathbf{R}^m)\}.$$

(ii) *If f^{qc} denotes the quasiconvex envelope of f, i.e. the largest quasiconvex function below f, then*

$$\inf_{W_v^{1,p}} I = \min_{W_v^{1,p}} \bar{I},$$

where

$$\bar{I}(u) = \int_\Omega f^{qc}(Du). \tag{4.4}$$

Moreover, a function $\bar{u}$ is a minimizer of $\bar{I}$ in $W_v^{1,p}$ if and only if it is a cluster point (with respect to weak convergence in $W^{1,p}$) of a minimizing sequence for I.

(iii) *For any $f : M^{m\times n} \to [-\infty, \infty)$ and every bounded domain U with $|\partial U| = 0$ one has*

$$f^{qc}(F) = \inf_{\varphi \in W_0^{1,\infty}} \frac{1}{|U|} \int_U f(F + D\varphi)dx. \tag{4.5}$$

The passage from I to $\bar{I}$ is called relaxation. It replaces a variational problem which may have no solution by one which has a solution. This sounds almost too good to be true and indeed there is a price to pay. The minimizers of $\bar{I}$ are in general only weak limits of a minimizing sequence of I, and important features of the sequence may be lost. If, for example, $\bar{I}$ has a homogeneous minimizer it is not clear whether minimizing sequences of I are (nearly) homogeneous or whether they involve an increasingly finer mixture of several states. A different approach, that keeps more information about the minimizing sequence is to derive a (relaxed) problem for the gradient Young measures generated by minimizing sequences (see Theorem 4.9 of the next section).

Physically, relaxation corresponds to the passage from a microscopic energy I to a macroscopic energy $\bar{I}$, which is obtained by averaging over fine scale oscillations; cf. the representation (4.5).

Theorem 4.6 *(regularity). Suppose that f is smooth, satisfies*

$$0 \le f(F) \le C(|F|^2 + 1)$$

and is uniformly quasiconvex, i.e. there exists $c > 0$ such that

$$\int_U [f(F + D\varphi) - f(F)]dx \ge c \int_U |D\varphi|^2, \quad \forall \varphi \in W_0^{1,\infty}(\Omega; \mathbf{R}^m).$$

Let $\bar{u} \in W^{1,2}(\Omega; \mathbf{R}^m)$ be a local minimizer of I, i.e.

$$I(\bar{u} + \varphi) \geq I(\bar{u}) \quad \forall \varphi \in C_0^\infty(\Omega).$$

Then there exists an open set Ω_0 of full measure such that

$$u \in C^\infty(\Omega_0).$$

4.3 Classification of gradient Young measures

Recall that a map $\nu : \Omega \to \mathcal{M}(M^{m\times n})$ is a $W^{1,\infty}$ gradient Young measure if it is the Young measure generated by a sequence of gradients Du_j, where $u_j \overset{*}{\rightharpoonup} u \quad W^{1,\infty}$ (see Definition 4.1).

Theorem 4.7 *([KP 91]) A (weak$*$ measurable) map $\nu : \Omega \to \mathcal{M}(M^{m\times n})$ is a $W^{1,\infty}$ gradient Young measure if and only if $\nu_x \geq 0$ a.e. and there exists a compact set $K \subset M^{m\times n}$ and $u \in W^{1,\infty}(\Omega; \mathbf{R}^m)$ such that the following three conditions hold.*

(i) $\mathrm{supp}\nu_x \subset K$ *for* a.e. x,

(ii) $\langle \nu_x, \mathrm{id}\rangle = Du$ *for* a.e. x,

(iii) $\langle \nu_x, f\rangle \geq f(\langle \nu_x, \mathrm{id}\rangle)$ *for* a.e. x *and all quasiconvex* $f : M^{m\times n} \to \mathbf{R}$.

Remarks. 1. The key point is (iii) which is in nice duality with the definition of quasiconvexity. Roughly speaking, quasiconvex functions satisfy Jensen's inequality for gradients, while gradient Young measures must satisfy Jensen's inequality for all quasiconvex functions.

2. Let $K \subset M^{m\times n}$ be compact. For future reference we define the set of nonnegative measures supported on K that satify condition (iii) of the theorem as

$$\begin{aligned}\mathcal{M}^{qc}(K) = \{\nu \in \mathcal{M}(M^{m\times n}) : \nu \geq 0, \mathrm{supp}\nu \subset K,& \\ \langle \nu, f\rangle \geq f(\langle \nu, \mathrm{id}\rangle) \quad \forall f : M^{m\times n} \to \mathbf{R} \quad \text{quasiconvex}\}.&\end{aligned} \tag{4.6}$$

By the theorem $\mathcal{M}^{qc}(K)$ consist exactly of the homogeneous gradient Young measures supported in K. Similarly one defines $\mathcal{M}^{rc}(K)$ and $\mathcal{M}^{pc}(K)$ using rank-one convex and polyconvex functions, respectively.

3. Every minor M is a quasiaffine function (i.e. (4.1) holds with equality; see Theorem 2.3(i)). Hence application of (iii) with $\pm M$ yields the minors relations

$$\langle \nu_x, M \rangle = M(\langle \nu_x, \mathrm{id} \rangle) \tag{4.7}$$

as a necessary condition for gradient Young measures. This condition in fact follows directly from Theorem 2.3 (ii) and does not require Theorem 4.7. The minors relations often prove very useful for problems with large symmetries that arise e.g. in models of microstructure in crystals (see e.g. [Bh 92]). They are, however, far from being sufficient in general.

Exercise. Find a nontrivial measure supported on three diagonal 2×2 matrices without rank-1 connections that satisfies (4.7) and compare with Theorem 2.5.

Hint: Look for matrices on the two hyperbolae given by $\{\det = 1\}$.

Proof of Theorem 4.7 (necessity). Conditions (i) and (ii) follow from basic facts about Young measures (see Theorem 3.1 (ii) and (iii)) while (iii) follows from Morrey's lower semicontinuity result (Theorem 4.3(ii)), applied to all open subsets U of Ω. Sufficiency is discussed in Section 4.9. □

To finish this section we briefly mention the analogous result for $p < \infty$ and its relation to relaxation and generalized solutions. This may be omitted on first reading.

Theorem 4.8 *([KP 94]) Let $p \in [1, \infty)$. A (weakly measurable) map $\nu : \Omega \to \mathcal{M}(M^{m\times n})$ is a $W^{1,p}$ gradient Young measure if and only if $\nu_x \geq 0$ a.e. and the following three conditions hold*

(i) $\int\limits_\Omega \int\limits_{M^{m\times n}} |F|^p d\nu_x(F) dx < \infty$;

(ii) $\langle \nu_x, \mathrm{id} \rangle = Du, \quad u \in W^{1,p}(\Omega; \mathbf{R}^m)$;

(iii) $\langle \nu_x, f \rangle \geq f(\langle \nu_x, \mathrm{id} \rangle)$ *for* a.e. x *and all quasiconvex* f *with* $|f|(F) \leq C(|F|^p + 1)$.

Young measures arise naturally as generalized solutions of variational problems that have no classical solution. To this end extend the functional

$$I(u) = \int\limits_\Omega f(Du) dx$$

on functions to a functional

$$J(\nu) = \int_\Omega \langle \nu_x, f \rangle dx$$

on Young measures. For $v \in W^{1,p}(\Omega; \mathbf{R}^m)$ consider the admissible classes

$$\begin{aligned}
\mathcal{A} &= \{u \in W^{1,p}(\Omega; \mathbf{R}^m) : u - v \in W_0^{1,p}(\Omega; \mathbf{R}^m),\\
\mathcal{G} &= \{\nu : \Omega \to \mathcal{M}(\mathbf{R}^m) : \nu \quad W^{1,p} \text{ gradient Young measure},\\
&\qquad \langle \nu_x, \mathrm{id} \rangle = Du(x), u \in \mathcal{A}\}.
\end{aligned}$$

Theorem 4.9 *Suppose that f is continuous and satisfies $c|F|^p \le f(F) \le C(|F|^p + 1), c > 0, p > 1$. Then*

$$\inf_{\mathcal{A}} I = \min_{\mathcal{G}} J(\nu).$$

Moreover the minimizers of J are Young measures that are generated by gradients of minimizing sequences of I.

In particular, I has a minimizer in $\mathcal{A}$ if and only if there exists a minimizer ν of J such that ν_x is a Dirac mass for a.e. x.

4.4 Convex hulls and resolution of Problem 3

To return to the setting of Sections 1 and 2 we extend the different notions of convexity from functions to sets. We first recall that the quasiconvex (convex, polyconvex, rank-1 convex) envelope or hull of a function $f : M^{m\times n} \to \mathbf{R}$ is the largest quasiconvex (convex, polyconvex, rank-1 convex) function below f and is denoted by f^{qc} ($f^c = f^{**}, f^{pc}, f^{rc}$). Similarly the quasiconvex hull of a set $K \subset M^{m\times n}$ is defined via sublevel sets as

$$K^{qc} = \{F \in M^{m\times n} : f(F) \le \inf_K f, \quad \forall\, f : M^{m\times n} \to \mathbf{R} \text{ quasiconvex}\},$$

with similar definitions for K^c, K^{pc}, K^{rc}. Note that K^c is the *closed* convex hull. A set is called quasiconvex if $K = K^{qc}$. In the case of rank-1 convexity one can also define a hull by pointwise operations rather than by sublevel sets. A set K is called *lamination convex* if the conditions $A, B \in K$ and $\mathrm{rk}(B - A) = 1$ imply that convex combinations of A and B are in K. The lamination convex hull K^{lc} of K is the smallest lamination convex set containing K. It is easy to verify that K^{lc} can equivalently be defined by inductively adding rank-1 segments:

$$K^{lc} := \bigcup_{i=1}^{\infty} K^{(i)}, \qquad K^{(1)} := K,$$

$$K^{(i+1)} := K^{(i)} \cup \{\lambda A + (1-\lambda)B : A, B \in K^{(i)}, \mathrm{rk}(B-A) = 1, \lambda \in (0,1)\}.$$

One has the following inclusions (see Lemma 4.3):

$$K^{lc} \subset K^{rc} \subset K^{qc} \subset K^{pc} \subset K^{c}. \tag{4.8}$$

The example in Section 2.5 shows that in general $K^{lc} \neq K^{rc}$. In this example $K^{lc} = K, K^{rc} \supset K \cup \{\mathrm{diag}(\lambda, \mu) : |\lambda| \leq 1, |\mu| \leq 1\}$. The characterization of laminates (see Section 4.6) as well as recent work of Matoušek and Plecháč [MP 98] suggest that K^{rc} is the more natural object, but more difficult to handle (Matoušek and Plecháč use the terms set-theoretic rank-1 convex hull and functional rank-1 convex hull to distinguish K^{lc} and K^{rc}).

The polyconvex hull is closest to the ordinary (closed) convex hull and is in fact the intersection of a convex set with a nonconvex constraint. Let $\mathbf{M}(F)$ denote the vector of all minors of F (see Section 4.1) and let

$$\hat{K} = \{\boldsymbol{M}(F) : F \in K\}.$$

Exercise. Show that

$$K^{pc} = \{F : \boldsymbol{M}(F) \in (\hat{K})^c\} \tag{4.9}$$

and moreover

$$K^{pc} = \{\langle \nu, \mathrm{id} \rangle : \nu \in \mathcal{M}^{pc}(K)\}.$$

With this notation in place we have the following abstract resolution of Problem 3 (see Section 1.4). Recall that the set K^{app} (interpreted as the macroscopically stress free affine deformations) was the set of all matrices F such that there exists a sequence u_j bounded in $W^{1,\infty}(\Omega; \mathbf{R}^m)$, such that

$$\mathrm{dist}(Du_j, K) \to 0 \qquad \text{in measure in } \Omega, \tag{4.10}$$

$$u_j = Fx \qquad \text{on } \partial\Omega, \tag{4.11}$$

and that $\mathcal{M}^{qc}(K)$ denotes the set of homogeneous gradient Young measures (see (4.6)).

Theorem 4.10 *Suppose that K is compact and denote by dist_K the distance function from K. Then*

(i) $K^{app} = K^{qc}$,

(ii) $K^{qc} = \{\operatorname{dist}_K^{qc} = 0\}$,

(iii) K^{qc} *is the set of barycentres of homogeneous gradient Young measures:*

$$K^{qc} = \{\langle \nu, \mathrm{id} \rangle : \nu \in \mathcal{M}^{qc}(K)\}.$$

Proof. After dilation we may assume that $|\Omega| = 1$.

(i) To show that $K^{app} \subset K^{qc}$ let $F \in K^{app}$ and let $f : M^{m\times n} \to \mathbf{R}$ be quasiconvex and suppose that $\{u_j\}$ is bounded in $W^{1,\infty}$ and satisfies (4.10) and (4.11). We may assume that $\inf_K f = 0$ and we need to show $f(F) \leq 0$. Since f is continuous (see Remark 3 after Definition 4.2) we have $f^+(Du_j) \to 0$ in measure, and $|f^+(Du_j)| \leq C$ since Du_j is bounded in L^∞. Quasiconvexity, (4.11) and dominated convergence yield

$$|\Omega| f(F) \leq \liminf_{j\to\infty} \int_\Omega f(Du_j) dx \leq \liminf_{j\to\infty} \int_\Omega f^+(Du_j) dx = 0.$$

To prove the converse inclusion $K^{qc} \subset K^{app}$ let $F \in K^{qc}$. Then $\operatorname{dist}_K^{qc}(F) = 0$ by definition of K^{qc}. In view of the representation formula for $\operatorname{dist}_K^{qc}$ (Theorem 4.5 (iii)) there exist $\varphi_j \in W_0^{1,\infty}(\Omega; \mathbf{R}^m)$ such that

$$0 = \operatorname{dist}_K^{qc}(F) = \lim_{j\to\infty} \int_\Omega \operatorname{dist}_K(F + D\varphi_j) dx.$$

The functions $u_j(x) = Fx + \varphi_j$ thus satisfy (4.10) and (4.11). The problem is that a priori Du_j only needs to be bounded in L^1 (in fact weakly relatively compact in L^1) and may not be bounded in L^∞. Zhang's lemma (see Lemma 4.21 (ii) below) assures that u_j can be modified on small sets such that (4.10) and (4.11) hold and Du_j is bounded in L^∞.

(ii) The inclusion $\subset$ follows from the definition of K^{qc}. On the other hand we have just shown that $\operatorname{dist}_K^{qc}(F) = 0$ implies $F \in K^{app} = K^{qc}$.

(iii) Let $\nu \in \mathcal{M}^{qc}(K)$. By definition $f(\langle \nu, \mathrm{id} \rangle) \leq \langle \nu, f \rangle \leq \sup_K f$ and hence $\langle \nu, \mathrm{id} \rangle \in K^{qc}$. Suppose conversely that $F \in K^{qc}$. We need to show that there exists $\nu \in \mathcal{M}^{qc}(K)$ with $\langle \nu, \mathrm{id} \rangle = F$. After an affine transformation we may assume $F = 0$. By part (i) there exists a sequence u_j (bounded in $W^{1,\infty}$) that satisfies (4.10) and (4.11). Passing to a subsequence we may assume that $\{Du_j\}$ generates the Young measure ν and $Du_j \stackrel{*}{\rightharpoonup} Du$ in $L^\infty(\Omega; M^{m\times n})$. By the divergence theorem

$$\int_\Omega \langle \nu_x, \mathrm{id} \rangle dx = \int_\Omega Du \, dx = 0. \tag{4.12}$$

To obtain a homogeneous Young measure we define the average $\mathrm{Av}\,\nu$ by duality as the unique Radon measure that satisfies

$$\langle \mathrm{Av}\,\nu, f \rangle = \frac{1}{|\Omega|} \int_\Omega \langle \nu_x, f \rangle dx \quad \forall f \in C_0(M^{m\times n}).$$

By Theorem 4.7 we have $\nu_x \in \mathcal{M}^{qc}(K)$ for a.e. x and hence $\mathrm{Av}\,\nu \in \mathcal{M}^{qc}(K)$. Moreover (4.12) yields $\langle \mathrm{Av}\nu, \mathrm{id} \rangle = 0$ as desired. □

4.5 The two-well problem

To see what the various convexity notions can do to understand microstructure in crystals we consider the two-well problem in two dimensions. This is the simplest multiphase problem consistent with the rotational symmetry and was analyzed completely in a beautiful paper by Šverák [Sv 93a]. Let

$$K = SO(2)A \cup SO(2)B \subset M^{2\times 2}, \quad \det B \geq \det A > 0. \tag{4.13}$$

Various normalizations are possible. Multiplication by A^{-1}, polar decomposition and diagonalization show, for example, that it suffices to consider

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \quad 0 < \lambda \leq \mu, \quad \lambda\mu \geq 1. \tag{4.14}$$

The first step towards the resolution of Problems 1-3 is to look for rank-1 connections in K.

Exercise. Prove the following classification.

(i) If $\lambda > 1$ then there are no rank-1 connections in K;

(ii) if $\lambda = 1$ (and $A \neq B$) each matrix in K is rank-1 connected to exactly one other matrix in K;

(iii) if $\lambda < 1$ each matrix in K is rank-1 connected to exactly two other matrices in K.

Theorem 4.11 *Suppose that K given by (4.13) contains no rank-1 connections. Then every Young measure $\nu : \Omega \to \mathcal{M}(M^{2\times 2})$ with* $\mathrm{supp}\nu_x \subset K$ *is a constant Dirac mass. Moreover*

$$K^{lc} = K^{rc} = K^{qc} = K^{pc} = K \tag{4.15}$$

Remark. It is not known whether the same result holds for $K = SO(3)A \cup SO(3)B \subset M^{3\times 3}$; some special cases are known ([Sv 93a], [Ma 92]).

Proof. The crucial observation is that

$$\det(F - G) > 0 \quad \forall F, G \in K, F \neq G. \tag{4.16}$$

By symmetry and $SO(2)$ invariance it suffices to verify this for $G = \mathrm{Id}$. The inequality clearly holds for $F = B$ (by the above exercise) and hence by connectedness and the absence of rank-1 connections for $G \in SO(2)B$. Similarly $\det(\mathrm{Id} - (-\mathrm{Id})) > 0$ and hence by connectedness (4.16) holds also for all other $G \in SO(2)$.

To determine K^{qc} consider first a homogeneous gradient Young measure ν supported in K and let $\bar{\nu} = \langle \nu, \mathrm{id} \rangle$ denote its barycentre. We have for $F, G \in M^{2\times 2}$

$$\det(F - G) = \det F - \mathrm{cof} F : G + \det G,$$

where $F : G = tr F^t G = \sum_{i,j} F_{ij} G_{ij}$. The minors relations yield

$$\begin{aligned} 0 &\leq \int_{M^{2\times 2}\times M^{2\times 2}} \det(F - G) d\nu(F) d\nu(G) \\ &= \int_{M^{2\times 2}\times M^{2\times 2}} (\det F - \mathrm{cof} F : G + \det G) d\nu(F) d\nu(G) \\ &= \int_{M^{2\times 2}} (\det \bar{\nu} - \mathrm{cof}\ \bar{\nu} : G + \det G) d\nu(G) \\ &= \det \bar{\nu} - \mathrm{cof}\ \bar{\nu} : \bar{\nu} + \det \bar{\nu} = \det(\bar{\nu} - \bar{\nu}) = 0. \end{aligned}$$

Hence the first inequality must be an equality, and (4.16) implies that the product measure $\nu \otimes \nu$ is supported on the diagonal of $M^{2\times 2} \times M^{2\times 2}$. Hence ν must be a Dirac mass. This implies $K^{qc} = K$ by Theorem 4.10. Since the argument used only the minors relations we even have $K^{pc} = K$.

Now let $\nu : \Omega \to \mathcal{M}(M^{m\times n})$ be an arbitrary gradient Young measure with $\text{supp}\nu_x \subset K$ a.e. By the above argument $\nu_x = \delta_{Du(x)}$ and $Du(x) \in K$ a.e. We show that $Du \equiv \text{const}$. To this end observe first that (4.16) can be strengthened to

$$\det(X - Y) \geq c|X - Y|^2, \quad c > 0, \quad \forall X, Y \in K. \tag{4.17}$$

Indeed by compactness and $SO(2)$ invariance it suffices to verify that the tangent space of $SO(2)$ at the identity contains no rank-1 connections. This is obvious. Now let e be a unit vector in $\mathbf{R}^2$, and for $0 < h < 1$ consider the translates $v(x) = u(x + he)$ and a cut-off function $\varphi \in C_0^\infty(\Omega)$. Since the determinant is a null Lagrangian (see Theorem 2.3(i)) integration of (4.17) yields

$$\begin{aligned}
& c\int_\Omega \varphi^2 |Du - Dv|^2\,dx \leq \int_\Omega \det[\varphi(Du - Dv)]\,dx \\
\leq\ & \int_\Omega \det[D(\varphi(u - v))]\,dx - \int_\Omega \operatorname{cof} D(\varphi(u - v)) : (u - v) \otimes D\varphi\,dx \\
& + \int_\Omega \det[(u - v) \otimes D\varphi]\,dx \\
\leq\ & \frac{c}{2}\int_\Omega \varphi^2 |Du - Dv|^2\,dx + C\int_\Omega |D\varphi|^2 |u - v|^2\,dx.
\end{aligned}$$

Hence the difference quotients $\varphi\frac{Du-Dv}{h}$ are uniformly bounded in L^2 and thus $Du \in W^{1,2}_{\text{loc}}(\Omega; M^{m\times n})$. Therefore Du can only take values in one connected component of K, and the assertion follows from Theorem 2.4. $\square$

To consider the case where K has rank-1 connections it is convenient to introduce new coordinates on $M^{2\times 2}$. Since A and B are not conformally equivalent (as $\lambda < \mu$), for every matrix F there exist a unique pair $(y, z) \in \mathbf{R}^2 \times \mathbf{R}^2$ such that

$$F = \begin{pmatrix} y_1 & -y_2 \\ y_2 & y_1 \end{pmatrix} + \begin{pmatrix} z_1 & -z_2 \\ z_2 & z_1 \end{pmatrix} B.$$

Theorem 4.12 *Suppose that K is given by (4.13), (4.14) and that $\lambda < 1$. Then*

$$K^{lc} = K^{rc} = K^{qc} = K^{pc}$$

and

$$\begin{array}{ll} K^{pc} = K^c \cap \{\det = 1\}, & \textit{if } \det B = 1, \\ K^{pc} = \{F = (y,z) : |y| \leq \frac{\det B - \det F}{\det B - 1}, \quad |z| \leq \frac{\det F - 1}{\det B - 1}\} & \\ \qquad \textit{if } \det B > 1. & \end{array}$$

To characterize the polyconvex hull we use the following

Proposition 4.13 *The convex hull of the set $\mathcal{K} \subset \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}$ given by $\mathcal{K} = \{(y, 0, a) : |y| = 1\} \cup \{(0, z, b) : |z| = 1\}$ is given by*

$$\mathcal{K}^c = \begin{cases} \{(y,z,a), |y| + |z| \leq 1\} & \textit{if } a = b, \\ \{(y,z,t), |y| \leq \frac{b-t}{b-a}, |z| \leq \frac{t-a}{b-a}\} & \textit{if } a < b. \end{cases}$$

Proof. This is obvious for $n = 1$, and the general case follows by invariance under $(y, z, t) \to (Ry, Qz, t), R, Q \in SO(n)$. □

Proof of Theorem 4.12. The formula for K^{pc} follows from the characterization (4.9) and Proposition 4.13. In view of the general relation (4.8) between the different convex hulls it only remains to show that $K^{lc} = K^{pc}$.

First case: $\det B = 1$.

Let $F \in K^{pc} = K^c \cap \{\det = 1\}$. If $F = (y, z) \in \partial K^c$, then by Proposition 4.13 we have $|y| + |z| = 1$. If $y = 0$ or $z = 0$ then $F \in K$ and we are done. If $y \neq 0, z \neq 0$ then we can consider $G(t) = (ty/|y|, (1-t)z/|z|)$. Let $g(t) = \det G(t)$. Then g is quadratic in t and $g(0) = g(1) = g(|y|) = 1$. Hence $g \equiv 1$ and $t \to G(t)$ must be a rank-1 line. This shows that F is a rank-1 combination of $(y/|y|, 0) \in K$ and $(0, z/|z|) \in K$.

If $F \in \text{int} K^c \cap \{\det = 1\}$ then there exist $a, n \in S^1$ such that $\text{cof} F : a \otimes n = Fn \cdot a = 0$. Hence the determinant is constant on the line $F + ta \otimes n$ and this rank-1 line intersects ∂K^c for positive and negative values of t. Therefore every matrix in $\text{int} K^c \cap \{\det = 1\}$ is a rank-1 combination of two matrices in $\partial K^c \cap \{\det = 1\}$ and thus belongs to K^{lc}. This finishes the proof for $\det B = 1$.

Second case: $\det B > 1$.

Since every rank-1 half-line through an interior point of K^{pc} intersects ∂K^{pc} it suffices to show $\partial K^{pc} \subset K^{lc}$. Let $\bar{F} = (\bar{y}, \bar{z}) \in \partial K^{pc}$ and define

$$\begin{aligned} f(y,z) &= (\det B - 1)|y| - \det B + \det F, \\ g(y,z) &= (\det B - 1)|z| + 1 - \det F. \end{aligned}$$

The polyconvex hull is given by $\{f \leq 0\} \cap \{g \leq 0\}$ and thus $f(\bar{y}, \bar{z}) = 0$ or $g(\bar{y}, \bar{z}) = 0$. For convenience we assume the latter, the other case is analogous.

If $f(\bar{F}) = 0$ then $|\bar{y}| + |\bar{z}| = 1$. Moreover f and g are quadratic functions on the segment $t(\bar{y}/|\bar{y}|, 0) + (1-t)(0, \bar{z}/|\bar{z}|)$, and vanish at $t = 0, 1, |\bar{y}|$. Hence they vanish identically on the segment which therefore is a rank-1 segment. Thus $F \in K^{lc}$.

Now suppose that $f(\bar{F}) < 0$. Using the $SO(2)$ invariance we may assume that $z_2 = 0$. For definiteness we suppose $z_1 > 0$, the case $z_1 < 0$ is analogous. Note that the linear space $\{z_2 = 0\}$ agrees with $\{F_{12} + F_{21} = 0\}$. We claim that there exists a rank-1 line in $\{z_2 = 0\}$ through $\bar{F}$ on which g vanishes (as long as $z_1 > 0$). One way to see this is to consider $\tilde{g}(y, z) = (\det B - 1)z_1 + 1 - \det F$ and to note that $\tilde{g} = 0$ defines a one sheeted hyperboloid H in the three dimensional space $\{z_2 = 0\}$. Hence through each point in H there exist two lines that lie on H and thus must be rank-1 lines since $\det F$ is an affine function on these lines. Alternatively one can consider $(y(t), z(t)) = F(t) = \bar{F} + t(\mu - \lambda)\ a \otimes Pa$, with $Pa = (-a_1, a_2)$. Then $z_2(t) = 0$, $\dot{z}_1 = |a|^2 > 0$ and $\tilde{g}$ is affine on the line $t \to F(t)$. A short calculation shows that $\frac{d}{dt} g(F(t))_{|t=0} = (Qa, a)$ and the quadratic form

$$Q = (\det B - 1)\mathrm{Id} + \frac{1}{2}(\mu - \lambda)[(\mathrm{cof}\bar{F})P + P(\mathrm{cof}\bar{F})^T]$$

is indefinite and hence has a nontrivial kernel.

Consider thus the rank-1 line $F(t) = \bar{F} + t(\mu - \lambda) a \otimes Pa$ on which z_2 and g vanish.

Let $t_0 < 0$ be defined by $z_1(t_0) = 0$. Since $g(F(t_0)) = 0$ we deduce that $F(t_0) = (y(0), 0) \in K$. On the other hand $f(F(0)) < 0$ and using the fact that g vanishes on $F(t)$ we have $f(F(t)) = (\det B - 1)(|y(t)| + |z(t)| - 1) \to \infty$ as $t \to \infty$. Hence there exist $t_1 > 0$ such that $f(F(t_1)) = g(F(t_1)) = 0$ and therefore $F(t_1) \in K^{lc}$ by the considerations above. Thus $\bar{F} = F(0) \in K^{lc}$ and the proof is finished. □

4.6 Are all microstructures laminates?

Theorems 4.7 and 4.10 completely classify gradient Young measures $\mathcal{M}^{qc}(K)$ and quasiconvex hulls K^{qc} and thus lead to an abstract solution of problems 2 and 3 in Section 1.4. The catch is that very few quasiconvex functions are known and that the abstract results are therefore of limited use to understand specific sets K. A manageable necessary condition is given by the minors relations (4.7). In this section we discuss the issue of sufficient conditions, i.e. constructions of (homogeneous) gradient Young measures supported on a given set K. The simplest case is $K = \{A, B\}$. If A and B are rank-1 connected every convex combination

$$\nu = \lambda\delta_A + (1-\lambda)\delta_B, \quad \lambda \in [0,1],$$

is a (homogeneous) gradient Young measure. It arises as a limit of a sequence of gradients Du_j arranged in a fine lamellar pattern (see Fig. 13).

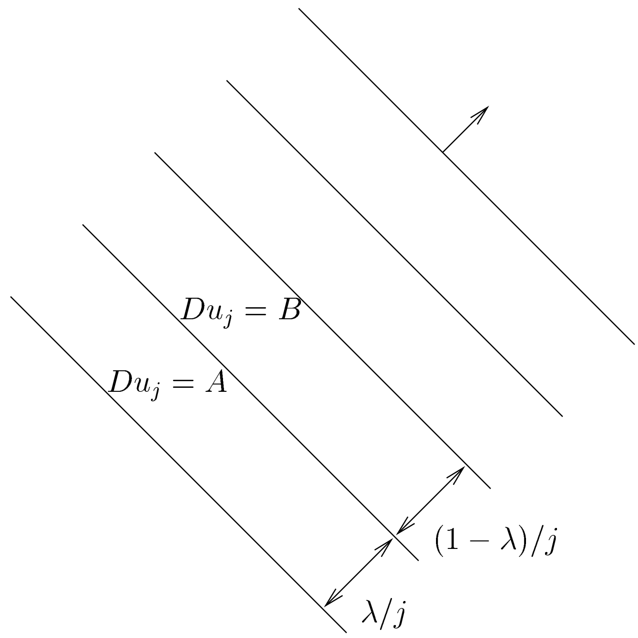

Figure 13: Fine layering of the rank-1 connected matrices A and B generates the homogeneous gradient Young measure $\lambda\delta_A + (1-\lambda)\delta_B$.

We saw in Section 2.5 that this construction can be iterated for larger sets K. More precisely let C be a matrix that is rank-1 connected to $\lambda A + (1-\lambda)B$. Then every convex combination

$$\nu = \mu(\lambda\delta_A + (1-\lambda)\delta_B) + (1-\mu)\delta_C \tag{4.18}$$

is a (homogeneous) gradient Young measure (see Figure 8).

This construction can be iterated and motivates the following definition.

Definition 4.14 *([Da 89]) For a finite family of pairs* $(\lambda_i, F_i) \in (0,1) \times M^{m\times n}$ *the condition* (H_l) *is defined inductively as follows.*

(i) *Two pairs* $(\lambda_1, F_1), (\lambda_2, F_2)$ *satisfy* (H_2) *if*

$$\mathrm{rk}(F_2 - F_1) \le 1, \quad \lambda_1 + \lambda_2 = 1.$$

(ii) *A family* $\{(\lambda_i, F_i)\}_{i=1,\dots,l}$ *satisfies* (H_l) *if, after possible renumbering*

$$\mathrm{rk}(F_l - F_{l-1}) = 1 \tag{4.19}$$

and the new family $\{(\tilde\lambda_i, \tilde F_i)\}_{i=1,\dots,l-1}$ *given by*

$$\tilde F_{l-1} = \frac{\lambda_{l-1}}{\lambda_{l-1}+\lambda_l} F_{l-1} + \frac{\lambda_l}{\lambda_{l-1}+\lambda_l} F_l, \quad \tilde\lambda_{l-1} = \lambda_{l-1} + \lambda_l, \tag{4.20}$$

$$(\tilde\lambda_i, \tilde F_i) = (\lambda_i, F_i) \text{ if } i \le l-2, \tag{4.21}$$

satisfies (H_{l-1}).

If we call the process defined by (4.19), (4.20) and (4.21) contraction then the family $\{(\lambda_i, F_i)\}_{i=1,\dots,l}$ satisfies (H_l) if it can be inductively contracted to $(1, \bar F)$ where $\bar F = \sum \lambda_i F_i$ is the barycentre. Note that the F_i may take the same value for different i. To see that this can be useful consider the 8 matrices $\{A_1, \dots A_4\}$ and $\{I_1, \dots I_4\}$ in the four gradient example in Section 2.5. The family $(1/2, A_1)$, $(1/4, A_2)$, $(1/8, A_3)$, $(1/16, A_4)$, $(1/32, A_1)$, $(1/32, I_1)$ satisfies (H_6), but the family obtained by combining the two pairs involving A_1 to $(17/32, A_1)$ does not satisfy (H_5).

Definition 4.15 *A (probability) measure* ν *on* $M^{m\times n}$ *is called a laminate of finite order if there exists a family* $\{(\lambda_i, F_i)\}_{i=1,\dots,l}$ *that satisfies* (H_l) *and*

$$\nu = \sum_{i=1}^{l} \lambda_i \delta_{F_i}.$$

A (probability) measure ν *is a laminate if there exists a sequence* ν_j *of laminates of finite order with support in a fixed compact set such that*

$$\nu_j \overset{*}{\rightharpoonup} \nu \text{ in } \mathcal{M}(M^{m\times n}).$$

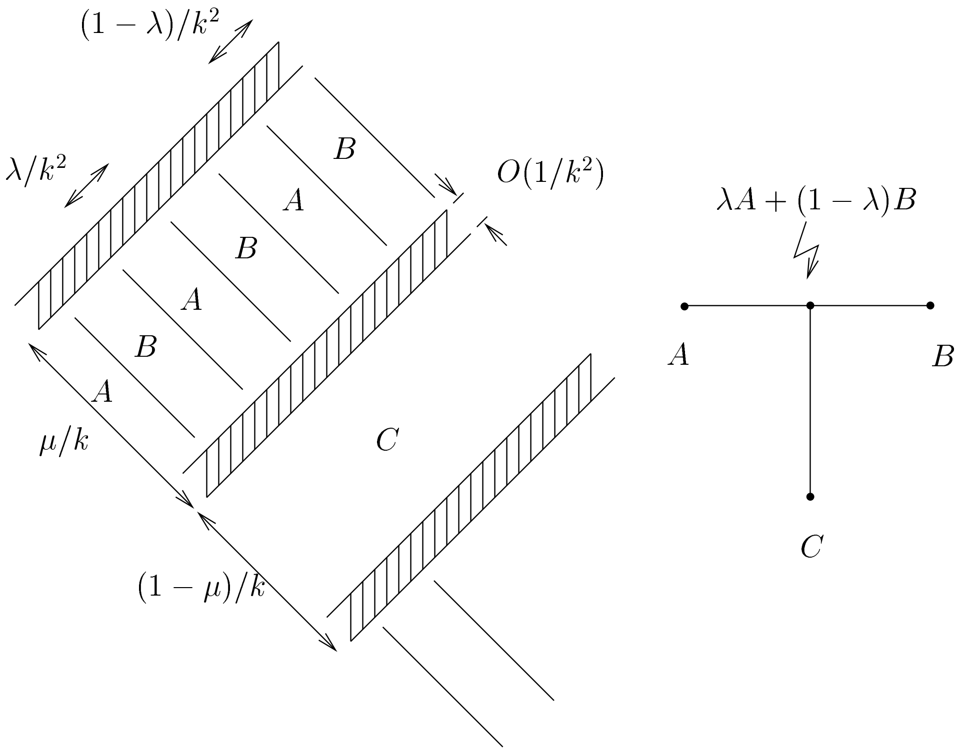

Figure 14: An order-2 laminate that generates (4.18) and the corresponding rank-1 connections.

Example. Again in the context of the four gradient example in Section 2.5 the measures

$$\frac{1}{2}\delta_{A_1} + \frac{1}{4}\delta_{A_2} + \frac{1}{8}\delta_{A_3} + \frac{1}{16}\delta_{A_4} + \frac{1}{16}\delta_{I_2}$$

or

$$(1 - (\frac{1}{16})^j)(\frac{8}{15}\delta_{A_1} + \frac{4}{15}\delta_{A_2} + \frac{2}{15}\delta_{A_3} + \frac{1}{15}\delta_{A_4}) + (\frac{1}{16})^j\delta_{I_2}$$

are laminates of finite order, while $\frac{8}{15}\delta_{A_1} + \frac{4}{15}\delta_{A_2} + \frac{2}{15}\delta_{A_3} + \frac{1}{15}\delta_{A_4}$ is a laminate but not a laminate of finite order.

Condition (H_l) implies that for every rank-1 convex function $f : M^{m\times n} \to \mathbf{R}$ one has

$$\langle \nu, f\rangle \geq f(\langle \nu, \mathrm{id}\rangle)$$

for all laminates of finite order ν. Since (finite) rank-1 convex functions are continuous the same inequality holds for all laminates ν. Pedregal [Pe 93] showed that this property characterizes laminates.

Theorem 4.16 *A compactly supported probability measure $\nu \in \mathcal{M}(M^{m\times n})$ is a laminate if and only if*

$$\langle \nu, f\rangle \geq f(\langle \nu, \mathrm{id}\rangle)$$

for all rank-1 convex functions $f : M^{m\times n} \to \mathbf{R}$. In other words, the laminates supported on a compact set K are given exactly by $\mathcal{M}^{rc}(K)$.

The question raised in the title of this subsection may now be stated more precisely:

Are all gradient Young measures laminates?

In view of Theorem 4.16 this may be concisely stated as

$$\mathcal{M}^{rc} \stackrel{?}{=} \mathcal{M}^{qc}.$$

This would clearly be true if rank-1 convexity implied quasiconvexity. Conversely if $\mathcal{M}^{rc} = \mathcal{M}^{qc}$ then rank-1 convexity would imply quasiconvexity in view of the definition of $\mathcal{M}^{rc}$ and the fact that $f^{qc}(F) = \inf\{\langle \nu, f\rangle : \nu \in M^{qc}, \langle \nu, \mathrm{id}\rangle = F\}$ (one equality follows from the definition of $\mathcal{M}^{qc}$; for the other use Theorem 4.5 (iii) for $\Omega = (0,1)^n$, extend φ periodically, let $\varphi_k(x) = k^{-1}\varphi(kx)$ and note that $\{D\varphi_k\}$ generates a homogeneous gradient Young measure).

In the next section we discuss Šverák's example that shows that rank-1 convexity does not imply quasiconvexity if the target dimension satisfies $m \geq 3$.

4.7 Šverák's counterexample

Theorem 4.17 *(Šverák [Sv 92a]) Suppose that $m \geq 3, n \geq 2$. Then there exists a function $f : M^{m\times n} \to \mathbf{R}$ which is rank-1 convex but not quasiconvex.*

Using this result Kristensen recently showed that there is no local condition that implies quasiconvexity. This finally resolves, for $m \geq 3$, the conjecture carefully expressed by Morrey in his fundamental paper [Mo 52], p. 26: 'In fact, after a great deal of experimentation, the writer is inclined to think that there is no condition of the type discussed, which involves f and only a finite number of its derivatives, and which is both necessary and sufficient for quasi-convexity in the general case.'

To state Kristensen's result let us denote by $\mathcal{F}$ the space of extended real-valued functions $f : M^{m\times n} \to [-\infty, \infty]$. An operator $\mathcal{P} : C^\infty(M^{m\times n}) \to \mathcal{F}$ is called local if the implication

$$f = g \text{ in a neighbourhood of } F \Longrightarrow \mathcal{P}(f) = \mathcal{P}(g) \text{ in a neighbourhood of } F$$

holds.

Theorem 4.18 *([Kr 97a]) Suppose that $m \geq 3, n \geq 2$. There exists no local operator $\mathcal{P} : C^\infty(M^{m\times n}) \to \mathcal{F}$ such that*

$$\mathcal{P}(f) = 0 \Longleftrightarrow f \text{ is quasiconvex.}$$

By contrast, the local operator

$$\mathcal{P}_{rc}(f)(F) = \inf\{D^2 f(F)(a \otimes b, a \otimes b) : a \in \mathbf{R}^m, b \in \mathbf{R}^n\}$$

characterizes rank-1 convexity. At the end of this subsection we will give an argument of Šverák that proves Theorem 4.18 for $m \geq 6$.

Most research before Šverák's result focused on choosing a particular rank-1 convex integrand f (e.g. the Dacorogna-Marcellini example given by (4.2)) and trying to prove or disprove that there exists a function $u \in W_0^{1,\infty}(\Omega; \mathbf{R}^m)$ and $F \in M^{m\times n}$ such that

$$\int_\Omega f(F + Du)dx < \int_\Omega f(F)dx. \tag{4.22}$$

Šverák's key idea was to first fix a function u and to look for integrands f that satisfy (4.22) but are rank-1 convex. He made the crucial observation that the linear space spanned by gradients of trigonometric polynomials contains very few rank-1 direction and hence supports many rank-1 convex functions.

To proceed, it is useful to note that quasiconvexity can be defined using periodic test functions rather than functions that vanish on the boundary.

Proposition 4.19 *A continuous function $f : M^{m\times n} \to \mathbf{R}$ is quasiconvex if and only if*

$$\int_Q f(F + Du)dx \geq f(F)$$

for all Lipschitz functions u that are periodic on the unit cube Q and all $F \in M^{m\times n}$.

Proof. Sufficiency of the condition is clear since it suffices to verify condition (4.1) for Q (see Remark 2 after Definition 4.2). To establish necessity consider a periodic Lipschitz function u and cut-off functions $\varphi_k \in C_0^\infty((-k,k)^n)$ such that $0 \le \varphi_k \le 1$, $\varphi_k = 1$ on $(-(k-1),(k-1))^n$ and $|D\varphi| \le C$. If we let $v_k = \varphi_k u, w_k(x) = \frac{1}{k} v_k(kx)$ then quasiconvexity implies that

$$(k-1)^n \int_Q f(F+Du)dy \ge \int_{(-k,k)^n} f(F+Dv_k)dx - Ck^{n-1}$$

$$= k^n \int_Q f(F+Dw_k)dx - Ck^{n-1} \ge k^n f(F) - Ck^{n-1}.$$

Division by k^n yields the assertion as $k \to \infty$. □

Proof of Theorem 4.17. Consider the periodic function $u : \mathbf{R}^2 \to \mathbf{R}^3$

$$u(x) = \frac{1}{2\pi} \begin{pmatrix} \sin 2\pi x^1 \\ \sin 2\pi x^2 \\ \sin 2\pi(x^1+x^2) \end{pmatrix}.$$

Then

$$Du(x) = \begin{pmatrix} \cos 2\pi x^1 & 0 \\ 0 & \cos 2\pi x^2 \\ \cos 2\pi(x^1+x^2) & \cos 2\pi(x^1+x^2) \end{pmatrix}$$

and

$$L := \operatorname{span}\{Du(x)\}_{x\in\mathbf{R}^2} = \left\{ \begin{pmatrix} r & 0 \\ 0 & s \\ t & t \end{pmatrix} : r,s,t \in \mathbf{R} \right\}.$$

The only rank-1 lines in L are lines parallel to the coordinate axes. In particular the function $g(F) = -rst$ is rank-1 convex (in fact rank-1 affine) on L. On the other hand

$$\int_{(0,1)^2} g(Du(x)) = -\frac{1}{4} < 0 = g(0). \tag{4.23}$$

To prove the theorem it only remains to show that g can be extended to a rank-1 convex function on $M^{3\times 2}$. Whether this is possible is unknown. There is, however, a rank-1 convex function that almost agrees with g in L and this is enough. Let P denote the orthogonal projection onto L and consider the quartic polynomial

$$f_{\epsilon,k}(F) = g(PF) + \epsilon(|F|^2 + |F|^4) + k|F - PF|^2.$$

We claim that for every $\epsilon > 0$ there exists a $k(\epsilon) > 0$ such that $f_{\epsilon,k(\epsilon)}$ is rank-1 convex. Suppose otherwise. Then there exists an $\epsilon > 0$ such that $f_{\epsilon,k}$ is not rank-1 convex for any $k > 0$. Hence there exist $F_k \in M^{m\times n}$, $a_k \in \mathbf{R}^m$, $b_k \in \mathbf{R}^n$, $|a_k| = |b_k| = 1$ such that

$$D^2 f_{\epsilon,k}(F_k)(a_k \otimes b_k, a_k \otimes b_k) \leq 0.$$

Now

$$D^2 f_{\epsilon,k}(F)(X,X) =$$
$$D^2 g(PF)(PX,PX) + 2\epsilon|X|^2 + \epsilon(4|F|^2|X|^2 + 8|F:X|^2) + k|X - PX|^2.$$

The term $D^2g(PF)$ is linear in F while the third term on the right hand side is quadratic and positive definite. Hence F_k is bounded as $k \to \infty$, and passing to a subsequence if needed, we may assume $F_k \to \bar{F}, a_k \to \bar{a}, b_k \to \bar{b}$. Since $D^2 f_{\epsilon,k} \geq D^2 f_{\epsilon,j}$ for $k \geq j$ we deduce

$$D^2 g(P\bar{F})(P\bar{a}\otimes\bar{b}, P\bar{a}\otimes\bar{b}) + 2\epsilon + j|\bar{a}\otimes\bar{b} - P\bar{a}\otimes\bar{b}|^2 \leq 0 \ \forall j. \qquad (4.24)$$

Thus $P(\bar{a}\otimes\bar{b}) = \bar{a}\otimes\bar{b}$, i.e. $\bar{a}\otimes\bar{b} \in L$. Therefore $t \mapsto g(P(\bar{F} + t\bar{a}\otimes\bar{b}))$ is affine, and the first term in (4.24) vanishes. This yields the contradiction $\epsilon \leq 0$.

Thus there exist $k(\epsilon)$ such that $f_\epsilon := f_{\epsilon,k(\epsilon)}$ is rank-1 convex. By (4.23) and the definition of u, the function f_ϵ is not quasiconvex as long as $\epsilon > 0$ is sufficiently small. □

An immediate consequence of Šverák's result and the considerations in the previous section is that there exist gradient Young measures that are not laminates. In fact the measure ν defined by averaging $\delta_{Du(x)}$, i.e.

$$\langle \nu, h \rangle = \int_{(0,1)^2} h(Du(x))dx, \quad \forall h \in C_0(M^{m\times n}),$$

provides an example, since $\langle \nu, f_\epsilon\rangle = \langle \nu, g\rangle + C\epsilon < f_\epsilon(\langle \nu, \mathrm{id}\rangle)$ (for small $\epsilon > 0$).

The following modification, due to James, provides an even simpler example and a nice illustration of the failure of quasiconvexity for g (or more

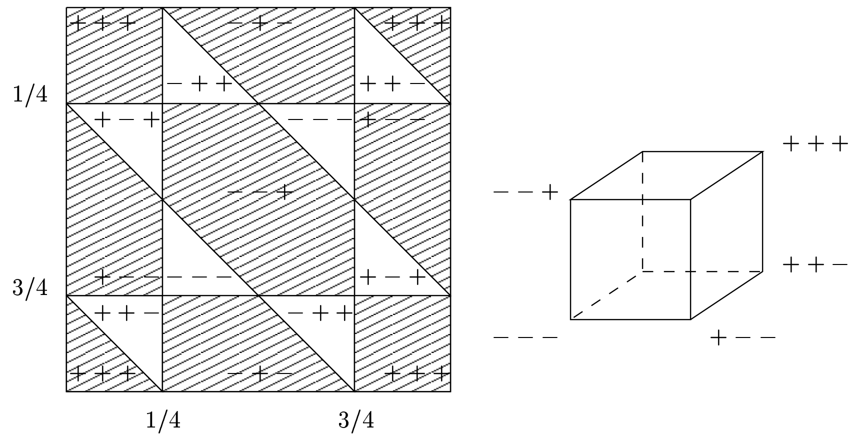

Figure 15: The gradients in James' modification of Šverák's example. Regions of positive parity are shaded. The picture on the right shows the rank-1 connections between the eight gradients.

precisely f_ϵ). Let $s : \mathbf{R} \to \mathbf{R}$ denote the periodic sawtooth function with mean zero and $s' = 1$ on $(0, 1/4) \cup (3/4, 1)$, $s' = -1$ on $(1/4, 3/4)$ and define

$$\tilde{u}(x) = \begin{pmatrix} s(x^1) \\ s(x^2) \\ s(x^1 + x^2) \end{pmatrix}.$$

Then $D\tilde{u} \in L$ and $D\tilde{u}$ takes 8 values which can be denoted by $(+++), (++-), \ldots$ according to the signs of $\partial_1 \tilde{u}^1$, $\partial_2 \tilde{u}^2$ and $\partial_1 \tilde{u}^3 = \partial_2 \tilde{u}^3$. We say that $D\tilde{u}$ has positive parity if the number of $-$ signs is even. The analogue of (4.23) can be proved by inspecting Figure 15.

$$\int_{(0,1)^2} g(D\tilde{u}(x))dx = |\{\text{parity} D\tilde{u} = -\}| - |\{\text{parity} D\tilde{u} = +\}| = -\frac{1}{2}.$$

In particular

$$\nu = \frac{3}{16}(\delta_{+++} + \delta_{+--} + \delta_{-+-} + \delta_{--+}) + \frac{1}{16}(\delta_{++-} + \delta_{+-+} + \delta_{-++} + \delta_{---})$$

is a gradient Young measure that is not a laminate. Indeed for every laminate that only involves the eight matrices $(\pm\pm\pm)$ one has

$$\nu(\text{positive parity}) = \nu(\text{negative parity})$$

since (an arbitrarily small perturbation of) g and $-g$ can be extended to rank-1 convex functions on $M^{3\times 2}$.

Šverák's example leaves open the question whether there exists a (compact) set $K \subset M^{3\times 2}$ that is rank-1 convex (i.e. $K^{rc} = K$) but not quasiconvex. For the gradient set above one has $K^{rc} = K^{qc} = K^c =$ unit cube in L. Using a variant of James' modification Milton [Mi 98] has recently shown that there exists a set $K \in M^{3\times 2}$ of seven matrices that satisfies $K^{rc} \neq K^{qc}$. His motivation arose from the relation between quasiconvexification and optimal composites. His example shows that certain composites cannot be obtained by successive lamination.

Šverák showed that the complex version of the original example yields a set $K \subset M^{6\times 2}$ with $K^{rc} \neq K^{qc}$. With the usual identification $\mathbf{R}^2 \simeq \mathbf{C}$ via $z = x + iy$ we define

$$K = \left\{ \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \\ z_3 & z_3 \end{pmatrix} : z_i \in \mathbf{C}, |z_i| = 1, z_3 = z_1 z_2 \right\}, \tag{4.25}$$

$L = \mathrm{span} K$, P orthogonal projection $M^{6\times 2} \to L$. The periodic function $w : \mathbf{R}^2 \to \mathbf{C}^3$, given by

$$w(x) = \begin{pmatrix} e^{ix^1} \\ e^{ix^2} \\ e^{i(x^1+x^2)} \end{pmatrix}$$

satisfies $Dw \in K$. Hence $0 \in K^{qc}$ (use e.g. that $\mathrm{dist}_K^{qc}(0) = 0$ by Proposition 4.19).

We claim that

$$K^{qc} = K^{pc} = K \cup \{0\}.$$

To prove this consider on L the function

$$g(z_1, z_2, z_3) = |z_1 z_2 - z_3|^2 + |\bar{z}_2 z_3 - z_1|^2 + |z_3 \bar{z}_1 - z_2|^2$$

and note that g vanishes exactly on $K \cup \{0\}$. Now g can easily be extended to a polyconvex function f on $\mathbf{C}^{3\times 2} \simeq M^{6\times 2}$ with $f > 0$ outside L. If

$$F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \\ F_{31} & F_{32} \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$$

we may take

$$\begin{aligned} f(F) &= |\det(F_1, F_2) - F_{31}|^2 + |\det(\bar{F}_2, F_3) + F_{11}|^2 \\ &\quad +|\det(F_3, \bar{F}_1) + F_{22}|^2 + |F - PF|^2. \end{aligned}$$

Therefore $K^{pc} \subset K \cup \{0\}$ and equality holds since $0 \in K^{qc} \subset K^{pc}$. Moreover either $K^{rc} = K$ or $K^{rc} = K \cup \{0\}$. The following result shows that rank-1 convexity is a local condition and hence $K^{rc} = K \neq K^{qc}$.

Lemma 4.20 *Let K_1 and K_2 be disjoint compact sets and suppose that $K = K_1 \cup K_2$ is rank-1 convex. Then both K_1 and K_2 are rank-1 convex.*

Proof. See [Pe 93], Thm. 5.1 or [MP 98], Prop. 2.8.

There is also a simple direct proof that $K^{rc} = K$. Note that $f \in C^\infty$ and $Df^2(0) \geq c\,\mathrm{Id}$, $c > 0$. Indeed $f(0) = Df(0) = 0$ and thus

$$D^2 f(0)(F, F) = \lim_{t \to \infty} \frac{2}{t^2} f(tF) = |F_{31}|^2 + |F_{11}|^2 + |F_{22}|^2 + |F - PF|^2$$

and the right hand side vanishes only if $F = 0$. Hence there exists a neighbourhood $B(0, \epsilon)$ and a new function $\tilde{f}$ such that $\tilde{f} > 0$ in $B(0, \epsilon)$, $D^2 \tilde{f} \geq \frac{c}{2}\mathrm{Id}$ in $B(0, 2\epsilon)$, $\tilde{f} = f$ outside $B(0, \epsilon)$. Then $\tilde{f}$ is locally polyconvex and hence rank-1 convex and $\{\tilde{f} \leq 0\} = K$. Thus $K^{rc} = K$.

Note that $\tilde{f}$ is in particular locally quasiconvex (i.e. for every point there is a neighbourhood in which $\tilde{f}$ agrees with a quasiconvex function) but not quasiconvex since $\tilde{f}(0) > 0$ and $\int_{T^2} \tilde{f}(Dw) = \int_{T^2} f(Dw) = 0$. This proves Kristensen's theorem for $m \geq 6$.

Kružik [Kr 97] used Šverák's counterexample to show that there exists an integrand $f : M^{3\times 2} \to \mathbf{R} \cup \{+\infty\}$ that is not quasiconvex such that $F \mapsto \tilde{f}(F) = f(F^T)$ is quasiconvex. Recall that

$$L = \left\{ \begin{pmatrix} r & 0 \\ 0 & s \\ t & t \end{pmatrix} \ : \ r, s, t \in \mathbf{R} \right\},$$

and let

$$f(F) = \begin{cases} -rst & \text{if } F \in L, \\ +\infty & \text{else.} \end{cases}$$

Then f, and hence $\tilde{F}$, is rank-1 convex since $\{f < \infty\} = L$ is convex and f is rank-1 convex on L. We have already seen that f is not quasiconvex and it only remains to show that

$$\int_{(0,1)^3} \tilde{f}(F + Du)\, dx \geq \tilde{f}(F)$$

for all periodic (Lipschitz) functions $u : \mathbf{R}^3 \to \mathbf{R}^2$. We may assume that $(F + Du)^T \in L$ a.e. Since $\int_{(0,1)^3} Du = 0$ by periodicity we deduce that $F^T \in L$ and $(Du)^T \in L$ a.e. Thus

$$\partial_2 u^1 = \partial_1 u^2 = 0, \quad \partial_3(u^1 - u^2) = 0.$$

Therefore u^1 is independent of x^2, while u^2 is independent of x^1, and differentiation of the second identity yields $\partial_1\partial_3 u^1 = \partial_2\partial_3 u^2 = 0$. Thus

$$\begin{aligned} u^1 &= a(x^1) + b(x^3), \quad u^2 = c(x^2) + d(x^3), \\ Du &= \begin{pmatrix} a'(x^1) & 0 & b'(x^3) \\ 0 & c'(x^2) & d'(x^3) \end{pmatrix}, \end{aligned}$$

and an application of Fubini's theorem in connection with the rank-1 convexity of $\tilde{f}$ yields the desired estimate.

By a more refined argument one can show that the function

$$f_{\varepsilon,k}(F) = f(PF) + \varepsilon(|F|^2 + |F|^4) + k|F - PF|^2$$

considered above provides a finite-valued counterexample if $\varepsilon > 0$ is small enough and $k \geq k(\varepsilon)$. To show that

$$\int_{(0,1)^3} \tilde{f}_{\varepsilon,k}(F + Du) - \tilde{f}_{\varepsilon,k}(F) + D\tilde{f}_{\varepsilon,k}(F)Du\, dx \geq 0,$$

one introduces $v = (v^1, v^2, v^3)$ and $w = (w^1, w^2, w^3)$ by

$$P(D\varphi)^T = \begin{pmatrix} v^1 & 0 \\ 0 & v^2 \\ v^3 & v^3 \end{pmatrix}, \ (D\varphi)^T - P(D\varphi)^T = \begin{pmatrix} 0 & w^2 \\ w^1 & 0 \\ w^3 & -w^3 \end{pmatrix}$$

and observes that the differential operator

$$A(Dv) = (\partial_2 v^1, \partial_3 v^1, \partial_1 v^2, \partial_3 v^2, \partial_1 v^3, \partial_2 v^3)$$

can be expressed as a linear combination of derivatives of w. Hence

$$\|A(Dv)\|_{W^{-1,2}(Q)} \leq C\|(D\varphi)^T - P(D\varphi)^T\|_{L^2(Q)}$$

and the crucial ingredient in the proof are the estimates

$$\begin{aligned} \left|\int_{(0,1)^3} v^1 v^2 v^3 \, dx\right| &\leq C\|v\|_{L^4}^2 \|A(Dv)\|_{W^{-1,2}}, \\ \left|\int_{(0,1)^3} v^i v^j\right| &\leq C\|v\|_{L^2}\|A(Dv)\|_{W^{-1,2}}, \quad \text{for } i \neq j, \end{aligned}$$

which are proved by a suitable decomposition of the (discrete) Fourier transforms $\mathcal{F}v^i$ into a part that is supported in a narrow cone near the i-th coordinates axis and a part that vanishes near that axis. The second part is then easily estimated in terms of $A(Dv)$.

4.8 Proofs: lower semicontinuity and relaxation

Proof of Theorem 4.4(i) ($W^{1,\infty}w^*slsc$ is equivalent to quasiconvexity of the integrand). To establish necessity of quasiconvexity let $Q = (0,1)^n$, $\varphi \in W_0^{1,\infty}(Q, \mathbf{R}^m)$, extend φ 1-periodically to $\mathbf{R}^n$ and let

$$u_j(x) = Fx + \frac{1}{j}\varphi(jx), \text{ for } x \in \Omega.$$

Then $u_j \overset{\star}{\rightharpoonup} u$ in $W^{1,\infty}(\Omega; \mathbf{R}^m)$, $u = Fx$ and

$$f(Du_j) \overset{\star}{\rightharpoonup} \text{const} = \int_Q f(F + D\varphi(y))dy \quad \text{in } L^\infty(\Omega),$$

cf. Section 3.2 a). The necessity of quasiconvexity follows.

To prove sufficiency consider $u_j \overset{\star}{\rightharpoonup} u$ in $W^{1,\infty}(\Omega; \mathbf{R}^m)$ and suppose first that $u(x) = Fx$. If $u_j - u$ was zero on $\partial\Omega$ the assertion would follow from the definition of quasiconvexity. For general u_j consider a compactly contained subdomain $\Omega' \subset\subset \Omega$, a cut-off function $\eta \in C_0^\infty(\Omega)$ with $\eta = 1$ on Ω' and let

$$v_j = u + \eta(u_j - u).$$

Since $u_j \to u$ locally uniformly in Ω by the Sobolev embedding theorem (or by the Arzela-Ascoli theorem) and since $|Du_j| \leq C$ we may assume that

$|Dv_j| \leq C'$ for $j \geq j_0(\eta)$. If we let $M = \sup\{|f(F)| : |F| \leq C + C'\}$ and use quasiconvexity we obtain

$$\begin{aligned}\liminf_{j\to\infty} I(u_j) &\geq \liminf_{j\to\infty} \left[\int_{\Omega} f(Dv_j)\, dx + \int_{\Omega\setminus\Omega'} (f(Du_j) - f(Dv_j))\, dx \right] \\ &\geq |\Omega| f(F) - 2M|\Omega \setminus \Omega'|.\end{aligned}$$

Since $\Omega' \subset\subset \Omega$ was arbitrary the assertion follows for $u = Fx$ and similarly for piecewise affine u.

For arbitrary $u \in W^{1,\infty}(\Omega; \mathbf{R}^m)$ the result is established by approximation as follows. For compactly contained subdomains $\Omega' \subset\subset \Omega'' \subset\subset \Omega$ there exist v_k such that v_k is piecewise affine in Ω', $u = v_k$ in $\Omega \setminus \Omega''$, $|Dv_k| \leq C$, $Dv_k \to Du$ in measure (and hence in all $L^p, p < \infty$). To construct such v_k first approximate u in Ω'' by a C^1 function and then consider piecewise linear approximations on a sufficiently fine (regular) triangulation. Let $u_{j,k} = u_j + v_k - u$. Then

$$u_{j,k} \overset{*}{\rightharpoonup} v_k \quad \text{in } W^{1,\infty}(\Omega; \mathbf{R}^m) \text{ as } j \to \infty, \tag{4.26}$$

$$|Du_{j,k}| \leq C \tag{4.27}$$

Hence, by the previous result and the dominated convergence theorem

$$\begin{aligned}\lim_{k\to\infty} \liminf_{j\to\infty} \int_{\Omega'} f(Du_{j,k})dx &\geq \lim_{k\to\infty} \int_{\Omega'} f(Dv_k)\, dx \\ = \int_{\Omega'} f(Du)dx &\geq \int_{\Omega} f(Du) - C|\Omega \setminus \Omega'|\,.\end{aligned}$$

On the other hand by (4.27), the uniform continuity of f on compact sets and the convergence of Dv_k in measure

$$\lim_{k\to\infty} \sup_j \int_{\Omega'} |f(Du_{j,k}) - f(Du_j)|\, dx = 0.$$

Hence

$$\liminf_{j\to\infty} \int_{\Omega} f(Du_j)\, dx \geq \int_{\Omega} f(Du)\, dx - 2C|\Omega \setminus \Omega'|,$$

and the assertion follows since Ω' was arbitrary. $\square$

Proof of Theorem 4.5(iii) (formula for f^{qc}).
Let

$$Qf(F,U) := \inf_{\varphi \in W_0^{1,\infty}} \frac{1}{|U|} \int_U f(F + D\varphi)\, dx.$$

We have to show that $f^{qc}(F) = Qf(F,U)$. A simple scaling and covering argument shows that Qf is independent of U. By the definition of quasiconvexity $Qf \geq Qf^{qc} = f^{qc}$. To prove the converse inequality $Qf \leq f^{qc}$ it suffices to show that Qf is quasiconvex since $Qf \leq f$. We first claim that

$$\begin{gathered} \frac{1}{|U|} \int_U Qf(F + D\psi)\, dx \geq Qf(F), \\ \forall \psi \in W_0^{1,\infty}(\Omega; \mathbf{R}^m),\, \psi \text{ piecewise affine.} \end{gathered} \tag{4.28}$$

Let U be a finite union (up to a null set) of disjoint open subsets U_i such that ψ is affine on U_i and let $\epsilon > 0$. By the definition of Qf (applied to U_i) there exist $\varphi_i \in W_0^{1,\infty}(U_i; \mathbf{R}^m)$ such that

$$Qf(F + D\psi) \geq \frac{1}{|U_i|} \int_{U_i} f(F + D\psi + D\varphi_i)\, dx - \epsilon \quad \text{ on } U_i.$$

Set $\varphi = \psi + \sum \varphi_i \in W_0^{1,\infty}(U; \mathbf{R}^m)$. Rearranging terms we find

$$\begin{aligned} \int_U Qf(F + D\psi)\, dx &\geq \int_U f(F + D\varphi)\, dx - \epsilon |U| \\ &\geq Qf(F) - \epsilon |U|, \end{aligned}$$

and assertion (4.28) follows as $\epsilon > 0$ was arbitrary. Now (4.28) is enough to conclude that Qf is rank-1 convex and therefore locally Lipschitz continous (see Remark 3 after Definition 4.2). Hence Qf is quasiconvex by (4.28) and density arguments and therefore $f^{qc} = Qf$.

So far we have assumed that Qf does not take the value $-\infty$. If $Qf(F + D\psi) = -\infty$ on U_i then an obvious modification of the argument above shows that (4.28) still holds. Hence Qf is rank-1 convex (see the proof of Lemma 4.3) and one easily concludes that $f^{qc} \equiv Qf \equiv -\infty$ since the rank-1 directions span the space of all matrices. $\square$

4.9 Proofs: classification

The main point is to show that Jensen's inequality for quasiconvex functions characterizes homogeneous gradient Young measures (see Lemma 4.23). The proof relies on the Hahn-Banach separation theorems and the representation (4.5) for f^{qc}. The extension to nonhomogeneous Young measures uses mainly generalities about measurable maps, in particular their approximation by piecewise constant ones.

An important technical tool of independent interest is a truncation result for sequences of gradients sometimes known as Zhang's lemma. (Closely related results were obtained previously by Acerbi and Fusco based on earlier work of Liu.) It implies that every gradient Young measure supported on a compact set $K \subset M^{m\times n}$ can be generated by a sequence $\{Dv_j\}$ whose L^∞ norm can be bounded in terms of K alone. For the rest of this section we adopt the following conventions:

$$K \quad \text{is a compact set in } M^{m\times n},$$

$$U, \Omega \quad \text{are bounded domains in } \mathbf{R}^n, |\partial\Omega| = |\partial U| = 0.$$

Lemma 4.21 *(Zhang's lemma). Let* $|K|_\infty = \sup\{|F| : F \in K\}$.

(i) *Let* $u_j \in W^{1,1}_{\mathrm{loc}}(\mathbf{R}^n; \mathbf{R}^m)$ *and suppose that*

$$\mathrm{dist}(Du_j, K) \to 0 \text{ in } L^1(\mathbf{R}^n). \tag{4.29}$$

Then there exists a sequence $v_j \in W^{1,1}_{\mathrm{loc}}(\mathbf{R}^n; \mathbf{R}^m)$ *such that*

$$|Dv_j| \le c(n,m)|K|_\infty, \tag{4.30}$$

$$|\{u_j \neq v_j\}| \to 0. \tag{4.31}$$

(ii) *Let* $U \in \mathbf{R}^n$ *be a bounded domain and let* $u_j \in W^{1,1}_{\mathrm{loc}}(U; \mathbf{R}^m)$. *Suppose that*

$$\mathrm{dist}(Du_j, K) \to 0 \text{ in } L^1(U), \quad u_j \to u \text{ in } L^1_{\mathrm{loc}}(U). \tag{4.32}$$

Then there exist $v_j \in W^{1,1}_{\mathrm{loc}}(U; \mathbf{R}^m)$ *such that*

$$|Dv_j| \le c(n,m)|K|_\infty, \tag{4.33}$$

$$|\{u_j \neq v_j\}| \to 0, \quad v_j = u \text{ near } \partial U.$$

Remarks. Estimates (4.29) and (4.32) can be replaced by the stronger assertion $\text{dist}(Dv_j, K^c) \to 0$ in L^∞, see [Mu 97b]. Note also that the assertion $|\{u_j \neq v_j\}| \to 0$ implies

$$|\{Du_j \neq Dv_j\}| \to 0,$$

since for any Sobolev function $Du = 0$ a.e. on $\{u = 0\}$.

Proof. Part (i) is essentially Lemma 3.1 in [Zh 92]. Alternatively it follows from (the proof of) Theorem 6.6.3 in [EG 92], pp. 254-255, with $\lambda = 3C|K|_\infty$. Part (ii) follows by a standard localization argument, see [Mu 97b] for the details. □

Now suppose that $\{u_j\}$ is bounded in $W^{1,\infty}(\Omega; \mathbf{R}^m)$ and $\{Du_j\}$ generates the (gradient) Young measure ν. Then $Du_j \overset{*}{\rightharpoonup} Du$ in L^∞, $Du(x) = \langle \nu_x, \text{id} \rangle$ and $u_j \to u$ (locally) in L^∞. We call u the underlying deformation of ν. The Young measure $\nu : \Omega \to \mathcal{M}(M^{m\times n})$ is called homogeneous if it is constant in Ω (up to a null set). As usual we identify constant maps with their values and view the set $H(\Omega)$ of homogeneous gradient Young measures as a subset of $\mathcal{M}(M^{m\times n})$. By $H(\Omega, K)$ we denote the set of homogeneous gradient Young measures supported on K.

Lemma 4.22 *We have*

(i) *If* $\nu \in H(\Omega, K)$ *and* $\langle \nu, \text{id} \rangle = 0$ *then there exists a sequence* $u_j \in W_0^{1,\infty}(\Omega; \mathbf{R}^m)$ *such that* Du_j *generates* ν *and satisfies* $|Du_j| \le C|K|_\infty$.

(ii) $H(\Omega, K)$ *is weak* compact in* $\mathcal{M}(M^{m\times n})$.

(iii) *The set* $H(\Omega, K)$ *is independent of* Ω. *If* ν *is a gradient Young measure with* $\text{supp}\nu(x) \subset K$ a.e. *whose underlying deformation* u *agrees with an affine map on* $\partial\Omega$ *(in the sence of* $W_0^{1,\infty}$*) then the average* $\text{Av}\nu$ *defined by*

$$\langle \text{Av}\nu, f \rangle = \frac{1}{|\Omega|} \int_\Omega \langle \nu_x, f \rangle dx$$

belongs to $H(K)$.

(iv) *The set* $H_F(K) = \{\nu \in H(K) : \langle \nu, \text{id} \rangle = F\}$ *is weak* closed and convex.*

Proof. Assertion (i) follows from the definition of $H(\Omega, K)$ and part (ii) of Zhang's lemma. The proof of (ii) uses (i) and a diagonalization argument. Note that $H(\Omega, K)$ is contained in the weak$*$ compact set $\mathcal{P}(K)$ of probability measures on K. Hence the weak$*$ topology is metrizable on $\mathcal{P}(K)$ and can be described by sequences. Suppose that $\nu_k \in H(\Omega, K)$ and $\nu_k \stackrel{*}{\rightharpoonup} \nu$. After subtraction of affine functions in the generating sequences for ν_k we may assume that $\langle \nu_k, \mathrm{id} \rangle = 0$. By (i) there exist $u_{k,j} \in W_0^{1,\infty}(\Omega; M^{m\times n})$ such that

$$\delta_{Du_{k,j}(\cdot)} \underset{j\to\infty}{\stackrel{*}{\rightharpoonup}} \nu_k \quad \text{in } L^\infty_w(\Omega; \mathcal{M}(M^{m\times n})), \quad |Du_{k,j}| \le C|K|_\infty.$$

Here we identified ν_k with the constant map $x \mapsto \nu_k$. Since the weak$*$ topology is metrizable on $L^\infty_w(\Omega; P(B(0, C|K|_\infty)))$ we can apply a standard diagonalization argument to find $j_k \to \infty$ such that

$$\delta_{Du_{k,j_k}(\cdot)} \stackrel{*}{\rightharpoonup} \nu \text{ in } L^\infty_w(\Omega; \mathcal{M}(M^{m\times n})).$$

Since $|Du_{k,j_k}| \le C$ we have $\nu \in H(\Omega, K)$. Thus $H(\Omega, K)$ is weak$*$ closed and therefore weak$*$ compact as a subset of $\mathcal{P}(K)$.

To prove (iii) consider first $v \in W_0^{1,\infty}(U; \mathbf{R}^m)$ and the trivial Young measure given by $\mu(x) = \delta_{Dv(x)}$. We claim that $\mathrm{Av}\mu$ is a homogeneous gradient Young measure (for all domains Ω). By the Vitali covering theorem there exist disjoint scaled copies $U_i = a_i + r_i U$ of U that are contained in the unit cube Q and cover it up to a null set. Define

$$w(x) = \begin{cases} r_i v\left(\frac{x-a_i}{r_i}\right) & \text{in } U_i, \\ 0 & \text{in } Q \setminus U_i, \end{cases}$$

extend w 1-periodically to $\mathbf{R}^n$, and let $w_k(x) = k^{-1}w(kx)$. Then for all continuous functions g (see Section 3.2 a))

$$g(Dw_k) \stackrel{*}{\rightharpoonup} \bar{g} \quad \text{in } L^\infty(\mathbf{R}^n),$$

where

$$\bar{g} = \int_Q g(Dw)dx = \frac{1}{|U|}\int_U g(Dv)dx = \langle \mathrm{Av}\mu, g\rangle.$$

Thus, for all Ω, $\delta_{Dw_k(\cdot)}$ converges to the homogeneous Young measure $Av\mu$ in $L^\infty_w(\Omega; \mathcal{M}(M^{m\times n}))$ in the weak$*$ topology. Hence $\mathrm{Av}\mu \in H(\Omega)$ as claimed.

Now let ν satisfy the assumption of (iii). We may suppose that $u \in W_0^{1,\infty}(\Omega; \mathbf{R}^m)$. By the definition of gradient Young measures and part (ii) of

Zhang's lemma there exists a sequence $u_j \in W_0^{1,\infty}(\Omega; \mathbf{R}^m)$ such that $|Du_j| \leq R$ and

$$\delta_{Du_j(\cdot)} \overset{*}{\rightharpoonup} \nu \quad \text{in } L^\infty_w(\Omega; \mathcal{M}(M^{m\times n})).$$

Taking test functions of the form $1 \otimes g$ we see that

$$\mathrm{Av}\delta_{Du_j(\cdot)} \overset{*}{\rightharpoonup} \mathrm{Av}\nu.$$

By the considerations above $\mathrm{Av}\delta_{Du_j(\cdot)}$ belongs to $H(\Omega, \overline{B(0,R)})$ for all Ω. By (ii) the same holds for $\mathrm{Av}\nu$. Since $\mathrm{supp}\nu(x) \subset K$ a.e. in fact $\mathrm{Av}\,\nu \in H(\Omega, K)$. If $\nu \in H(U, K)$ then $\mathrm{Av}\nu = \nu$ and hence $H(\Omega, K)$ is independent of Ω.

Regarding (iv) we may suppose $F = 0$. Let $\nu_1, \nu_2 \in H_0(K)$. Let $Q_1 = (0, \lambda) \times (0,1)^{n-1}, Q_2 = (\lambda, 1) \times (0,1)^{n-1}$. By (i) there exist sequences $\{Du_{i,j}\} \subset W_0^{1,\infty}(Q_i, \mathbf{R}^m), i = 1, 2$ that generate ν_i. Hence the gradients of

$$u_j(x) = \begin{cases} u_{1,j}(x), & x \in Q_1 \\ u_{2,j}(x), & x \in Q_2 \end{cases}$$

generate

$$\nu(x) = \begin{cases} \nu_1, & x \in Q_1 \\ \nu_2, & x \in Q_2 \end{cases}.$$

By (iii) we have

$$\lambda\nu_1 + (1-\lambda)\nu_2 = \mathrm{Av}\nu \in H_0(K). \qquad \square$$

Lemma 4.23 *(characterization of homogeneous gradient Young measures). We have*

$$H(K) = \mathcal{M}^{qc}(K).$$

Proof. Clearly $H(K) \subset \mathcal{M}^{qc}(K)$ by lower semicontinuity (see Theorem 4.4(i)). To prove the converse it suffices to consider measures with barycentre zero. Now $H_0(K)$ is weak$*$ closed and convex, and $C(K)$ is the dual of $(\mathcal{M}(K)$, weak$*)$ (see e.g. [Ru 73], Thm. 3.10). By the Hahn-Banach separation theorem it suffices to show that, for all $f \in C(K)$,

$$\langle \nu, f \rangle \geq \alpha \quad \forall \nu \in H_0(K), \tag{4.34}$$

implies

$$\langle \mu, f \rangle \geq \alpha \quad \forall \mu \in \mathcal{M}_0^{qc}(K).$$

Fix $f \in C(K)$, consider a continuous extension to $C_0(M^{m\times n})$ and let

$$f_k(F) = f(F) + k\mathrm{dist}^2(F, K).$$

We claim that

$$\lim_{k\to\infty} f_k^{qc}(0) \geq \alpha. \tag{4.35}$$

Once this is shown we are done since by definition every $\mu \in \mathcal{M}_0^{qc}(K)$ satisfies

$$\langle \mu, f\rangle = \langle \mu, f_k\rangle \geq \langle \mu, f_k^{qc}\rangle \geq f_k^{qc}(0).$$

Suppose now (4.35) was false. Then there exist $\delta > 0$ such that

$$f_k^{qc}(0) \leq \alpha - 2\delta, \quad \forall\, k.$$

By Theorem 4.5(iii) there exist $u_k \in W_0^{1,\infty}(Q; \mathbf{R}^m)$ such that

$$\int_Q f_k(Du_k)dy \leq \alpha - \delta. \tag{4.36}$$

In particular we may assume $u_k \rightharpoonup u$ in $W_0^{1,2}(Q; \mathbf{R}^m)$ and

$$\mathrm{dist}(Du_k, K) \to 0 \text{ in } L^1(Q).$$

By part (ii) of Zhang's lemma there exists $v_k \in W_0^{1,\infty}(Q; \mathbf{R}^m)$ such that

$$|Dv_k| \leq C, \quad |\{(Du_k \neq Dv_k\}| \to 0. \tag{4.37}$$

In particular a further subsequence of $\{Dv_k\}$ generates a gradient Young measure ν with $\mathrm{supp}\nu(x) \subset K$ and underlying deformation $u \in W_0^{1,\infty}(Q; \mathbf{R}^m)$. Thus $\mathrm{Av}\nu \in H_0(K)$ by Lemma 4.22 (iii). Since f is bounded from below we deduce from (4.37) and (4.34)

$$\liminf_{k\to\infty} \int_Q f_j(Du_k) \geq \liminf_{k\to\infty} \int_Q f_j(Dv_k)$$

$$= \int_Q \langle \nu_x, f_j\rangle dx = \langle \mathrm{Av}\nu, f_j\rangle \geq \alpha.$$

This contradicts (4.36) as $f_k \geq f_j$ if $k \geq j$, and (4.35) is proved. □

Proof of Theorem 4.7. Necessity of conditions (i) - (iii) was established in Section 4.3. To prove sufficiency we first consider the case that the underlying deformation vanishes. Let

$$A = \{\nu \in L^\infty_w(\Omega; \mathcal{M}(M^{m\times n})) : \nu(x) \in \mathcal{M}^{qc}_0(K) \text{ a.e.}\}$$

denote the set of maps that satisfy (i) - (iii) with $Du = 0$. We have to show that every element of A is a gradient Young measure.

To do so we use some generalities about measurable maps to approximate the elements of A by piecewise constant maps. First note that the set of subprobability measures $\mathcal{M}_1 = \{\mu \in \mathcal{M}(M^{m\times n}) : \nu \geq 0, \|\nu\| \leq 1\}$ is weak$*$ compact in $\mathcal{M}(M^{m\times n})$. Hence the weak$*$ topology is metrizable on $\mathcal{M}_1$. To define a specific metric let $\{f_i\} \subset C_0(M^{m\times n})$ be a countable dense set in the unit sphere of $C_0(M^{m\times n})$ and let

$$d(\mu, \mu') = \sum_{i=1}^\infty 2^{-i} |\langle \mu - \mu', f_i \rangle|.$$

The space $(\mathcal{M}_1, d)$ is a compact metric space. Since d induces the weak$*$ topology, a map $\nu : \Omega \to \mathcal{M}(M^{m\times n})$ that takes (a.e.) values in $\mathcal{M}_1$ is weak$*$ measurable if and only if $\nu : \Omega \to (\mathcal{M}_1, d)$ is measurable.

The set $\{\nu \in L^\infty_w(\Omega; \mathcal{M}(M^{m\times n})) : \nu(x) \in \mathcal{M}_1 \text{ a.e.}\}$ is also weak$*$ compact in $L^\infty_w(\Omega; \mathcal{M}(M^{m\times n}))$ (cf. the proof of Theorem 3.2). A metric $\tilde d$ that induces weak$*$ convergence on that set may be defined as follows. Let $\{h_j\}$ be a countable dense set in the unit ball of $L^1(\Omega)$ and let

$$\tilde d(\nu, \nu') = \sum_{i,j=1}^\infty 2^{-i-j} |\langle \nu - \nu', h_j \otimes f_i \rangle|.$$

It thus follows from Proposition 4.24 below that every $\nu \in A$ can be arbitrarily well approximated in $\tilde d$ by a map $\tilde\nu$ with the following properties. There exist finitely many disjoint open sets U_i with $|\partial U_i| = 0$ such that $\tilde\nu = \nu_i$ on $U_i, \nu_i \in \mathcal{M}^{qc}_0(K)$, $\tilde\nu = \delta_0$ on $\Omega \setminus \cup U_i$. Application of Lemma 4.22(i) to each U_i shows that $\tilde\nu$ is a gradient Young measure (extend the generating sequence by zero to $\Omega \setminus \cup U_i$). Hence the closure of gradient Young measures with support in $K' = K \cup \{0\}$ contains A. On the other hand the set of these Young measures is (weakly) compact (see the proof of Lemma 4.22 (ii)). Thus every $\nu \in A$ is a gradient Young measure. This finishes the proof if $Du = 0$.

The remaining case $Du \neq 0$ can now be easily treated by translation. For a measure μ define its push-forward unter translation by

$$\langle T_F \mu, f \rangle = \langle \mu, f(\cdot + F) \rangle$$

so that $T_F \delta_0 = \delta_F$. Now if ν satisfies the hypotheses of Theorem 4.7 and $\tilde{\nu}(x) = T_{-Du(x)}\nu(x)$ then $\tilde{\nu} \in A$. Hence there exists a sequence $\{Dv_j\}$ that generates $\tilde{\nu}$ and one easily verifies that $Du_j = Dv_j + Du$ generates ν (use e.g. Corollary 3.3 with $f(x, F) = g(Du(x) + F), g \in C_0(M^{m\times n})$). □

Proposition 4.24 *Let (X, d) be a compact metric space and $M \subset X$. Suppose that $\nu : \Omega \to X$ is measurable and $\nu(x) \in M$a.e. Then, for every $k \in \mathbf{N}$, there exists a finite number of disjoint open sets U_i with $|\partial U_i| = 0$ and values $\nu_i \in M$ such that the map*

$$\tilde{\nu} = \begin{cases} \nu_i & \text{on } U_i \\ \nu_0 & \text{on } \Omega \setminus U_i \end{cases}$$

satisfies

$$|\{x : d(\nu(x), \tilde{\nu}(x)) > \frac{1}{k}\}| < \frac{1}{k}.$$

Proof. By compactness X can be covered by a finite number of open balls B_i with radius $\frac{1}{2k}$. The sets $\tilde{E}_i = \nu^{-1}(B_i)$ are measurable. To obtain disjoint sets E_i, we define $E_1 = \tilde{E}_1$, $E_2 = \tilde{E}_2 \setminus E_1$, etc. If $|E_i| > 0$ then there exist $x_i \in E_i$ such that $\nu_i := \nu(x_i) \in M$. There exist disjoint compact sets $K_i \subset E_i$ such that

$$\sum |E_i \setminus K_i| < 1/k; \tag{4.38}$$

if $|E_i| = 0$ we take $K_i = \emptyset$. The K_i have positive distance and thus there exist disjoint open sets $U_i \supset K_i$ with $|\partial U_i| = 0$ (consider e.g. suitable sublevel sets of the distance function of K_i). Now $E_i \supset \tilde{E}_i \supset K_i$ and thus $d(\nu(x), \nu_i) < 1/k$ in K_i. The assertion follows from (4.38). □

5 Exact solutions

Approximate solutions are characterized by the quasiconvex hull K^{qc} and set $\mathcal{M}^{qc}(K)$ of Young measures. The construction of exact solutions is more delicate. In view of the negative result for the two-gradient problem (see Proposition 2.1) it was widely believed that exact solutions are rather rare. Recent results suggest that many exact solutions exist but that they have to be very complicated. This is reminiscent of rigidity and flexibility results for isometric immersions and other geometric problems (see [Na 54]; [Ku 55]; [Gr 86], Section 2.4.12).

To illustrate some of the difficulties consider again the two-dimensional two-well problem (see Section 4.5)

$$Du \in K \;\text{ a.e. in } \Omega, \quad u = Fx \text{ on } \partial\Omega, \tag{5.1}$$

$$K = SO(2)A \cup SO(2)B, \tag{5.2}$$

$$A = \text{ Id }, \quad B = \operatorname{diag}(\lambda, \mu), \quad 0 < \lambda < 1 < \mu, \lambda\mu \geq 1. \tag{5.3}$$

If we ignore boundary conditions the simplest solutions of $Du \in K$ are simple laminates, see Figure 16. A short analysis of the rank-1 connections in K shows that such laminates are perpendicular to one of the normals n_1 or n_2, determined by the two solutions of the equation

$$QA - B = a \otimes n. \tag{5.4}$$

There is, however, no obvious way to combine the two laminates (see Fig. 17). It was thus believed that the problem (5.1) – (5.3) has no nontrivial solutions. This is false. The construction of nontrivial solutions is based on Gromov's method of convex integration.

5.1 Existence of solutions

First, one observes that the open version of the two-gradient problem admits a solution. Here and for the rest of this section we say that a map $u : \Omega \to \mathbf{R}^m$ is piecewise linear if it is Lipschitz continuous and if there exist finitely or

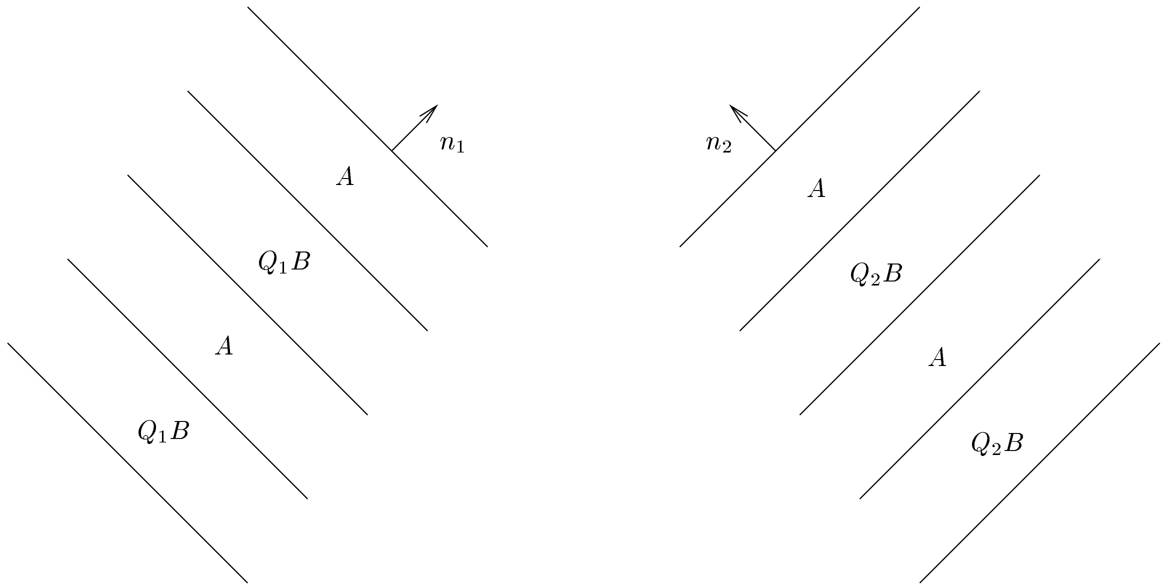

Figure 16: Two possible laminates for the two-well problem.

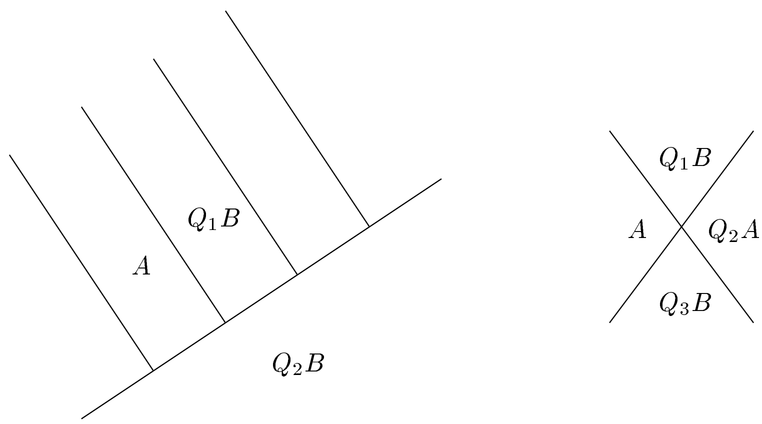

Figure 17: None of the above constructions satisfies the rank-1 condition across every interface.

countably many disjoint open sets Ω_i whose union has full measure in Ω such that $u_{|\Omega_i}$ is affine.

Lemma 5.1 *([MS 96]). Suppose that* $\mathrm{rk}(B-A) = 1, F = \lambda A+(1-\lambda)B, \lambda \in (0,1)$. *Then, for a bounded domain* Ω *and every* $\delta > 0$ *there exists a piecewise linear map* u *such that*

$$u(x) = Fx \quad \text{on } \partial\Omega$$

$$\mathrm{dist}(Du, \{A, B\}) < \delta,$$

$$\sup |u(x) - Fx| < \delta.$$

Remark. It is even possible to handle certain constraints. If $n = m = 2$ and $\det A = \det B = c \neq 0$ then one can achieve $\det Du = c$. How many constraints can be handled is a largely open problem.

Proof. The construction has some similarities with Fonseca's work, in particular her proof of Theorem 2.4 of [Fo 88]. There are some differences, however, so I give the proof in [MS 96] which is slightly simpler. We will first construct a solution for a special domain U. The argument will then be finished by an application of the Vitali covering theorem.

By an affine change of variables we may assume without loss of generality that

$$A = -\lambda a \otimes e_n, \quad B = (1-\lambda) a \otimes e_n, \quad F = 0, \quad \text{and} \quad |a| = 1.$$

Let $\epsilon > 0$, let

$$V = (-1,1)^{n-1} \times ((\lambda - 1)\epsilon, \lambda\epsilon)$$

and define $v : V \to \mathbf{R}^m$ by

$$v(x) = -\epsilon\lambda(1-\lambda)a + \begin{cases} -\lambda a x_n & \text{if } x_n < 0, \\ (1-\lambda) a x_n & \text{if } x_n \geq 0. \end{cases}$$

Then $Dv \in \{A, B\}$ and $v = 0$ at $x_n = \epsilon(\lambda - 1)$ and $x_n = \epsilon\lambda$, but v does not vanish on the whole boundary ∂V. Next let

$$h(x) = \epsilon\lambda(1-\lambda)a \sum_{i=1}^{n-1} |x_i|.$$

Then h is piecewise linear and $|Dh| = \epsilon\lambda(1-\lambda)\sqrt{n-1}$. Set

$$\tilde{u} = v + h.$$

Note that $\tilde{u} \geq 0$ on ∂V and let

$$U = \{x \in V : \tilde{u}(x) < 0\}.$$

Then

$$\begin{array}{rcl} \tilde{u}_{|U} \quad \text{is piecewise linear} & , & \tilde{u}_{|\partial U} = 0, \\ \operatorname{dist}(D\tilde{u}, \{A, B\}) & \leq & \epsilon\lambda(1-\lambda)\sqrt{n-1}, \\ |\tilde{u}| & \leq & \epsilon\lambda(1-\lambda). \end{array}$$

By the Vitali covering theorem one can exhaust Ω by disjoint scaled copies of U. More precisely there exist $x_i \in \mathbf{R}^n$ and $r_i > 0$ such that the sets

$$U_i = x_i + r_i U$$

are mutually disjoint and $|\Omega \setminus \cup_i U_i| = 0$. Define u by

$$u(x) = \begin{cases} r_i \tilde{u}(\frac{x-x_i}{r_i}) & \text{if} \quad x \in U_i, \\ 0 & \text{else.} \end{cases}$$

Note that

$$Du(x) = D\tilde{u}(\frac{x - x_i}{r_i}) \quad \text{if} x \in \Omega_i.$$

It follows that u is piecewise linear, that $u_{|\partial\Omega} = 0$ and that $\operatorname{dist}(Du, \{A, B\}) < \delta$ for a suitable $\epsilon > 0$. Moreover by choosing $r_i \leq 1$ one can also obtain the estimate for $|u - Fx|$. □

Lemma 5.1 can be easily iterated, and using the notion of the lamination convex hull of a set (see Section 4.4) one obtains the following result (see [MS 96] for the details).

Lemma 5.2 *Suppose that $U \subset M^{m\times n}$ is open. Let $v : \Omega \to \mathbf{R}^m$ be piecewise affine and Lipschitz continuous and suppose $Dv \in U^{lc}$ a.e. Then there exist $u : \Omega \to \mathbf{R}^m$ such that*

$$Du \in U \text{ a.e. in } \Omega, \quad u = v \text{ on } \partial\Omega.$$

The crucial step is the passage from open to compact sets $K \subset M^{m\times n}$. Following Gromov we say that a sequence of sets U_i is an in-approximation of K if

(i) the U_i are open and contained in a fixed ball

(ii) $U_i \subset U_{i+1}^{lc}$

(iii) $U_i \to K$ in the following sense: if $F_{i_k} \in U_{i_k}$, $i_k \to \infty$ and $F_{i_k} \to F$, then $F \in K$.

Theorem 5.3 *([Gr 86], p. 218; [MS 96]). Suppose that K admits an in-approximation $\{U_i\}$. Let $v \in C^1(\Omega; \mathbf{R}^m)$ with*

$$Dv \in U_1.$$

Then there exists a Lipschitz map u such that

$$Du \in K \in \Omega \text{ a.e.}, \quad u = v \text{ on } \partial\Omega.$$

Proof. The proof uses a sequence of approximations obtained by successive application of Lemma 5.2. To achieve strong convergence each approximation uses a much finer spatial scale than the previous one, similar to the construction of continuous but nowhere differentiable functions. This is one of the key ideas of convex integration.

We first construct a sequence of piecewise linear maps u_i that satisfy

$$\begin{array}{rcllrcll} Du_i & \in & U_i & \text{a.e,} & & & & \\ \sup|u_{i+1} - u_i| & < & \delta_{i+1}, & & u_{i+1} & = & u_i & \text{on } \partial\Omega, \\ \sup|u_1 - v| & < & \delta/4, & & u_1 & = & v & \text{on } \partial\Omega. \end{array}$$

To construct u_1 note that if Ω' is open and $\Omega' \subset\subset \Omega$ then $\operatorname{dist}(Dv(x), \partial U_1) \geq c(\Omega') > 0$ for all $x \in \Omega'$. Hence it is easy to obtain $u_1|\Omega'$ by introducing a sufficiently fine triangulation. Now exhaust Ω by an increasing sequence of sets $\Omega_i \subset\subset \Omega$.

To construct u_{i+1} and δ_{i+1} from u_i and δ_i we proceed as follows. Let

$$\Omega_i = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > 2^{-i}\}.$$

Let ρ be a usual mollifying kernel, i.e. let ρ be smooth with support in the unit ball and $\int \rho = 1$. Let

$$\rho_\epsilon(x) = \epsilon^{-n}\rho(x/\epsilon).$$

Since the convolution $\rho_\epsilon * Du_i$ converges to Du_i in $L^1(\Omega_i)$ as $\epsilon \to 0$ we can choose $\epsilon_i \in (0, 2^{-i})$ such that

$$||\rho_{\epsilon_i} * Du_i - Du_i||_{L^1(\Omega_i)} < 2^{-i}. \tag{5.5}$$

Let

$$\delta_{i+1} = \delta_i \epsilon_i. \tag{5.6}$$

Use Lemma 5.2 to obtain u_{i+1} such that $Du_{i+1} \in U_{i+1}$, $u_{i+1} = u_i$ on $\partial\Omega$ and

$$\sup_\Omega |u_{i+1} - u_i| < \delta_{i+1}. \tag{5.7}$$

Since $\delta_{i+1} \leq \delta_i/2$ we have

$$\sum_{i=1}^{\infty} \delta_i \leq \delta/2.$$

Thus

$$u_i \to u_\infty \quad \text{uniformly},$$

and u_∞ is Lipschitz since the u_i are uniformly Lipschitz (by (ii) in the definition of an in-approximation). Moreover $u_\infty = v$ on $\partial\Omega$.

It only remains to show that $Du_\infty \in K$. The key point is to ensure strong convergence of Du_i. Since $||D\rho_\epsilon||_{L^1} \leq C/\epsilon$ we deduce from (5.7) and (5.6)

$$\begin{array}{rcl} ||\rho_{\epsilon_k} * (Du_k - Du_\infty)||_{L^1(\Omega_k)} & = & ||D\rho_{\epsilon_k} * (u_k - u_\infty)||_{L^1(\Omega_k)} \\ \leq \frac{C}{\epsilon_k} \sup |u_k - u_\infty| & \leq & \frac{C}{\epsilon_k} \sum_{j=k+1}^{\infty} \delta_j \\ \leq 2\frac{C}{\epsilon_k}\delta_{k+1} & \leq & C'\delta_k. \end{array} \tag{5.8}$$

Taking into account (5.5) it follows that

$$\begin{array}{rcl} ||Du_k - Du_\infty||_{L^1(\Omega)} & \leq & C'\delta_k + 2^{-k} + ||\rho_{\epsilon_k} * Du_\infty - Du_\infty||_{L^1(\Omega_k)} \\ & + & ||Du_k - Du_\infty||_{L^1(\Omega \setminus \Omega_k)}. \end{array}$$

Since Du_k and Du_∞ are bounded we obtain $Du_k \to Du_\infty \quad \text{in } L^1(\Omega)$. Therefore there exists a subsequence u_{k_j} such that

$$Du_{k_j} \to Du_\infty \quad \text{a.e.}$$

It follows from the definition of an in-approximation that

$$Du_\infty \in K \quad \text{a.e.}$$

Hence $u = u_\infty$ has the desired properties. □

For the two-well problem (5.1) - (5.3) one can construct an in-approximation using the explicit formula in Theorem 4.12. The details can be found in [MS 96]; for a different approach based on Baire's theorem see Dacorogna and Marcellini [DM 96a], [DM 96b], [DM 97].

Theorem 5.4 *Suppose that $\lambda\mu > 1$. Then the two-well problem* (5.1) - (5.3) *has a solution if*

$$F \in \mathrm{int} K^{lc},$$

where

$$K^{lc} = \left\{ F = (y,z) : |y| \le \frac{\lambda\mu - \det F}{\lambda\mu - 1}, |z| \le \frac{\det F - 1}{\lambda\mu - 1} \right\}.$$

Remark. A similar result holds if $\lambda\mu = 1$ provided that in the definition of in-approximation and interior one considers relatively open sets subject to the constraint $\det F = 1$. One only needs to use the remark after Lemma 5.1 to achieve $\det Du = 1$, provided that $\det A = \det B = 1$.

A more detailed analysis shows that in the definition of in-approximation one can replace the lamination convex hull which is based on explicit rank-1 connections by the rank-1 convex hull defined by duality with functions (see Section 4.4). This has a striking consequence for the four-gradient example

$$K = \left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \pm \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

discussed in Section 2.6, see in particular Figure 9. For any matrix

$$F \in K^{rc} \supset \left\{ \begin{pmatrix} F_{11} & 0 \\ 0 & F_{22} \end{pmatrix} : \ |F_{11}| \le 1, |F_{22}| \le 1 \right\}$$

and any open neighbourhood $U \supset K$ there exists a map $u : \Omega \to \mathbf{R}^2$ such that

$$\begin{array}{lll} Du \in U & \text{a.e.} & \text{in } \Omega, \\ u = Fx & & \text{on } \partial\Omega. \end{array}$$

This is true despite the fact that small neigbourhoods contain no rank-1 connections so at first glance there seems to be no way to start the construction.

This obstacle is overcome by first constructing a (piecewise linear) map that satisfies $Dv \in U^{rc}$ a.e. and $Dv \in U$ except on a set of small measure. One can then show that the exceptional set can be inductively removed.

The major outstanding problem is whether in the definition of an in-approximation one can replace the lamination convex hull (or rank-1 convex hull) by the quasiconvex hull. One key step would be to resolve the following question.

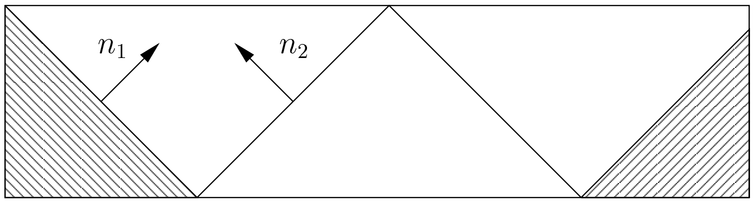

Figure 18: Structure of solutions with finite perimeter. The normals n_1, n_2 are determined by (5.4).

Conjecture 5.5 *Let K be a compact quasiconvex set, i.e. $K^{qc} = K$ and let $\nu \in \mathcal{M}^{qc}(K)$. Then for every open set $U \supset K$ there exists a sequence $u_j : (0,1)^n \to \mathbf{R}^m$ such that Du_j generates ν and $Du_j \in U$ a.e.*

The conjecture is true for compact convex sets [Mu 97a]; this refines Zhang's Lemma (see Lemma 4.21) which implies the existence of u_j such that $Du_j \in B(0,R)$ for a sufficiently large ball.

5.2 Regularity and rigidity

The construction outlined above yields very complicated solutions of the two-well problem (5.1) - (5.3). This raises the question whether the geometry of the solutions can be controlled. Consider the set

$$E = \{x \in \Omega : \ Du(x) \in SO(2)A\}$$

where Du takes values in one connected component of K (or one phase in the applications to crystals). The perimeter of a set $E \subset \Omega \subset \mathbf{R}^n$ is defined as

$$\text{Per}E = \sup \left\{ \int_E \text{div}\, \varphi \, dx : \varphi \in C_0^1(\Omega, \mathbf{R}^n), \ |\varphi| \le 1 \right\}.$$

For smooth or polyhedral sets this agrees with the $(n-1)$ dimensional measure of ∂E.

Theorem 5.6 *([DM 95]). If u is a solution of* (5.1) - (5.3) *and if* $\text{Per}E < \infty$ *then u is locally a simple laminate and ∂E consists of straight line segments that can only intersect at $\partial\Omega$.*

The proof combines geometric and measure-theoretic ideas. The geometric idea is that the Gauss curvature $K(g)$ of the pull-back metric $g = (Du)^T Du$ should vanish (in a suitable sense). Since g only takes two values this should give information on E.

One key step in the implementation of this idea is a finite perimeter version of Liouville's theorem on the rigidity of infinitesimal rotations (cf. Theorem 2.4). In this framework connected components are replaced by indecomposable components. A set A of finite perimeter is indecomposable if for every $A_1 \subset A$ with $\mathrm{Per} A = \mathrm{Per} A_1 + \mathrm{Per} A \setminus A_1$ the set A_1 or $A \setminus A_1$ has zero measure. It can be shown that each set of finite perimeter is a union of at most countably many indecomposable components.

Theorem 5.7 *Suppose that $u : \Omega \subset \mathbf{R}^n \to \mathbf{R}^n$ belongs to $W^{1,\infty}(\Omega; \mathbf{R}^n)$ and that $\det Du \geq c > 0$. Suppose further that $E \subset \Omega$ has finite perimeter and*

$$Du \in SO(n) \qquad \text{a.e. } in\ E.$$

Then Du is constant on each indecomposable component of E.

To finish the proof of Theorem 5.6 one can decompose Du as $e^{i\Theta} g^{1/2}$ (where $g = (Du)^T Du \in \{A^T A, B^T B\}$) and analyze the jump conditions at the boundary of each indecomposable component to deduce that Θ only takes two values and solves (in the distributional sense) a wave equation with characteristic directions n_1 and n_2.

B. Kirchheim recently devised more flexible measure-theoretic arguments, and combining them with algebraic ideas he established a generalization of Theorem 5.6 to the three-well problem $K = \bigcup_{i=1}^{3} SO(3) U_i$ in three dimensions with $U_1 = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_2)$, $U_2 = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_1)$, $U_3 = \mathrm{diag}(\lambda_2, \lambda_2, \lambda_1)$, $\lambda_i > 0$. A major additional difficulty in this case is that the gauge group $SO(3)$ is not abelian and one cannot hope to derive a linear equation for a quantity like Θ in the two-dimensional situation.

6 Length scales and surface energy

Minimization of the continuum elastic energy is a drastic simplification, in particular if a very fine mixture of phases is observed. It neglects interfacial energy as well as discreteness effects due to the atomic lattice. It is therefore not surprising that elastic energy minimization often predicts an infinitesimally fine mixture of phases (in the sense of a nontrivial Young measure), whereas in any real crystal all microstructures are of finite size.

Nonetheless elastic energy minimization does surprisingly well. It often correctly predicts the phase proportions and in combination with considerations of rank-1 compatibility the orientation of phase interfaces. It recovers in particular the predictions of the crystallographic theory of martensite. In fact one of the major achievements was to realize that the predictions of that theory can be understood as consequences of energy minimization. This allows one to bring to bear the powerful methods of the calculus of variations in the analysis of microstructures.

The problem that elastic energy minimization does not determine the length scale and fine geometry of the microstructure remains. It can be overcome by introducing a small amount of interfacial energy or higher gradient terms. One expects these contributions which penalize rapid changes to be small since otherwise a very fine structure would not arise in the first place. The most popular functionals are

$$I^\epsilon(u) = \int_\Omega W(Du)dx + \int_\Omega \epsilon^2 |D^2 u|^2 dx \tag{6.1}$$

and

$$J^\epsilon(u) = \int_\Omega W(Du)dx + \int_\Omega \epsilon |D^2 u| dx. \tag{6.2}$$

The second functional allows for jumps in the gradient, and $|D^2u|$ is understood as the total variation of a Radon measure.

The small parameter $\epsilon > 0$ introduces a length scale and as $\epsilon \to 0$ both models approach (at least formally) pure elastic energy minimization. More realistic models should of course involve anisotropic terms in D^2u or more generally terms of the form $h(Du, \epsilon D^2 u)$. Even the basic models (6.1) and (6.2) are, however, far from being understood for maps $u : \Omega \subset \mathbf{R}^3 \to \mathbf{R}^3$. In the following we discuss briefly two simple scalar models which already show some of the interesting effects generated by the interaction of elastic energy and surface energy.

6.1 Selection of periodic structures

As a simple one-dimensional counterpart of the two-well problem consider the problem

$$\text{Minimize } I(u) = \int_0^1 (u_x^2 - 1)^2 + u^2 \, dx \tag{6.3}$$

subject to periodic boundary conditions. Clearly $I(u) > 0$ since the conditions $u = 0$ a.e. and $u_x = \pm 1$ a.e. are incompatible. On the other hand $\inf I = 0$, since a sequence of finely oscillating sawtooth functions u_j can achieve $u_{jx} \in \{\pm 1\}, u_j \to 0$ uniformly. For any such sequence u_{jx} generates the (unique) Young measure $\nu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ (see Section 3.2b)). Note that there are many 'different' sequences that generate this Young measure. Minimizers of the singularly perturbed functional

$$I^\epsilon(u) = \int_0^1 \epsilon^2 u_{xx}^2 + (u_x^2 - 1)^2 + u^2 \, dx$$

yield a very special minimizing sequence for I.

Theorem 6.1 *If $\epsilon > 0$ is sufficiently small then every minimizer of I^ϵ (subject to periodic boundary conditions) is periodic with minimal period $P^\epsilon = 4(2\epsilon)^{1/3} + \mathcal{O}(\epsilon^{2/3})$.*

A more detailed analysis shows that the minimizers u^ϵ look approximately like a sawtooth function with slope ± 1 and involve *two* small length scales: the sawtooth has period $\sim \epsilon^{1/3}$, and its corners are rounded off on a scale $\sim \epsilon$ (see Fig. 19).

The heuristics behind the proof of Theorem 6.1 is simple and relies on two observations. First, the condition $I^\epsilon(u^\epsilon) \to 0$ enforces that u^ϵ is almost a sawtooth function with slopes ± 1. Second, a key observation of Modica and Mortola is that the first two terms of the energy combined essentially count (ϵ times) the number of changes in the slope from 1 to -1 and vice versa. Indeed the arithmetic geometric mean inequality yields for any interval $(a, b) \subset (0, 1)$

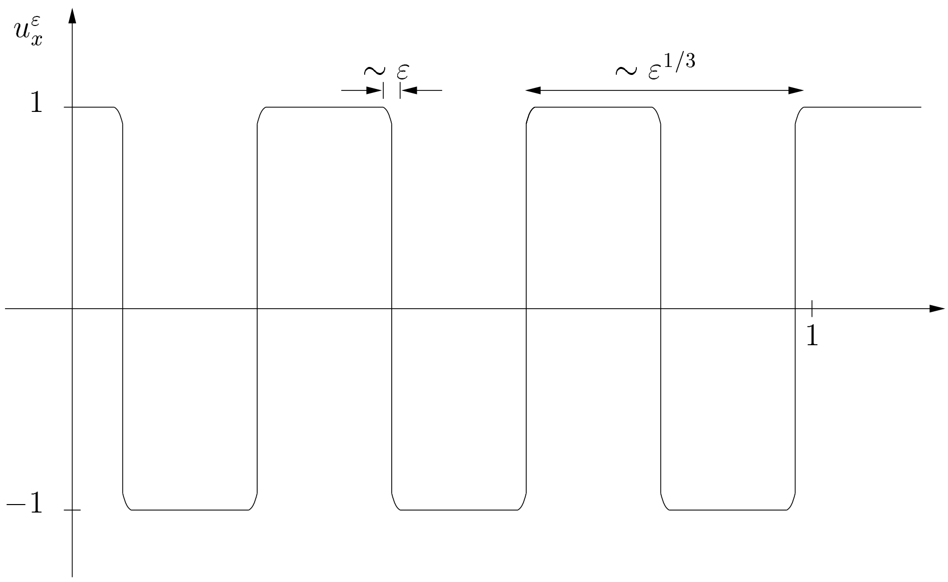

Figure 19: Sketch of u_x^ϵ for a minimizer of I^ϵ

over which u_x changes sign

$$\int_a^b \epsilon^2 u_{xx}^2 + (u_x^2 - 1)^2 \, dx \geq \int_a^b 2\epsilon |(u_x^2 - 1) u_{xx}| \, dx$$

$$\geq \quad \epsilon \, | \int_a^b H'(u_x) \, dx | \geq \epsilon \, |H(u_x(b)) - H(u_x(a))|$$

$$\approx \quad \epsilon \, |H(1) - H(-1)|,$$

where $H'(t) = 2|t^2 - 1|$. On the other hand the above estimates can be made sharp if one choose u as a solution of the ODE $\epsilon u_{xx} = (u_x^2 - 1)^2$, e.g. $u_x = \tanh \frac{x - x_0}{\epsilon}$.

The two observations strongly suggest that (6.3) is essentially equivalent to the following "sharp-interface problem"

$$\text{Minimize} \qquad \epsilon A_0 N + \int_0^1 u^2 \, dx \qquad \text{among periodic function with } |u_x| = 1. \tag{6.4}$$

Here N denotes the number of sign changes of u_x and $A_0 = H(1) - H(-1) = 8/3$. For fixed N (6.4) is a discrete problem, and a short calculation shows that in this case periodically spaced sign changes of u_x are optimal and the second term in the energy becomes $\frac{1}{12}N^{-2}$. Minimization over N yields the assertion.

The actual proof of Theorem 6.1 uses the expected analogy between (6.4) and (6.3) only as a guiding principle and proceeds by careful approximations and estimates for odes. Nonetheless it would be very useful to relate (6.4) and (6.3) in a rigorous way, also as a test case for higher dimensional problems where the fine ode methods are not available. Conventional Γ-convergence methods do not apply since the problem involves two small length scales and the passage from (6.3) to (6.4) corresponds to removing only the faster one (i.e. the smoothing of the sawtooth's corners). Recently G. Alberti and the writer developped a new approach that allows one to do that. One of the main ideas is to introduce a new variable y that corresponds to the slower scale and to view

$$v^\epsilon(x, y) := \epsilon^{-1/3} u(x + \epsilon^{1/3} y)$$

as a map V^ϵ from $(0, 1)$ into a suitable function space X via $V^\epsilon(x) = v^\epsilon(x, \cdot)$. One can endow X with a topology that makes it a compact metric space and study of the Young measure ν generated by V^ϵ. For each $x \in (0, 1)$ the measure ν_x is a probability measure on the function space X. If u^ϵ is a sequence of (almost) minimizers of I^ϵ then one can show that ν_x is supported on translates of sawtooth functions with the optimal period $4\ 2^{1/3}$.

One easily checks that the asymptotic behaviour is the same for minimizers of (6.4), and this gives a precise meaning to the assertion that (6.3) and (6.4) are asymptotically equivalent.

This approach is inspired by the idea of two-scale convergence ([Al 92], [E 92], [Ng 89]). A crucial difference is that two-scale convergence usually only applies if the period of the microstructure is fixed and possible phase shifts are controlled. This is the case if, for example, the solutions are of the form $\tilde{u}(x, \frac{x}{\epsilon^{1/3}})$ where $\tilde{u}$ is periodic in the second variable.

6.2 Surface energy and domain branching

Consider the two-dimensional scalar model problem (see [KM 92] for the relation with three-dimensional elasticity)

$$I(u) = \int_0^1 \int_0^L u_x^2 + (u_y^2 - 1)^2 dx\, dy \overset{!}{\to} \min,$$

$$u = 0 \text{ on } x = 0. \tag{6.5}$$

The integrand is minimized at $Du = (u_x, u_y) = (0, \pm 1)$. The preferred gradients are incompatible with the boundary condition. The infimum of I subject to (6.5) is zero but not attained. The gradients Du_j of any minimizing sequence generate the Young measure $\frac{1}{2}\delta_{(0,-1)} + \frac{1}{2}\delta_{(0,1)}$. One possible construction of a minimizing sequence is as follows (see Fig. 20). Let s_h be a periodic sawtooth function with period h and slope ± 1 and let $u(x, y) = s_h(y)$ for $x \geq \delta$, $u(x, y) = \frac{x}{\delta} s_h(y)$ for $0 \leq x < \delta$. Then consider a limit $h \to 0, \delta \to 0$ such that h/δ remains bounded. Similar reasoning applies if we replace (6.5) by the condition that u vanishes on the whole boundary of $[0, L] \times [0, 1]$.

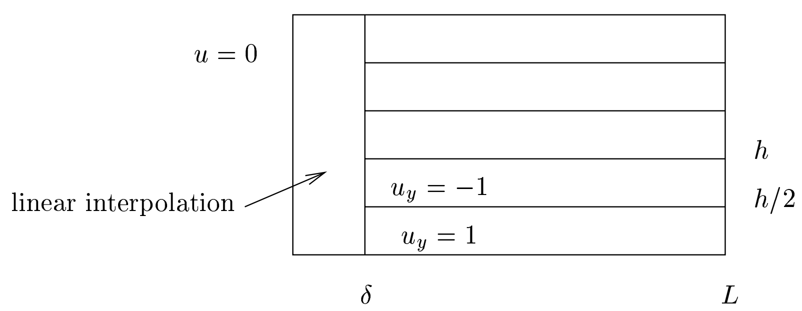

Figure 20: Construction of a minimizing sequence.

To understand the influence of regularizing terms on the length scale and the geometry of the fine scale structure we consider

$$I^\epsilon(u) = \int_0^1 \int_0^L u_x^2 + (u_y^2 - 1)^2 + \epsilon^2 u_{yy}^2 dx\, dy,$$

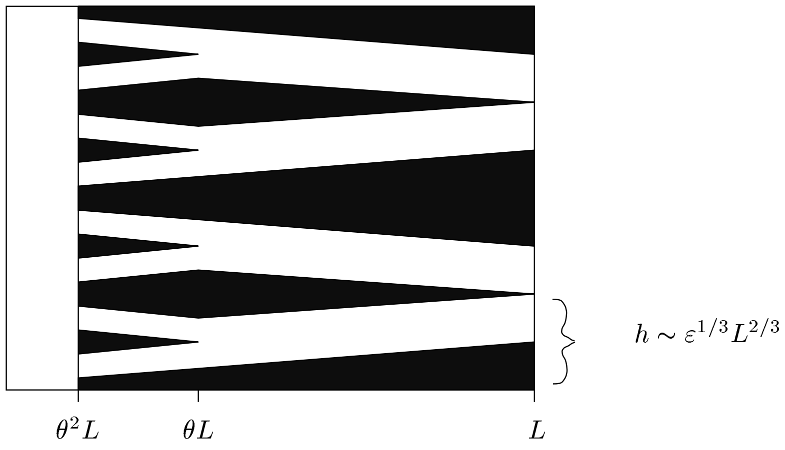

Figure 21: The self-similar construction with $1/4 < \Theta < 1/2$. Only two generations of refinement are shown.

subject to (6.5). Instead of the second derivatives in y one can consider other regularizing terms, e.g. $|D^2u|^2$. The derivatives in y are, however, the most important ones, since we expect that fine scale oscillations arise mainly in the y direction. It was widely believed that for small $\epsilon > 0$ the minimizers of I^ϵ look roughly like the construction $u_{h,\delta}$ depicted in Figure 20 (with the corners of the sawtooth 'rounded off' and optimal choices $\delta(\epsilon), h(\epsilon)$). This is false. Indeed a short calculation shows that $\delta(\epsilon) \sim (\epsilon L)^{1/2}$, $h(\epsilon) \sim (\epsilon L)^{1/2}$ and $I^\epsilon(u^\epsilon_{h,L}) \sim \epsilon^{1/2}L^{1/2}$. On the other hand one has

Theorem 6.2 *([Sch 94]) For* $0 < \epsilon < 1$ *there exists constants* $c, C > 0$ *such that*

$$c\epsilon^{2/3}L^{1/3} \leq \min_{u=0 \text{ at } x=0} I^\epsilon \qquad \leq C\epsilon^{2/3}L^{1/3}.$$

The upper bound is obtained by a smooth version of the self-similar construction depicted in Figure 21.

The mathematical issues become clearer if we again replace I^ϵ by a sharp interface version

$$J^\epsilon(u) = \int_0^L \int_0^1 u_x^2 + \epsilon|u_{yy}|dydx \tag{6.6}$$

subject to

$$|u_y| = 1 \text{ a.e.} \tag{6.7}$$

Thus $y \mapsto u(x,y)$ is a sawtooth function and $\int_0^1 |u_{yy}| dy$ denotes twice the number of jumps of u_y. Minimization of (6.6) subject to (6.7) is in fact a purely geometric problem for the set

$$E = \{(x,y) : u_y(x,y) = 1\}.$$

The first term in J^ϵ is a nonlocal energy in terms of E, while the second is essentially the length of ∂E (more precisely its projection to the x-axis; as before we consider this to be the essential part since oscillations occur mainly in the y direction). The functional and the constraint are invariant under the scaling

$$u_\lambda(x,y) = \lambda^{-1} u(\lambda^{3/2} x, \lambda y)$$

which suggests a self-similar construction with $\Theta = \left(\frac{1}{2}\right)^{3/2}$.

Theorem 6.3 *([KM 94]). For* $0 < \epsilon < 1$ *one has*

$$c\epsilon^{2/3} L^{1/3} \le \min_{(6.5)(6.7)} J^\epsilon \le C\epsilon^{2/3} L^{1/3}.$$

Moreover, if $\bar{u}$ *is a minimizer of* J^ϵ *subject to* (6.5), (6.7) *then*

$$c\epsilon^{2/3} l^{1/3} \le \int_0^1 \int_0^l \bar{u}_x^2 + \epsilon |\bar{u}_{yy}| dx dy \le C\epsilon^{2/3} l^{1/3}. \tag{6.8}$$

The scaling in (6.8) is exactly the scaling predicted by the self-similar construction with $\theta = (\frac{1}{2})^{3/2}$.

The prediction of refinement of the microstructure (domain branching) towards the boundary $x = 0$ in the simple model (6.5–(6.7) inspired new experimental investigations ([Sch 93]). In closely related models for magnetization domains in ferromagnetic materials domain branching is experimentally well established ([Li 44], [Hu 67], [Pr 76]), a rigorous mathematical analysis is just beginning to emerge ([CK 97b], [CKO 97]). Already a quick look at some of the sophisticated constructions in [Pr 76] suggests that a lot is to be discovered.

7 Outlook

There are many other interesting aspects of microstructure and I can mention only three areas: alternative descriptions of microstructure, dynamics and computation.

7.1 Alternative descriptions of microstructure

Young measures are but one way to describe microstructure and to extract 'relevant' information from a sequence of rapidly oscillating functions. They determine the asymptotic local distribution of function values but contain no information about the direction, length scale or fine geometry of the oscillations. As we saw in Section 3.3 the Young measure does not suffice to determine the limits of natural nonlocal quantities such as the magnetostatic energy or the self-correlation function.

There is an intense search for new objects that record additional information, see [Ta 95] for a survey. One such object was introduced independently by Tartar [Ta 90] and Gérard [Ge 91] under the names 'H-measure' and 'microlocal defect measure', respectively. They show that for every sequence $\{u_j\}$ that converges to zero weakly in $L^2(\Omega)$ there exists a subsequence $\{u_{j_k}\}$ and a Radon measure μ on $\bar{\Omega} \times S^{n-1}$ (the H-measure of $\{u_{j_k}\}$) such that for every pseudo-differential operator A of order zero with (sufficiently regular) symbol $a(x, \xi)$ one has

$$\langle A u_{j_k}, u_{j_k} \rangle_{L^2} \to \int_{\bar{\Omega} \times S^{n-1}} a d\mu.$$

For $\mathbf{R}^m$-valued sequences one similarly obtains a matrix-valued (hermitian) measure $\nu = (\nu^{ij})_{1 \le i,j \le m}$. The H-measure suffices, for example, to compute the limit of the micromagnetic energy discussed in Example 1 of Section 3.3 (the corresponding matrix valued symbol is just $a(\xi) = \frac{\xi}{|\xi|} \otimes \frac{\xi}{|\xi|}$). Other applications of the H-measure include small amplitude homogenization, compensated compactness with variable coefficients, compactness by averaging in kinetic equations and the propagation of energy concentrations in linear hyperbolic systems.

Two outstanding open problems are the relation between H-measures and Young measures (see [MT 97], [Ta 95] for partial results) and a useful generalization that allows one to compute limits of nonquadratic quantities; even the case of trilinear expressions is open.

The H-measure tracks the energy of oscillations depending on the direction, but regardless of length scales. P. Gèrard [Ge 90] introduced a variant of the H-measure, called semiclassical measure, that allows one to study the effect of oscillations on a typical length scale $h_j \to 0$ (see also [LP 93]). A completely different approach to analyze the detailed behaviour on small length scales was briefly discussed at the end of Section 6.1.

7.2 Dynamics

Three fundamental questions are:

> Can realistic dynamics create microstructure?
>
> Can one deduce a law for the evolution of microstructure from the macroscopic laws, and possibly reasonable additional assumptions?
>
> What is a 'good' evolution law for interfaces in complex microstructures and how can one model hysteresis?

A typical setting for the first question is a dynamical system that admits a Liapunov function (such as energy or entropy) for which there exist no classical minimizer. Will the dynamics drive the Liapunov function to its infimum and hence create fine scale oscillations or will the dynamics generate compact orbits (in a suitable energy space) and thus prevent global minimization of the Liapunov function?

The papers [Ba 90a] and [BHJPS 91] give nice surveys; Friesecke and McLeod [FM 96] solved a longstanding problem by showing that one-dimensional viscoelastic dynamics with a nonconvex elastic energy does not generate microstructure.

A precise setting for the second question is as follows. Consider a sequence of rapidly oscillating initial data that generate a certain Young measure (or H-measure, semiclassical measure, ...). Is the Young measure of the solution at a later time determined by the Young measure of the initial data? In physical language this is closely related to the idea of coarse-graining. Given an evolution law for a very complex pattern are there simpler laws for certain gross quantities such as the local phase average (= Young measure = one-point statistics)? If this can be achieved it can lead not only to new insights but also to huge savings in computer time and more reliable results since it is no longer needed to resolve the finest scale of the pattern.

A typical obstacle in attacking these questions is the closure problem. Often the time derivatives of certain moments of the Young measure involve higher moments. Even worse, sometimes the time derivative of the Young measure involves terms that depend on two-point or higher correlations which cannot be determined from the Young measure (see Example 2 of Section 3.3).

The first results on the evolution of Young measures and creation or non-creation of oscillations were obtained by Tartar for kinetic models and more general semilinear hyperbolic systems ([Ta 80], [Ta 81], [Ta 84], [Ta 86], [Ta 87], [MPT 85]), see [Jo 83], [JMR 95], [Mi 97] for further developments. In [FBS 94] Tartar's ideas were used to study the evolution of Young measures for a viscoelastically damped wave equation with nonmonotone stress-strain relation. Theil [Th 97] recently obtained very sharp results on this problem by a modification of the method that relies on transport theory rather than on a study of the moments of the measure. Otto derived equations for the evolution of microstructure in unstable two-phase flow through porous media [Ot 95] and in magnetic fluids [Ot 98]; further references on the evolution of microstructure include [De 96] and [HR 94].

Regarding the third question about evolution laws and hysteresis Chu and James observed that the hysteresis curves obtained in cyclic biaxial loading of a Cu-Al-Ni single crystal cannot be explained by usual kinetic laws. One alternative approach is based on metastability induced by lack of rank-1 connections [BCJ 95], another interesting route is explored in [ACJ 96]: the energy landscape in function space contains many local minima (that correspond to different microstructures), and the study of the effective evolution laws for such 'wiggly' potentials yields surprising conclusions already in a simple model; see also [Kin 97]. For other views of hysteresis, see the survey article of Huo and I. Müller [HM 93], the recent monograph of Brokate and Sprekels [BS 96] and the series of lectures [Br 94]; general references for hysteresis include [KP 89] and [Vi 94].

7.3 Computation

The computation of microstructure by numerical energy minimization is a very challenging task, see Luskin [Lu 96] for a recent survey. If microstructure is numerically observed, it often forms on the scale of the underlying mesh. Hence calculations are notoriously mesh dependent unless (expensive) regularizations are included or special care is taken.

So far most numerical schemes do not make use of analytical insights, except for scalar problems where relaxation leads not only to a drastic speed up but also to more accurate results ([CP 97]). Some other exceptions are discussed in Sections 1 and 7 of [Lu 96]. One difficulty in using analytical information in higher dimensions is that quasiconvexity while being the natural convexity notion (see Section 4.2) is still largely mysterious and no efficient algorithm for the computation of the quasiconvex hull is known. At least for rank-convex hulls there has been some progress in [MP 98] and [Do 97].

One important issue is how to represent microstructures numerically in an efficient way. Currently mostly finite element approaches are used but they require a lot of unknowns to represent simple microstructures such as an order 2 laminate (cf. Fig. 14). Ideally a good representation should both yield a high compression ratio and be well adapted to the numerical algorithm. The search for better analytical objects to describe microstructure discussed above may well be relevant here.

7.4 Some solved and unsolved problems

The following table gives an overview of the state of Problems 1 and 2 for sets K without rank-one connections. Further examples and references can be found in [Ba 90b], [BFJK 94], [Sv 95].

Acknowledgements

The writing of this article and the work reported here would not have been possible without the support and encouragement of many colleagues and friends. Particular thanks go to G. Alberti, J.M. Ball, K. Bhattacharya, G. Dolzmann, I. Fonseca, G. Friesecke, R.D. James, D. Kinderlehrer, B. Kirchheim, J. Kristensen, R.V. Kohn, V. Šverák and L. Tartar. I am also very greatful to U. Knieper and M. Krieger for their excellent typing of these notes.

I thank R. Conti, S. Hildebrandt, and M. Struwe for their invitation to prepare these lectures and for providing such a delightful and stimulating atmosphere at the C.I.M.E. summer school in Cetraro.

In the preparation of the lectures I greatly benefited from an invitation by J. Lukeš and J. Maly to present a similar set of lectures at a Spring School in Paseky in 1995.

K	Nontrivial exact soln.	Nontrivial GYM	Method	References
$\{A, B\}$	No	No	Minors or Cauchy-Riemann eqns.	[BJ 87] [Sv 91a]
$\{A, B, C\}$	No	No	Nonconvex solns. of Monge-Ampère	[Sv 91b] [Sv 92b]
finite	?	Yes	Example with four matrices	[AH 86], [Ta 93], [CT 93], [BFJK 94]
countable	Yes	Yes	Example	[Sv]
$SO(2); \mathbf{R}SO(2)$	No; holomorphic fns.	No	Cauchy-Riemann eqns.	[Sv 91a]
$SO(n); \mathbf{R}SO(n)$	No;	No	Minors;degenerate	[Ki 88]
$n \geq 3$	Möbius maps	No	elliptic eqns.	[Re 68a]
$\cup_{i=1}^{k} SO(2)A_i$	No	No	Minors or elliptic regularity	[Sv 93a]
$SO(3)A \cup SO(3)B$	?	?	Special cases known	[Ma 92], [Sv 93a]
scalar conservation laws	No	No	compensated compactness; div-curl lemma	[Ta 79b]
$m \times 2$ elliptic		Yes, $m \geq 6$		[Sv 95]
systems	Yes, $m = 2$	Yes, $m = 2$		[MS 98]

Table 1: Nontrivial exact solutions and nontrivial gradient Young measures (GYM) for incompatible sets K (i.e. $\mathrm{rk}(A - B) \neq 1 \; \forall A, B \in K$).

Notes

Here I have collected some additional references to the literature without any pretention to be exhaustive or impartial.

Chapter 1 The idea to use nonlinear continuum theory for elastic crystals and solid-solid phase transformations goes back to Ericksen [Er 75], [Er 77], [Er 80], [Er 84], [Er 89] (see also [Gu 83], [Ja 81], [Pa 81], [Pi 84]) and was developed in the context of the calculus of variations by Ball and James ([BJ 87], [BJ 92]), Chipot and Kinderlehrer [CK 88], Fonseca [Fo 87] and subsequently by many others. There is a similar theory for micromagnetism ([Br 63], [DS 93], [JK 90]) and magnetostriction ([Br 66], [JK 93]).

The analytical foundations of the theory go back to the fundamental work of Morrey ([Mo 52], [Mo 66]) on lower semicontinuity (extended by Reshetnyak [Re 67], [Re 89] to problems in quasiconformal geometry and by Ball [Ba 77] to nonlinear elasticity) and to the pioneering work of Tartar on compensated compactness (partly in collaboration with Murat) and on weak convergence as a tool to pass from microscopic to macroscopic descriptions. His work in the seventies is summarized in the seminal paper [Ta 79b], some more recent developments are discussed in [Ta 90], [Ta 93] and [Ta 95], and a comprehensive treatment will appear in [Ta 98]. While the current notes focus mostly on variational problems Tartar's approach is more general. In view of applications to nonlinear partial differential equations in continuum mechanics he considers general combinations of pointwise constraints $w \in K \subset \mathbf{R}^d$ a.e. (these usually arise from constitutive equations) and differential constraints $\sum_{j,k} a_{ijk}\partial_j w^k = 0$ (or in a compact set of $W^{-1,p}$; these correspond to the balance laws). The situation considered in the current notes corresponds to the constaint curl $w = 0$.

Partially motivated by Eshelby's classic work on ellipsoidal inclusions [Es 57, Es 59, Es 61], Khachaturyan, Roitburd and Shatalov ([Kh 67], [KSh 69], [Ro 69], [Ro 78]) developed already in the sixties a theory of microstructure based on energy minimization in the context of linear elasticity; see [Kh 83] for a comprehensive treatment. Comparisons between the linear and the nonlinear theory appear in [BJ 92], [Bh 93] and [Ko 89].

For lack of space I have not been able to discuss the close relation between the variational approach to microstructures and the theory of optimal design and optimal composites. Some constructions used in optimal design closely resemble observed phase arrangements in solid-solid phase transitions. Early work in this direction includes [Ta 75], [Ta 79b], [Mu 78] and [KL 71]. A number of important papers that were previously difficult to access have recently appeared in English translation in [CK 97a]. Further references can be found both in the introduction and the individual articles of that volume as well as in the forthcoming books [Mi 98] and [Ta 98].

The approach to microstructures via energy minimization provides a new foundation for the crystallographic theory of martensite ([BM 54], [WLR 53]) and has found important applications which include: new criteria for the reversible shape-memory effect based on the possibility of self-accommodation of the transformed phase [Bh 92], bounds for the recoverable strains in polycrystals and their dependence on the symmetry of the phase transformation and material texture ([BK 96], [BK 97], [BRL 97]), a proposed design of micromachines that are based on thin films of shape-memory materials [BhJ 97] and the discovery of a new magnetostrictive material with greatly enlarged magnetostrictive constant [JW 97].

The book by Pitteri and Zanzotto [PZ 97] and the forthcoming book by Ball and James [BJ 97] as well as the collection of reviews [AMM 98] give an overview of the theory and engineering applications. More on the mathematical side, the recent book of Pedregal [Pe 97] reviews the relevance of microstructure and Young measures in various areas of application, while Roubíček's book [Ro 97] focuses more on the functional analytic aspects. Evans' notes [Ev 90] are an excellent introduction to the application of weak convergence methods to partial differential equations. Many further examples can be found in [BFJK 94] and [Sv 95].

The experimental observations described in Section 2.2 are discussed in detail in [CJ 95] and in Chu's thesis [Ch 93]; a careful comparison of theory and experiment for a variety of solid-solid phase transformations was undertaken by Hane [Ha 97].

Chapter 2 The connection with the Cauchy-Riemann equation appears in [Sv 91a]. A counterpart of Theorem 2.4 holds for quasiconformal maps, i.e. $K = \mathbf{R}^+ SO(n), n \geq 3$. In this case one is led to degenerate elliptic equations and Reshetnyak's work ([Re 68a], [Re 89]) was a breakthrough in the study of quasiconformal and quasiregular maps by pde methods. Part (i) of Theorem 2.5 was proved in [Zh] and also follows from more general results in [JO 90]; the proof given is due to Kirchheim.

Lemma 2.7 is called the 'span restriction' in [BFJK 94] because it implies (in view of Corollary 3.2) that the span of the support of a nontrivial gradient

Young measure must contain a rank-1 line. The result is essentially a special case of Theorem 3 in Tartar's work [Ta 83]. It was probably known in some form to Serre [Se 83] and is implicit in [DP 85]. The use of elliptic theory is a common idea in the theory of microstructures, see e.g. [DP 85], [Ma 92], [Sv 93b], [Sv 95].

Chapter 3 Young measures (also known as parametrized measures, relaxed controls, chattering controls or generalized curves) were invented by L.C. Young [Yo 37]; his book [Yo 69] is a delightful read (see also McShane [MS 40] for early applications of the theory and [MS 89] for a personal review). The theory was generalized to much more general domains, target spaces and integrals by Berliocchi and Lasry [BL 73], Balder [Ba 84] and many others including Kristensen [Kr 94]; recent surveys with extensive references are [Va 90] and [Va 94]. Varifolds (see [Al 66], [Al 72], [Re 68b]) are a generalization of Young measures in a geometric setting. Tartar [Ta 79b, Ta 83] introduced Young measures as a fundamental tool for the study of oscillation effects as well as compactness and existence questions in nonlinear partial differential equations. His theory of compensated compactness allows one to derive nontrivial constraints on the Young measure from the combination of pointwise and differential constraints on the generating sequence. One of the early successes of the theory were applications to conservation laws ([Ta 83], [DP 85]); for other applications see e.g. [Ev 90], [Sv 95], [Ta 98].

The presentation here follows [Ba 89]; Section 3.3 is based on [BJ 94]. Another phenomenon that Young measures cannot detect are concentration effects. Varifolds, currents [FF 60, Fe 69, GMS 89, GMS 96] or H-measures [Ge 91, Ta 90] do better in this regard; see also [FMP 97]. There are various alternative proofs of the fundamental theorem: via disintegration of measures on $\Omega \times \mathbf{R}^d$ (see e.g. [BL 73] for a much more general setting and [Ev 90] for a short proof), via L^∞ weak$*$ precompactness of bounded sequences and the theory of multivalued maps (see [Sy 97]) or by consideration of countable dense sets of integrands f_j and test functions φ_k and diagonalization. Corollary 3.3 is a special case of results in [Ba 84].

Chapter 4 The fundamental connection between quasivonvexity and lower semicontinuity was discovered by Morrey (see [Mo 52], [Mo 66]). Dacorogna [Da 81], [Da 82a] discovered the relation between quasiconvexity and relaxation (see also [AF 84]); his book [Da 89] gives a comprehensive treatment of the different notions of convexity. The work of Acerbi and Fusco [AF 84] and

Marcellini [Ma 85] brought a major technical refinement with the coverage of Carathéodory integrands. Since then many further refinements and generalizations have been achieved; a selection is [ABF 96], [BFM 97], [Fo 96], [Kr 97b], [Ma 94], and many further references can be found there. For the connection between quasiconvexity, regularity and compactness see [Ev 86], [EG 87], [FH 85] and [GM 86].

Tartar has pointed out various weaknesses of quasiconvexity. First, quasiconvexity might not be necessary to obtain existence of minimizers. In view of Ekeland's variational principle [Ek 79] (which makes use of the Bishop-Phelps argument [BP 61]) one can choose minimizing sequences that satisfy in addition div$\sigma_k \to 0$ in $W^{-1,q}$ where $\sigma_k = \frac{\partial f}{\partial F}(Du_k)$ is the stress and hence one does not need to verify lower semicontinuity along arbitrary sequences. To my knowledge this line of thought has not been explored in detail. Secondly it is not clear (indeed rather doubtful) whether quasiconvexity implies the stability of equilibria, i.e. whether the conditions $u_k \overset{*}{\rightharpoonup} u$ in $W^{1,\infty}$, $\sigma_k \overset{*}{\rightharpoonup} \overline{\sigma}$ in L^∞ and div$\sigma_k \to$ div$\overline{\sigma}$ in $W^{-1,\infty}$ do imply $\overline{\sigma} = \frac{\partial f}{\partial F}(Du)$ (by contrast Jensen obtained a nice classification in the scalar case, see [Ta 79b], Theorem 23). Šverák has shown [Sv 95], [Sv 98] that the compactness arguments that are the cornerstone of the regularity theory for minimizers for (uniformly) quasiconvex integrals fail for solutions of the equilibrium equations. For arguments in favour of quasiconvexity, in addition to those in the text, see [BMa 84] (cf. also [Me 65], pp.128–131) and [BM 84], Theorem 5.1.

Sections 4.1 and 4.2 are partially based on [BJ 94]. In the definition of quasiconvexity often additional restrictions on the integrand are imposed. Hüsseinov [Hu 88], [Hu 95] realized that this is not necessary, see also [Fo 88]. Section 4.4 follows partially unpublished lectures by Šverák, see also [Sv 95]. Šverák's counterexample is reminiscent of a counterexample by Tartar in the theory of compensated compactness (see [Ta 79b], pp.185–186). The proof of the classification result follows roughly Kinderlehrer and Pedregal's original work [KP 91] (see also [Kr 94]). Some simplifications, in particular for the nonhomogeneous case, are based on discussions with Alberti. Sychev [Sy 97] recently presented independently a similar approach for the case $1 < p < \infty$. The idea to use the Hahn-Banach theorem to characterize Young measures appears e.g. in [Ta 79b], p.152, for the case without differential constraints; in a similar vein the Krein-Milman theorem is used in [BL 73], p.148. The proof of Theorem 4.4.(i) is by now standard (see [Mo 52]), the proof of Theorem 4.5(iii) is the same as Fonseca's [Fo 88], see also [Hu 88].

Truncation arguments that are closely related to what I called Zhang's lemma were used earlier by Acerbi and Fusco [AF 84], [AF 88], based on work of Liu [Li 77].

Gradient Young measures and quasiconvexity correspond to the constraint curl$v = 0$. As mentioned above, in continuum mechanics and electromagnetism one also meets more general systems of first order constraints $A(Dv) = 0$. If A satisfies a constant rank condition there is a largely parallel theory ([Da 82b], [FM 97]) (in an L^p-setting, $1 < p < \infty$) while the situation is widely open even in simple examples where this condition fails (see [Ta 93]).

Chapter 5 Most of the material is taken from [MS 96] and [Mu 97c]. The basic existence result is the theorem on p.218 of Gromov's book [Gr 86]. A detailed proof for a special case and the application to the two-well problem are described in [MS 96]. The case $K = O(3)$ is studied in [CP 95] (here some simplifications occur since $K^{lc} = \overline{\text{conv}} K$); results for more general isometric maps appear in [Gr 86], Chapter 2.4.11. For variable prescribed singular values see also [CPe 97].

Chapter 6 This chapter is based on [Mu 97c]. Theorem 6.1 is taken from [Mu 93]; the $\varepsilon^{1/3}$ scaling had been predicted earlier by Tartar based on matched asymptotic expansions. The Modica-Mortola inequality [MM 77a, MM 77b] was found shortly after De Giorgi had introduced the notion of Γ-convergence [DG 75], [DGF 75], but was initially somewhat overlooked. With the growing interest in the gradient theory of phase transitions since the mid-80's (see [BF 94], [Bo 90], [FT 89], [Gu 87] [KS 89], [Mo 87] and the references therein) it later became a crucial tool. Dal Maso's book [DM 93] is a good reference on Γ-convergence with a very useful commented bibliography. The influence of surface energy on phase transformations in crystals was studied in a series of papers by Parry and others [MP 86], [Pa 87a], [Pa 87b], [Pa 89], mostly in one-dimensional situations.

References

[ACJ 96] R. Abeyaratne, C. Chu and R.D. James, Kinetics of materials with wiggly energies: theory and application to the evolution of twinning microstructures in a Cu-Al-Ni shape-memory alloy, *Phil. Mag.* **A 73** (1996), 457–497.

[AF 84] E. Acerbi and N. Fusco, Semicontinuity problems in the calculus of variations, *Arch. Rat. Mech. Anal.* **86** (1984), 125–145.

[AF 88] E. Acerbi and N. Fusco, An approximation lemma for $W^{1,p}$ functions, in: *Material instabilities in continuum mechanics and related mathematical problems* (J. M. Ball, Ed.), Oxford Univ. Press, 1988, 1–5.

[Ah 66] V. Ahlfors, *Lectures on quasiconformal mappings*, Van Nostrand, 1966.

[AMM 98] G. Airoldi, I. Müller, S. Miyazaki (eds.), *Shape-memory alloys: from microstructure to macroscopic properties*, Trans. Tech. Pub., in press, 1998.

[AM 97] G. Alberti and S. Müller, in preparation.

[AD 92] J.J. Alibert and B. Dacorogna, An example of a quasiconvex function not polyconvex in dimension two, *Arch. Rat. Mech. Anal.* **117** (1992), 155–166.

[Al 92] G. Allaire, Homogenization and two-scale convergence, *SIAM J. Math. Anal.* **23** (1992), 1482–1518.

[Al 72] W.K. Allard, On the first variation of a varifold, *Ann. Math.* **95** (1972), 417–491.

[Al 66] F. J. Almgren, Plateau's problem: an invitation to varifold geometry, Benjamin, 1966.

[ABF 96] L. Ambrosio, G. Buttazzo and I. Fonseca, Lower semicontinuity problems in Sobolev spaces with respect to a measure, *J. Math. Pures Appl.* **75** (1996), 211–224.

[AH 86] R. Aumann and S. Hart, Bi-convexity and bi-martingales, *Israel J. Math.* **54** (1986), 159–180.

[Ba 84] E.J. Balder, A general approach to lower semicontinuity and lower closure in optimal control theory, *SIAM J. Control Opt.* **22** (1984), 570–598.

[Ba 77] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rat. Mech. Anal.* **63** (1977), 337–403.

[Ba 87] J.M. Ball, Does rank one convexity imply quasiconvexity?, in: *Metastability and incompletely posed problems*, IMA Vol. Math. Appl. **3** (S.S. Antman, J.L. Ericksen, D. Kinderlehrer and I. Müller, eds.), Springer, 1987.

[Ba 89] J.M. Ball, A version of the fundamental theorem for Young measures, in: *PDE's and Continuum Models of Phase Transitions* (M. Rascle, D. Serre, M. Slemrod, eds.), Lecture Notes in Physics **344**, Springer, 1989, 207–215.

[Ba 90a] J.M. Ball, Dynamics and minimizing sequences, in: *Problems involving change of type* (K. Kirchgässner, ed.), Lecture Notes in Physics **359**, Springer, 1990, 3–16.

[Ba 90b] J.M. Ball, Sets of gradients with no rank one connections, *J. Math. Pures Appl.* **69** (1990), 241–259.

[BCJ 95] J.M. Ball, C. Chu and R.D. James, Hysteresis in martensite phase transformations, in: Proc. ICOMAT-95, *J. de physique* IV **5**, colloque C8, 245–251.

[BHJPS 91] J.M. Ball, P.J. Holmes, R.D. James, R.L. Pego and P.J. Swart, On the dynamics of fine structure, *J. Nonlin. Sci.* **1** (1991), 17–70.

[BJ 87] J.M. Ball and R.D. James, Fine phase mixtures as minimizers of energy, *Arch. Rat. Mech. Anal.* **100** (1987), 13–52.

[BJ 92] 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** (1992), 389–450.

[BJ 94] J.M. Ball and R.D. James, *The mathematics of microstructure*, DMV-Seminar (unpublished lectures).

[BJ 97] J.M. Ball and R.D. James, book in preparation.

[BMa 84] J. M. Ball and J. E. Marsden, Quasiconvexity at the boundary, positivity of second variation and elastic stability, *Arch. Rat. Mech. Anal.* **86** (1984), 251–277.

[BM 84] J.M. Ball and F. Murat, $W^{1,p}$-quasiconvexity and variational problems for multiple integrals, *J. Funct. Anal.* **58** (1984), 225–253.

[BKK 98] J. M. Ball, B. Kirchheim and J. Kristensen, in preparation.

[BF 94] A.C. Barroso and I. Fonseca, Anisotropic singular perturbations – the vectorial case, *Proc. Roy. Soc. Edinburgh* **124A** (1994), 527–571.

[BL 73] H. Berliocchi and J.M. Lasry, Intégrandes normales et mesures paramétrées en calcul des variations, *Bull. Soc. Math. France* **101** (1973), 129–184.

[Bh 91] K. Bhattacharya, Wedge-like microstructure in martensites, *Acta metall.* **39** (1991), 2431–2444.

[Bh 92] K. Bhattacharya, Self-accommodation in martensite, *Arch. Rat. Mech. Anal.* **120** (1992), 201–244.

[Bh 93] K. Bhattacharya, Comparison of geometrically nonlinear and linear theories of martensitic transformation, *Cont. Mech. Thermodyn.* **5** (1993), 205–242.

[BFJK 94] K. Bhattacharya, N. Firoozye, R.D. James and R.V. Kohn, Restrictions on microstructure, *Proc. Roy. Soc. Edinburgh*, **124A** (1994), 843–878.

[BhJ 97] K. Bhattacharya and R.D. James, A theory of thin films of martensite materials with applications to microactuators, to appear in *J. Mech. Phys. Solids.*

[BK 96] K. Bhattacharya and R.V. Kohn, Symmetry, texture and the recoverable strain of shape-memory polycrystals, *Acta mater.* **44** (1996), 529–542.

[BK 97] K. Bhattacharya and R.V. Kohn, Elastic energy minimization and the recoverable strains of polycristalline shape memory materials, *Arch. Rat. Mech. Anal.* **139** (1997), 99–180.

[BP 61] E. Bishop and R. R. Phelps, A proof that all Banach spaces are subreflexive, *Bull. Amer. Math. Soc.* **67** (1961), 97–98.

[Bo 57] B. Bojarski, Generalized solutions to a system of first order differential equations of elliptic type with discontinuous coefficients (in Russian), *Mat. Sb.* **43** (1957), 451–503.

[Bo 90] G. Bouchitté, Singular perturbations of variational problems arising from a two phases transition model, *J. Appl. Math. Opt.* **21** (1990), 289–314.

[BFM 97] G. Bouchitté, I. Fonseca and L. Mascarenhas, *A global method for relaxation*, to appear in *Arch. Rat. Mech. Anal.*

[BM 54] J.S. Bowles and J.K. Mackanzie, The crystallography of martensite transformations, I and II, *Acta Met.* **2** (1954), 129–137, 138–147.

[Br 94] M. Brokate et al., *Phase transitions and hysteresis*, LNM **1584**, Springer, 1994.

[BS 96] M. Brokate and J. Sprekels, *Hysteresis and phase transitions*, Springer, 1996.

[Br 63] W.F. Brown, *Micromagnetics*, Wiley, 1963.

[Br 66] W.F. Brown, *Magnetoelastic interactions*, Springer Tracts in Natural Philosophy **9** (C. Truesdell, ed.), Springer, 1966.

[BRL 97] O. Bruno, F. Reitich and P. Leo, The overall elastic energy of polycrystalline martensite solids, submitted to *J. Mech. Phys. Solids.*

[CP 97] C. Carstensen and P. Plecháč, Numerical solution of the scalar double-well problem allowing microstructure, *Math. Comp.* **66** (1997), 997–1026.

[CT 93] E. Casadio-Tarabusi, An algebraic characterization of quasiconvex functions, *Ricerche Mat.* **42** (1993), 11–24.

[CPe 97] P. Celada and S. Perrotta, Functions with prescribed singular values of the gradient, preprint SISSA 56/97/M, 1997.

[CP 95] A. Cellina and S. Perrotta, On a problem of potential wells, *J. Convex Anal.* **2** (1995), 103–115.

[CK 97a] A. Cherkaev and R.V. Kohn (eds.), *Topics in the mathematical modelling of composites*, Birkhäuser, 1997.

[CK 88] M. Chipot and D. Kinderlehrer, Equilibrium configurations of crystals, *Arch. Rat. Mech. Anal.* **103** (1988), 237–277.

[CK 97b] R. Choksi and R.V. Kohn, Bounds on the micromagnetic energy of a uniaxial ferromagnet, preprint.

[CKO 97] R. Choksi, R.V. Kohn and F. Otto, in preparation.

[Ch 93] C. Chu, *Hysteresis and microstructure: a study of biaxial loading on compound twins of copper-aluminium-nickel single crystals*, Ph.D. thesis, University of Minnesota, 1993.

[CJ 95] C. Chu and R.D. James, Analysis of microstructures in Cu-14% Al-3.9% Ni by energy minimization, in: Proc. ICOMAT-95, *J. de physique* IV **5**, colloque C8, 143–149.

[Da 81] B. Dacorogna, A relaxation theorem and its applications to the equilibrium of gases, *Arch. Rat. Mech. Anal.* **77** (1981), 359–386.

[Da 82a] B. Dacorogna, Quasiconvexity and the relaxation of non convex variational problems, *J. Funct. Anal.* **46** (1982), 102–118.

[Da 82b] B. Dacorogna, Weak continuity and weak lower semicontinuity of non-linear functionals, Springer LNM **922**, Springer, 1982.

[Da 89] B. Dacorogna, *Direct methods in the calculus of variations*, Springer, 1989.

[DM 88] B. Dacorogna and P. Marcellini, A counterexample in the vectorial calculus of variations, in: *Material instabilities in continuum mechanics and related mathematical problems* (J.M. Ball, ed.), Oxford Univ. Press, 1988, 77–83.

[DM 96a] B. Dacorogna and P. Marcellini, Théorèmes d'existence dans les cas scalaire et vectoriel pour les équations de Hamilton-Jacobi, *C.R.A.S. Paris* **322** (1996), Serie I, 237–240.

[DM 96b] B. Dacorogna and P. Marcellini, Sur le problème de Cauchy-Dirichlet pour les systèmes d'équations non linéaires du premier ordre, *C.R.A.S. Paris* **323** (1996), Serie I, 599–602.

[DM 97] B. Dacorogna and P. Marcellini, General existence theorems for Hamilton-Jacobi equations in the scalar and vectorial cases, *Acta Math.* **178** (1997), 1–37.

[DM 93] G. Dal Maso, *An introduction to Γ-convergence*, Birkhäuser, 1993.

[De 96] S. Demoulini, Young measure solutions for a nonlinear parabolic equation of forward-backward type, *SIAM J. Math. Anal.* **27** (1996), 378–403.

[DG 75] E. De Giorgi, Sulla convergenza di alcune successioni di integrali del tipo dell'area, *Rend. Mat.* **8** (1975), 277–294.

[DGF 75] E. De Giorgi and T. Franzoni, Su un tipo di convergenze variationale, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **58** (1975), 842–850.

[DS 93] A. DeSimone, Energy minimizers for large ferromagnetic bodies, *Arch. Rat. Mech. Anal.* **125** (1993), 99–143.

[DP 85] R.J. DiPerna, Compensated compactness and general systems of conservation laws, *Trans. Amer. Math. Soc.* **292** (1985), 383–420.

[DPM 87] R.J. DiPerna and A.J. Majda, Oscillations and concentrations in weak solutions of the incompressible fluid equations, *Comm. Math. Phys.* **108** (1987), 667–689.

[Do 97] G. Dolzmann, Numerical computations of rank-one convex envelopes, preprint.

[DM 95] G. Dolzmann and S. Müller, Microstructures with finite surface energy: the two-well problem, *Arch. Rat. Mech. Anal.* **132** (1995), 101–141.

[E 92] Weinan E, Homogenization of linear and nonlinear transport equations, *Comm. Pure Appl. Math.* **45** (1992), 301–326.

[Ed 65] R.E. Edwards, *Functional analysis*, Holt, Rinehart and Winston, 1965.

[Ek 79] I. Ekeland, Nonconvex minimization problems, *Bull. Amer. Math. Soc.* **1** (1979), 443–474.

[Er 75] J.L. Ericksen, Equilibrium of bars, *J. Elasticity* **5** (1975), 191–201.

[Er 77] J.L. Ericksen, Special topics in nonlinear elastostatics, in: *Advances in applied mechanics* **17** (C.-S. Yih, ed.), Academic Press, 1977, 189–244.

[Er 79] J.L. Ericksen, On the symmetry of deformable elastic crystals, *Arch. Rat. Mech. Anal.* **72** (1979), 1–13.

[Er 80] J.L. Ericksen, Some phase transitions in elastic crystals, *Arch. Rat. Mech. Anal.* **73** (1980), 99–124.

[Er 84] J.L. Ericksen, The Cauchy and Born hypothesis for crystals, in: *Phase transformations and material instabilities in solids* (M.E. Gurtin, ed.), Academic Press, 1984, 61–77.

[Er 89] J.L. Ericksen, Weak martensitic transformations in Bravais lattices, *Arch. Rat. Mech. Anal.* **107** (1989), 23–36.

[Es 57] J.D. Eshelby, The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proc. Roy. Soc. London* **A 241** (1957), 376–396.

[Es 59] J.D. Eshelby, The elastic field outside an ellipsoidal inclusion, *Proc. Roy. Soc. London* **A 252** (1959), 561–569.

[Es 61] J.D. Eshelby, Elastic inclusions and inhomogenities, in: *Progress in Solid Mechanics* vol. II, North-Holland, 1961, 87–140.

[Ev 86] L.C. Evans, Quasiconvexity and partial regularity in the calculus of variations, *Arch. Rat. Mech. Anal.* **95** (1986), 227–252.

[Ev 90] L.C. Evans, *Weak convergence methods for nonlinear partial differential equations*, Am. Math. Soc., Providence, 1990.

[EG 87] L.C. Evans and R. Gariepy, Blow-up, compactness and partial regularity in the calculus of variations, *Indiana Univ. Math. J.* **36** (1987), 361–371.

[EG 92] L.C. Evans and R. Gariepy, *Measure theory and fine properties of functions*, CRC Press, 1992.

[FF 60] H. Federer and W.H. Fleming, Normal and integral currents, *Ann. Math.* **72** (1960), 458–520.

[Fe 69] H. Federer, *Geometric measure theory*, Springer, 1969.

[Fo 87] I. Fonseca, Variational methods for elastic crystals, *Arch. Rat. Mech. Anal.* **97** (1987), 189–220.

[Fo 88] I. Fonseca, The lower quasiconvex envelope of the stored energy function for an elastic crystal, *J. Math. Pures Appl.* **67** (1988), 175–195.

[Fo 89] I. Fonseca, Phase transitions of elastic solid materials, *Arch. Rat. Mech. Anal.* **107** (1989), 195–223.

[Fo 96] I. Fonseca, Variational techniques for problems in materials science, in: *Variational methods for discontinuous structures*, (R. Serapioni and F. Tomarelli, eds.), Birkhäuser, 1996, 161–175.

[FBS 94] I. Fonseca, D. Brandon and P. Swart, Dynamics and oscillatory microstructure in a model of displacive phase transformations, in: *Progress in partial differential equations, the Metz surveys* **3** (M. Chipot, ed.), Pitman, 1994, 130–144.

[FM 97] I. Fonseca and S. Müller, A-quasiconvexity, lower semicontinuity and Young measures, in preparation.

[FMP 97] I. Fonseca, S. Müller and P. Pedregal, Analysis of concentration and oscillation effects generated by gradients, to appear in *SIAM J. Math. Anal.*

[FT 89] I. Fonseca and L. Tartar, The gradient theory of phase transitions for systems with two potential wells, *Proc. Roy. Soc. Edinburgh* **111 A** (1989), 89–102.

[FM 96] G. Friesecke and J.B. McLeod, Dynamics as a mechanism preventing the formation of finer and finer microstructure, *Arch. Rat. Mech. Anal.* **133** (1996), 199–247.

[FH 85] N. Fusco and J. Hutchinson, Partial regularity of functions minimizing quasiconvex integrals, *manuscripta math.* **54** (1985), 121–143.

[Ge 90] P. Gérard, Mesures semi-classiques et ondes de Bloch, in: Equations aux dérivées partielles, exposé XVI, seminaire 1990–91, Ecole Polytechnique, Palaiseau.

[Ge 91] P. Gérard, Microlocal defect measures, *Comm. PDE* **16** (1991), 1761–1794.

[GM 86] M. Giaquinta and G. Modica, Partial regularity of minimizers of quasiconvex integrals, *Ann. IHP Analyse non linéaire* **3** (1986), 185–208.

[GMS 89] M. Giaquinta, G. Modica and J. Souček, Cartesian currents, weak diffeomorphisms and nonlinear elasticity, *Arch. Rat. Mech. Anal.* **106** (1989), 97–159; erratum **109** (1990), 385–392.

[GMS 96] M. Giaquinta, G. Modica and J. Souček, *Cartesian currents in the calculus of variations, part I*, preprint, 1996.

[Gr 86] M. Gromov, *Partial differential relations*, Springer, 1986.

[Gu 83] M.E. Gurtin, Two-phase deformations in elastic solids, *Arch. Rat. Mech. Anal.* **84** (1983), 1–29.

[Gu 87] M.E. Gurtin, Some results and conjectures in the gradient theory of phase transitions, in: *Metastability and Incompletely Posed Problems* (S.S. Antman, ed.), Springer, 1987, 135–146.

[Ha 97] K. Hane, *Microstructures in thermoelastic martensites*, Ph.D. thesis, University of Minnesota, 1997.

[HR 94] K.-H. Hoffmann and T. Roubíček, Optimal control of a fine structure, *Appl. Math. Optim.* **30** (1994), 113–126.

[Hu 67] A. Hubert, Zur Theorie der zweiphasigen Domänenstrukturen in Supraleitern und Ferromagneten, *Phys. status solidi* **24** (1967), 669–682.

[Hu 88] F. Hüsseinov, Continuity of quasiconvex functions and theorem on quasiconvexification, *Izv. Akad. Nauk Azerbaidzhan SSR Ser. Fiz.-Tekhn. Mat. Nauk* **8** (1988), 17–23.

[Hu 95] F. Hüsseinov, Weierstrass condition for the general basic variational problem, *Proc. Roy. Soc. Edinburgh* **125 A** (1995), 801–806.

[HM 93] Y. Huo and I. Müller, Nonequilibrium thermodynamics and pseudoelasticity, *Cont. Mech. Thermodyn.* **5** (1993), 163–204.

[IT 69] A. & C. Ionescu Tulcea, *Topics in the theory of liftings*, Springer, 1969.

[Ja 81] R.D. James, Finite deformation by mechanical twinning, *Arch. Rat. Mech. Anal.* **77** (1981), 143–176.

[JK 89] R.D. James and D. Kinderlehrer, Theory of diffusionless phase transitions, in: *PDEs and continuum models of phase transitions* (M. Rascle, D. Serre and M. Slemrod, eds.), Lecture Notes in Physics **344**, Springer, 1989.

[JK 90] R.D. James and D. Kinderlehrer, Frustration in ferromagnetic materials, *Cont. Mech. Thermodyn.* **2** (1990), 215–239.

[JK 93] R.D. James and D. Kinderlehrer, Theory of magnetostriction with applications to $Tb_xDy_{1-x}Fe_2$, *Phil. Mag. B* **68** (1993), 237–274.

[JW 97] R.D. James and M. Wuttig, Magnetostriction of martensite, to appear in *Phil. Mag. A*.

[JO 90] M. Jodeit and P.J. Olver, On the equation $\mathrm{grad} f = M\,\mathrm{grad} g$, *Proc. Roy. Soc. Edinburgh* **116 A** (1990), 341–358.

[Jo 83] J.L. Joly, Sur la propagation des oscillations per un système hyperbolique en dimension 1, *C.R.A.S. Paris* **296** (1983), 669–672.

[JMR 95] J. L. Joly, G. Métivier, J. Rauch, Trilinear compensated compactness and nonlinear geometric optics, *Ann. Math.* **142** (1995), 121–169.

[Kh 67] A. Khachaturyan, Some questions concerning the theory of phase transformations in solids, *Soviet Physics-Solid State* **8** (1967), 2163–2168.

[Kh 83] A. Khachaturyan, *Theory of structural transformations in solids*, Wiley, 1983.

[KSh 69] A. Khachaturyan and G. Shatalov, Theory of macroscopic periodicity for a phase transition in the solid state, *Soviet Physics JETP* **29** (1969), 557–561.

[Ki 88] D. Kinderlehrer, Remarks about equilibrium configurations of crystals, in: *Material instabilities in continuum mechanics and related mathematical problems* (J.M. Ball, ed.), Oxford Univ. Press, 1988, 217–242.

[Kin 97] D. Kinderlehrer, Metastability and hysteresis in active materials, submitted to: *Mathematics and control in smart structures* (V.K. Vardan and J. Chandra, eds.), 1997.

[KP 91] D. Kinderlehrer and P. Pedregal, Characterization of Young measures generated by gradients, *Arch. Rat. Mech. Anal.* **115** (1991), 329–365.

[KP 94] D. Kinderlehrer and P. Pedregal, Gradient Young measures generated by sequences in Sobolev spaces, *J. Geom. Analysis* **4** (1994), 59–90.

[Ki 97] B. Kirchheim, in preparation.

[KL 71] B. Klosowicz and K.A. Lurie, On the optimal nonhomogeneity of a torsional elastic bar, *Arch. Mech.* **24** (1971), 239–249.

[Ko 89] R.V. Kohn, The relationship between linear and nonlinear variational models of coherent phase transitions, in: *Trans. 7th Army Conf. on applied mathematics and computing*, (F. Dressel, ed.), 1989.

[KM 92] R.V. Kohn and S. Müller, Branching of twins near an austenite/twinned-martensite interface, *Phil. Mag.* **A 66** (1992), 697–715.

[KM 94] R.V. Kohn and S. Müller, Surface energy and microstructure in coherent phase transitions, *Comm. Pure Appl. Math.* **47** (1994), 405–435.

[KS 89] R.V Kohn and P. Sternberg, Local minimisers and singular perturbations, *Proc. Roy. Soc. Edinburgh* **111A** (1989), 69–84.

[KP 89] M.A. Krasnosel'skiĭ and A.V. Pokrovskiĭ, *Systems with hysteresis*, Springer, 1989.

[Kr 94] J. Kristensen, Finite functionals and Young measures generated by gradients of Sobolev functions, Ph.D. Thesis, Technical University of Denmark, Lyngby.

[Kr 97a] J. Kristensen, On the non-locality of quasiconvexity, to appear in *Ann. IHP Anal. non linéaire.*

[Kr 97b] J. Kristensen, Lower semicontinuity in spaces of weakly differentiable functions, preprint.

[Kr 97] M. Kružik, On the composition of quasiconvex functions and the transposition, preprint 1997.

[Ku 55] N.H. Kuiper, On C^1 isometric embeddings, I., *Nederl. Akad. Wetensch. Proc.* **A 58** (1955), 545–556.

[Li 44] E. Lifshitz, On the magnetic structure of iron, *J. Phys.* **8** (1944), 337–346.

[Li 77] F.-C. Liu, A Lusin type property of Sobolev functions, *Indiana Univ. Math. J.* **26** (1977), 645–651.

[LP 93] P.-L. Lions and T. Paul, Sur les mesures de Wigner, *Rev. Mat. Iberoamericana* **9** (1993), 553–618.

[Lu 96] M. Luskin, On the computation of crystalline microstructure, *Acta Num.* **5** (1996), 191–258.

[MP 86] J. H. Maddocks and G. Parry, A model for twinning, *J. Elasticity* **16** (1986), 113–133.

[Ma 94] J. Malý, Lower semicontinuity of quasiconvex integrals, *manuscripta math.* **85** (1994), 419–428.

[Ma 85] P. Marcellini, Approximation of quasiconvex functions, and lower semicontinuity of multiple integrals, *manuscripta math.* **51** (1985), 1-28.

[Ma 92] J.P. Matos, Young measures and the absence of fine microstructures in a class of phase transitions, *Europ. J. Appl. Math.* **3** (1992), 31–54.

[MP 98] J. Matoušek and P. Plecháč, On functional separately convex hulls, *J. Discrete Comput. Geom.* **19** (1998), 105–130.

[MPT 85] D.W. McLaughlin, G. Papanicolaou and L. Tartar, Weak limits of semilinear hyperbolic sytems with oscillating data, in: *Macroscopic modelling of turbulent flows*, Lecture Notes in physics **230**, Springer, 1985, 277–289.

[MS 40] E.J. McShane, Generalized curves, *Duke Math. J.* **6** (1940), 513–516.

[MS 89] E.J. McShane, The calculus of variations from the beginning through optimal control theory, *SIAM J. Control Opt.* **27** (1989), 916–939.

[Me 66] P.-A. Meyer, *Probability and potentials*, Blaisdell, 1966.

[Me 65] N. G. Meyers, Quasiconvexity and lower semicontinuity of multiple variational integrals of any order, *Trans. Amer. Math. Soc.* **119** (1965), 125–149.

[Mi 97] A. Mielke, Evolution equations for Young-measure solutions of semilinear hyperbolic problems, preprint, 1997.

[Mi 98] G. Milton, *The effective tensors of composites*, book in preparation.

[Mo 87] L. Modica, Gradient theory of phase transitions and minimal interface criterion, *Arch. Rat. Mech. Anal.* **98** (1987), 123–142.

[MM 77a] L. Modica and S. Mortola, Il limite nella Γ-convergenza di una famiglia di funzionali ellittici, *Boll. U.M.I* **14**-A (1977), 525–529.

[MM 77b] L. Modica and S. Mortola, Un esempio di Γ^--convergenza, *Boll. U.M.I* **14**-B (1977), 285–299.

[Mo 52] C.B. Morrey, Quasi-convexity and the lower semicontinuity of multiple integrals, *Pacific J. Math.* **2** (1952), 25–53.

[Mo 66] C.B. Morrey, *Multiple integrals in the calculus of variations*, Springer, 1966.

[Mu 93] S. Müller, Singular perturbations as a selection criterion for periodic minimizing sequences, *Calc. Var.* **1** (1993), 169–204.

[Mu 97a] S. Müller, *Microstructure and the calculus of variations*, Nachdiplomvorlesung ETH Zürich, in preparation.

[Mu 97b] S. Müller, A sharp version of Zhang's theorem on truncating sequences of gradients, to appear in *Trans. AMS.*

[Mu 97c] S. Müller, Microstructures, phase transitions and geometry, in: *Proc. ECM2, Budapest, 1996*, Birkhäuser, to appear.

[MS 96] S. Müller and V. Šverák, Attainment results for the two-well problem by convex integration, *Geometric analysis and the calculus of variations*, (J. Jost, ed.), International Press, 1996, 239–251.

[MS 97] S. Müller and V. Šverák, in preparation.

[MS 98] S. Müller and V. Šverák, in preparation.

[Mu 78] F. Murat, H-convergence, mimeographed notes, 1978, based partially on the Cours Peccot by L. Tartar, 1977; translation (with authors F. Murat and L. Tartar) in [CK 97a], 21–43.

[MT 97] F. Murat and L. Tartar, On the relation between Young measures and H-measures, in preparation.

[Na 54] J. Nash, C^1 isometric embeddings, *Ann. Math.* **60** (1954), 383–396.

[Ng 89] G. Nguetseng, A general convergence result for a functional related to the theory of homogenization, *SIAM J. Math. Anal.* **20** (1989), 608–623.

[Ot 95] F. Otto, Evolution of microstructure in unstable porous media flow: a relaxational approach, preprint SFB 256, Bonn, 1995.

[Ot 98] F. Otto, Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory, *Arch. Rat. Mech. Anal.* **141** (1998), 63–103.

[Pa 77] G. Parry, On the crystallographic point groups and Cauchy symmetry, *Math. Proc. Cambridge Phil. Soc.* **82** (1977), 165–175.

[Pa 81] G. Parry, On phase transitions involving internal strain, *Int. J. Solids Structures* **17** (1981), 361–378.

[Pa 87a] G. Parry, On internal variable models of phase transitions, *J. Elasticity* **17** (1987), 63–70.

[Pa 87b] G. Parry, On shear bands in unloaded crystals, *J. Mech. Phys. Solids* **35** (1987), 357–382.

[Pa 89] G. Parry, Stable phase boundaries in unloaded crystals, *Cont. Mech. Thermodyn.* **1** (1989), 305–314.

[Pe 93] P. Pedregal, Laminates and microstructure, *Europ. J. Appl. Math.* **4** (1993), 121–149.

[Pe 97] P. Pedregal, *Parametrized measures and variational principles*, Birkhäuser, 1997.

[Pe 87] R.L. Pego, Phase transitions in one-dimensional nonlinear viscoelasticity: admissibility and stability, *Arch. Rat. Mech. Anal.* **97** (1987), 353–394.

[Pi 84] M. Pitteri, Reconciliation of local and global symmetries of crystals, *J. Elasticity* **14** (1984), 175–190.

[PZ 97] M. Pitteri and G. Zanzotto, *Continuum models for phase transitions and twinning in crystals*, Chapman and Hall, forthcoming.

[Pr 76] I. Privorotskii, *Thermodynamic theory of domain structures*, Wiley, 1976.

[Re 67] Yu.G. Reshetnyak, On the stability of conformal mappings in multidimensional spaces, *Sib. Math. J.* **8** (1967), 69–85.

[Re 68a] Yu.G. Reshetnyak, Liouville's theorem under minimal regularity assumptions, *Sib. Math. J.* **8** (1968), 631–634.

[Re 68b] Yu.G. Reshetnyak, Weak convergence of completely additive vector functions on a set, *Sib. Math. J.* **9** (1968), 1039–1045.

[Re 89] Yu.G. Reshetnyak, *Space mapping with bounded distorsion*, Am. Math. Soc., 1989.

[Ro 69] A. Roitburd, The domain structure of crystals formed in the solid phase, *Soviet Physics-Solid State* **10** (1969), 2870–2876.

[Ro 78] A. Roitburd, Martensitic transformation as a typical phase transformation in solids, *Solid State Physics* **34** (1978), 317–390.

[Ro 97] T. Roubíček, *Relaxation in optimization theory and variational calculus*, de Gruyter, 1997.

[Ru 73] W. Rudin, *Functional analysis*, McGraw-Hill, 1973.

[Sch 94] C. Schreiber, Rapport de stage de D.E.A., ENS Lyon, 1994.

[Sch 93] D. Schryvers, Microtwin sequences in thermoelastic Ni_xAl_{100-x} martensite studied by conventional and high resolution transmission electron microscopy, *Phil. Mag.* **A 68** (1993), 1017–1032.

[Se 83] D. Serre, Formes quadratique et calcul de variations, *J. Math. Pures Appl.* **62** (1983), 177–196.

[Sv 91a] V. Šverák, Quasiconvex functions with subquadratic growth, *Proc. Roy. Soc. London* **A 433** (1991), 723–725.

[Sv 91b] V. Šverák, On regularity for the Monge-Ampère equations, preprint, Heriot-Watt University, 1991.

[Sv 92a] V. Šverák, Rank-one convexity does not imply quasiconvexity, *Proc. Roy. Soc. Edinburgh* **120** (1992), 185–189.

[Sv 92b] V. Šverák, New examples of quasiconvex functions, *Arch. Rat. Mech. Anal.* **119** (1992), 293–300.

[Sv 93a] V. Šverák, On the problem of two wells, in: *Microstructure and phase transitions*, IMA Vol. Appl. Math. **54** (D. Kinderlehrer, R.D. James, M. Luskin and J. Ericksen, eds.), Springer, 1993, 183–189.

[Sv 93b] V. Šverák, On Tartar's conjecture, *Ann. IHP Analyse non linéaire* **10** (1993), 405–412.

[Sv 95] V. Šverák, Lower semicontinuity of variational integrals and compensated compactness, in: Proc. ICM 1994 (S.D. Chatterji, ed.), vol. 2, Birkhäuser, 1995, 1153–1158.

[Sv 98] V. Šverák, in preparation.

[Sv] V. Šverák, personal communication.

[Sy 97] M.A. Sychev, A new approach to Young measure theory, relaxation and convergence in energy, preprint.

[Ta 75] L. Tartar, Problèmes de contrôle des coefficients dans des équations aux derivées partielles, in: *Control theory, numerical methods and computer systems modelling* (A. Bensoussan and J.-L. Lions, eds.), Springer, 1975, 420–426; translated (with authors F. Murat and L. Tartar) in [CK 97a], 1–8.

[Ta 79a] L. Tartar, Compensated compactness and partial differential equations, in: *Nonlinear Analysis and Mechanics: Heriot-Watt Symposium* Vol. IV, (R. Knops, ed.), Pitman, 1979, 136–212.

[Ta 79b] L. Tartar, Estimations de coefficients homogénéises, in: *Computing methods in applied sciences and engineering*, LMN **704**, Springer, 1979.

[Ta 80] L. Tartar, Some existence theorems for semilinear hyperbolic sytems in one space variable, report # 2164, Mathematics Research Center, University of Wisconsin, 1980.

[Ta 81] L. Tartar, Solutions oscillantes des équation de Carleman, seminar Goulaouic-Meyer-Schwartz 1980/81, exp. No. XII, Ecole Polytechnique, Palaiseau, 1981.

[Ta 83] L. Tartar, The compensated compactness method applied to systems of conservations laws, in: *Systems of Nonlinear Partial Differential Equations*, (J.M. Ball, ed.), NATO ASI Series, Vol. C111, Reidel, 1983, 263–285.

[Ta 84] L. Tartar, Etude des oscillations dans les équations aux derivées partielles non linéaires, in: *Trends and applications of pure mathematics to mechanics*, Lecture Notes in physics **195**, Springer, 1984, 384–412.

[Ta 86] L. Tartar, Oscillations in nonlinear partial differential equations: compensated compactness and homogenization, in:

Nonlinear systems of partial differential equations in applied mathematics, Amer. Math. Soc., 1986, 243–266.

[Ta 87] L. Tartar, Oscillations and asymptotic behaviour for two semilinear hyperbolic systems, in: *Dynamics of infinite-dimensional dynamical systems*, NATO ASI Ser. F, Springer, 1987.

[Ta 90] L. Tartar, H-measures, a new approach for studying homogenization, oscillations and concentration effects in partial differential equations, *Proc. Roy. Soc. Edinburgh* **A 115** (1990), 193–230.

[Ta 93] L. Tartar, Some remarks on separately convex functions, in: *Microstructure and phase transitions*, IMA Vol. Math. Appl. **54**, (D. Kinderlehrer, R.D. James, M. Luskin and J.L. Ericksen, eds.), Springer, 1993, 191–204.

[Ta 95] L. Tartar, Beyond Young measures, *Meccanica* **30** (1995), 505–526.

[Ta 98] L. Tartar, *Homogenization, compensated compactness and H-measures*, CBMS-NSF conference, Santa Cruz, June 1993, lecture notes in preparation.

[Th 97] F. Theil, Young measure solutions for a viscoelastically damped wave equation with nonmonotone stress-strain relation, to appear in *Arch. Rat. Mech. Anal.*.

[Va 90] M. Valadier, Young measures, in: *Methods of nonconvex analysis*, LNM **1446**, Springer, 1990.

[Va 94] M. Valadier, A course on Young measures, *Rend. Istit. Mat. Univ. Trieste* **26** (1994) suppl., 349–394.

[Vi 94] A. Visintin, *Differential models of hysteresis*, Springer, 1994.

[WLR 53] M. S. Wechsler, D. S. Liebermann, T. A. Read, On the theory of the formation of martensite, *Trans. AIME J. Metals* **197** (1953), 1503–1515.

[Yo 37] L.C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *Comptes Rendues de la Société des Sciences et des Lettres de Varsovie, classe III* **30** (1937), 212–234.

[Yo 69] L.C. Young, *Lectures on the calculus of variations and optimal control theory*, Saunders, 1969 (reprinted by Chelsea, 1980).

[Za 92] G. Zanzotto, On the material symmetry group of elastic crystals and the Born rule, *Arch. Rat. Mech. Anal.* **121** (1992), 1–36.

[Zh 92] K. Zhang, A construction of quasiconvex functions with linear growth at infinity, *Ann. S.N.S. Pisa* **19** (1992), 313–326.

[Zh] K. Zhang, Rank-one connections and the three-well problem, preprint.