Quasiconvex functions incorporating volumetric constraints are rank-one convex

June 15, 2007

Sergio Conti

Fachbereich Mathematik, Universität Duisburg-Essen
Lotharstr. 65, 47057 Duisburg, Germany

We prove that a quasiconvex function $W : \mathbb{M}^{n\times n} \to [0,\infty]$ which is finite on the set $\Sigma = \{F : \det F = 1\}$ is rank-one convex, and hence continuous, on Σ; and the same for constraints on minors. This implies that the rank-one convex envelope gives an upper bound on the quasiconvex envelope of any energy density modeling an incompressible material. Our result is based on the construction of an appropriate piecewise affine function u such that $\nabla u \in \Sigma$ almost everywhere.

1 Introduction

Quasiconvexity is a central notion in the vectorial calculus of variations, introduced by Morrey in 1952 [18]. Under suitable continuity and growth assumptions on the energy density $W : \mathbb{M}^{m\times n} \to \mathbb{R}$, the functional

$$u \mapsto \int_\Omega W(\nabla u) dx \tag{1.1}$$

is weakly lower semicontinuous in $W^{1,p}(\Omega;\mathbb{R}^m)$ if and only if W is quasiconvex. In parallel, in the theory of relaxation one is typically confronted with the issue of determining the quasiconvex envelope of the energy density W, i.e., the largest quasiconvex function below W. For a review of these concepts see, e.g., [7, 19, 3, 10].

A function $W : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ is said to be quasiconvex if

$$W(F) \le \frac{1}{|\Omega|} \int_\Omega W(F + \nabla\varphi)\, \mathrm{d}x \quad \text{for all } \varphi \in W_0^{1,\infty}(\Omega;\mathbb{R}^m) \tag{1.2}$$

(provided the integral exists) for all bounded, open, nonempty sets $\Omega \subset \mathbb{R}^n$ such that $|\partial\Omega| = 0$ [18, 5, 11, 19]. This definition involves minimizing an integral functional over the space of all Lipschitz functions, and is therefore implicit and difficult to handle directly. The search for more explicit conditions lead to the introduction of the related concepts of polyconvexity and rank-one convexity [18, 2]. A function $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ is polyconvex if it can be written as a convex function of F, its determinant, and its minors. All polyconvex functions are quasiconvex, since the determinant and the minors are null Lagrangians [2]. One says that a function $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ is rank-one convex if it is convex in all rank-one directions, i.e., if all one-dimensional restrictions $t \mapsto f(F + ta \otimes b)$ are convex. This is a local condition; for C^2 functions it can be directly written in terms of their second gradient (and of the distributional one in the general case). For finite-valued maps, i.e., functions $f : \mathbb{M}^{m\times n} \to \mathbb{R}$, rank-one convexity, much as ordinary convexity, implies that f is locally Lipschitz continuous. It is well known that finite-valued quasiconvex functions are rank-one convex (and in particular continuous). For higher regularity of quasiconvex envelopes see [4].

Treating extended-valued functions, i.e., maps $W : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$, is substiantially more complex. These functions are however often encountered in models from nonlinear elasticity. For example in the study of elastomeric materials, which are modeled as incompressible, one deals with functions which incorporate a constraint on the determinant, like $W(F) = \infty$ whenever $\det F \neq 1$. In this case it is still true that polyconvexity implies quasiconvexity, but the usual argument relating quasiconvexity to rank-one convexity fails. Indeed, quasiconvexity does not, in general, imply rank-one convexiy, as can be seen by the example provided by Ball and Murat [5],

$$W(F) = \begin{cases} 0 & \text{if } F = \pm e_1 \otimes e_1\,, \\ \infty & \text{else.} \end{cases} \tag{1.3}$$

We show here that for functions which are finite on the constant-determinant surface quasiconvexity implies rank-one convexity.

Theorem 1.1. *Let $W : \mathbb{M}^{n\times n} \to \mathbb{R} \cup \{\infty\}$ be quasiconvex, and*

$$\Sigma = \left\{F \in \mathbb{M}^{n\times n} : \det F = 1\right\}\,.$$

If $W(F) < \infty$ for all $F \in \Sigma$, then W is rank-one convex, and hence continuous, on the surface Σ.

The same holds in the case that the constraint is on a minor, and that W is finite on a relatively open subset of Σ, see Theorem 3.1 and Remark 3.2 below.

This result has several applications in the theory of relaxation, where one is interested in determining the quasiconvex envelope of W, i.e., the highest quasiconvex function below W. For finite-valued functions one typically determines a function f which is at the same time an upper bound on the rank-one convex envelope of W, and a lower bound on the polyconvex envelope of W. Then, since polyconvexity implies quasiconvexity which in turn implies rank-one convexity, all the three envelopes necessarily coincide with f. Theorem 1.1 permits to extend the same method to functions incorporating volumetric constraints. This leads to considerable simplifications in the proof of relaxation results with volumetric constraints, as for example those obtained for nematic and smectic elastomers and for models in plasticity [9, 1, 12]. In the mentioned papers the relation between rank-one convexity and quasiconvexity has been proven on a case-by-case basis using the convex integration result by Müller and Šverák discussed below [22]. Generalizing the argument used in those works it is possible to show that, if the function W is quasiconvex and is *assumed* to be continuous of Σ, then it is rank-one convex (even if, to the best of my knowledge, this general fact is never explicitly stated in the literature). The present argument instead permits to *prove* continuity on Σ, as well as rank-one convexity.

Convex integration is a general strategy to prove existence of solutions, which is useful for cases where the functional is not lower semicontinuous. This approach, first developed by Nash, Kuiper and later by Gromov in the context of differential geometry [24, 17, 13], was extended by Müller and Šverák to Lipschitz solutions [21, 23], including in particular the case that a volumetric constraint is present [22]. They first constructed smooth test functions obeying the constraint, and then modified them so that they became affine on each of infinitely many pieces. The modification was done by an iterative argument, which in each step restored the volumetric constraint exploiting a result by Dacorogna and Moser, which shows that for any Hölder regular f with average one there are diffeomorphisms $u : \Omega \to \Omega$ such that $\det \nabla u = f$ and $u(x) = x$ on $\partial\Omega$ [8]. The construction presented in Theorem 2.1 below refines their result, and in particular Theorem 6.1 of [22], by presenting a completely elementary explicit piecewise affine construction, such that the gradient takes finitely many values, which directly satisfy the constraint. This simplifies not only the argument, avoiding the need to pass through the Dacorogna-Moser diffeomorphism, but also the result, producing functions with a simpler structure (see discussion after Theorem 2.1 below). In particular, the present test functions permit to prove that, if the function W is quasiconvex and finite on the unit-determinant constraint, then it is continuous. The existence of a piecewise affine construction for the two-dimensional case was first noticed by Müller and Šverák [22, Remark 2 after Th. 6.1]; their idea was then worked out in detail in [6, Lemma A.2]. We give below a simplification of the latter proof, and show how it can be extended to higher dimension and to constraints on minors. In the two-dimensional case, a different construction fulfilling the determinant constraint was recently obtained by Kirchheim, which uses infinitely many gradients, and with the additional property that for boundary data of the form $C = \lambda A + (1-\lambda)B$, the gradient takes values in the set $\{A, B\} \cup B_\varepsilon(C)$ [14]. The latter construction has the advantage that, combining with the convex integration techniques by Müller and Šverák, piecewise affine solutions of the partial differential inclusion can be obtained.

Notation. We identify $\mathbb{R}^n$ with the subset of $\mathbb{R}^m$, $m > n$, such that the last $m-n$ components are zero, and denote by e_i the canonical basis. Analogously for $\mathbb{M}^{m\times n}$ and $\mathbb{M}^{m'\times n'}$. We denote by $\mathrm{Id}_r = \sum_{i=1}^r e_i \otimes e_i$ the $r \times r$ identity matrix; $\mathrm{Id}_r \in \mathbb{M}^{m\times n}$ whenever $n, m \geq r$. In particular Id_r is the projection on the first r components in $\mathbb{R}^n$, $n \geq r$. Finally, $\det_r : \mathbb{M}^{m\times n} \to \mathbb{R}$, for $n, m \geq r$, denotes the determinant of the first $r \times r$ block; $|A| = (\sum_{ij} A_{ij}^2)^{1/2}$ the Euclidean norm of a matrix (or a vector).

2 Piecewise affine constructions

We present here our basic construction. The main result of this section is the following.

Theorem 2.1. *Let $n, m \in \mathbb{N}$, $2 \leq r \leq \min\{n, m\}$, and A, $B \in \mathbb{M}^{m\times n}$, with* $\mathrm{rank}(A-B) = 1$. *Assume that there are matrices $P \in \mathbb{M}^{m\times m}$ and $Q \in \mathbb{M}^{n\times n}$ such that*

$$\det_r PAQ = \det_r PBQ = t\,,$$

for some $t \neq 0$.

If $(A-B)Q\mathrm{Id}_r \neq 0$, then for any $\lambda \in [0,1]$ and any $\varepsilon > 0$ there is a finite set

$$K \subset \left\{F \in \mathbb{M}^{m\times n} : \det_r PFQ = t\,,\ \min\{|F-A|, |F-B|\} < \varepsilon\right\} \qquad (2.1)$$

with the following property: For any $\delta > 0$ and any open set $\Omega \subset \mathbb{R}^n$ there is a function $u \in W^{1,\infty}(\Omega;\mathbb{R}^m)$ such that

(i) $u(x) = (\lambda A + (1-\lambda)B)x$ *on* $\partial\Omega$;

(ii) $|u(x) - (\lambda A + (1-\lambda)B)x| \leq \delta$ *a.e. on* Ω;

(iii) $\nabla u \in K$ *a.e. on* Ω;

(iv) $|\{x \in \Omega : \nabla u(x) \notin \{A, B\}\}| \leq \delta|\Omega|$;

(v) There is a polyhedron $\widetilde{\Omega} \subset \mathbb{R}^n$, depending on ε and δ, such that if $\Omega = \widetilde{\Omega}$ the function u can be taken affine on each of finitely many symplexes covering $\widetilde{\Omega}$;

(vi) $\#K \leq 10 \cdot 4^{n-2}$.

If instead $(A-B)Q\mathrm{Id}_r = 0$ (and the other assumptions still hold) then for any $\varepsilon > 0$, $\lambda \in [0,1]$ and $\delta > 0$ there is a set K with $\#K \leq 30 \cdot 4^{n-2}$ obeying (2.1) and such that for any open set $\Omega \subset \mathbb{R}^n$ there is a function $u \in W^{1,\infty}(\Omega;\mathbb{R}^m)$ obeying (i)-(v).

Notice that in the typical case $m = n = r$ the condition $(A-B)Q\mathrm{Id}_r \neq 0$ always holds. Only if $r < n$ the exceptional case $(A-B)Q\mathrm{Id}_r = 0$ can appear; in that situation the set K depends also on δ.

For a comparison, we recall that Theorem 6.1 of [22] gave under the same assumptions a construction for the case that

$$K = \left\{F \in \mathbb{M}^{m\times n} : \det_r PFQ = t\,,\ \mathrm{dist}(F, [A,B]) < \varepsilon\right\},$$

i.e., for an infinite set, and with matrices close to the entire segment $[A,B] = \{sA + (1-sB) : s \in [0,1]\}$. In that case the distinction between ε and δ became irrelevant, and points (v), (vi) of course did not apply. Here we keep the two parameters ε and δ distinct, since in proving Theorem 3.1 below we shall take the limit $\delta \to 0$ keeping the finite set K fixed.

The proof is separated in several steps. We first focus on a special case (i.e., on special A, B, P, Q, Ω), on a special domain Ω, and ignore points (iv) and (ii). For this case, we give an explicit construction in two dimensions, and extend it inductively to higher dimension. In Section 2.3 we then modify the construction to fulfill point (iv) as well, introducing a large region where the function coincides with a simple laminate. In Section 2.4 we then show how the general case can be reduced to the mentioned special case.

2.1 Construction for special matrices

Let $n \geq 2$ and $\lambda \in (0,1)$. In this and the next section we consider the special case

$$A = \mathrm{Id}_n + (1-\lambda) e_2 \otimes e_1 \,, \qquad \text{and} \qquad B = \mathrm{Id}_n - \lambda e_2 \otimes e_1 \,, \tag{2.2}$$

which obey $\lambda A + (1-\lambda) B = \mathrm{Id}_n$, $\det_r A = \det_r B = 1$ for all r, $1 \leq r \leq n$. For example, for $n = 3$ we have

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1-\lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ -\lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Given a number $h \in (0,\infty)$ we define the "single laminate" $v^L : \mathbb{R}^n \to \mathbb{R}^n$ by

$$v^L(x) = x + \begin{cases} (1-\lambda) x_1 e_2 & \text{if } |x_1| \leq \lambda h \,, \\ \lambda (h - x_1) e_2 & \text{if } \lambda h < x_1 < h \,, \\ \lambda(-h - x_1) e_2 & \text{if } -h < x_1 < -\lambda h \,, \\ 0 & \text{else} \end{cases} \tag{2.3}$$

(see Figure 1). It is clear that $v^L \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^n)$ and

$$\nabla v^L(x) = \begin{cases} A & \text{if } |x_1| \leq \lambda h, \\ B & \text{if } \lambda h < |x_1| \leq h, \\ \mathrm{Id}_n & \text{else.} \end{cases}$$

Lemma 2.2. *Let $\lambda \in (0,1)$, $\varepsilon > 0$, $n \geq 2$, A and $B \in \mathbb{M}^{n\times n}$ as in (2.2). Then there is a finite set*

$$K \subset \big\{ F \in \mathbb{M}^{n\times n} : \det_r F = 1 \textit{ for all } 2 \leq r \leq n \,, \\ \min\{|F-A|, |F-B|\} < \varepsilon \big\} \tag{2.4}$$

such that one can find a polyhedron $\Omega \subset \mathbb{R}^n$ and a map $u \in W^{1,\infty}(\Omega;\mathbb{R}^n)$, with the following properties:

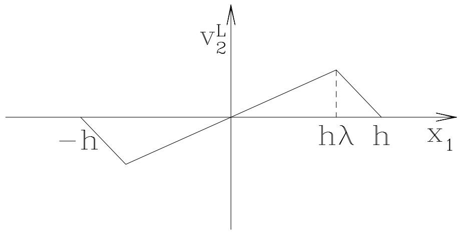

FIGURE 1: Representation of $v_2^L(x_1e_1)$, as defined in (2.3).

(i) $u(x) = x$ *on* $\partial\Omega$*;*

(ii) $\nabla u \in K$ *a.e.;*

(iii) *The domain has the form*

$$\Omega = \mathrm{conv}\left(\{\pm h_i e_i\}_{i=1\ldots n}\right) , \tag{2.5}$$

for some numbers $h_i > 0$ *(which may depend on* ε*,* n *and* λ*);*

(iv) *The function* u *coincides with a single laminate on the central segment* $[-h_1e_1, h_1e_1]$*, i.e.,*

$$u(x_1e_1) = v^L(x_1e_1) \qquad \text{for } x_1 \in (-h_1, h_1) , \tag{2.6}$$

with v^L *as in (2.3);*

(v) *The domain* Ω *can be subdivided into* $5 \cdot 2^{n-1}$ *symplexes such that* u *is affine on each of them;*

(vi) *For any* $i \geq 3$*, and all* x*, one has* $u_i(x) = x_i$*;*

(vii) *For all* x *with* $x_2 = 0$*, one has* $u_1(x) = x_1$*;*

(viii) $\#K \leq 5 \cdot 2^{n-2}$*.*

Condition (vi) implies that, e.g. for $n = 6$, ∇u has the form

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and implies in particular that $\det_r \nabla u$ has the same value for all r, $2 \le r \le n$. The conditions is equivalent to stating that $F^T e_i = e_i$ for all $i \ge 3$, and all $F \in K$.

Proof. We shall prove the case $n = 2$ separately by an explicit construction in Lemma 2.3 below. Here we prove the result for larger n by induction on the dimension.

Let $n \ge 3$, and $\Omega_{n-1} \subset \mathbb{R}^{n-1}$, $u_{n-1} \in W^{1,\infty}(\Omega_{n-1}; \mathbb{R}^{n-1})$ be the result of the Lemma applied in dimension $n-1$, with the same λ and $\varepsilon_{n-1} = \varepsilon/2$. The function u_{n-1} is affine on each of the $n-1$-dimensional symplexes $t_1 ... t_{5\cdot 2^{n-2}}$ composing Ω_{n-1}.

Consider the two points $\pm e_n = (0, \dots, 0, \pm 1)$, and let $\widetilde{\Omega}$ be the convex hull of

$$\Omega_{n-1} \cup \{e_n, -e_n\} \ .$$

(recall that we identify $\Omega_{n-1} \subset \mathbb{R}^{n-1}$ with $\Omega_{n-1} \times \{0\} \subset \mathbb{R}^n$). The set $\widetilde{\Omega}$ can be subdivided in $5 \cdot 2^{n-1}$ symplexes, each with one of the $t_i \times \{0\}$ as basis and one of $\{e_n, -e_n\}$ as vertex. Precisely, we set

$$T_i^+ = \mathrm{conv}\,(t_i, e_n) \ , \qquad T_i^- = \mathrm{conv}\,(t_i, -e_n) \ .$$

We define

$$v(x', 0) = \begin{pmatrix} u_{n-1}(x') \\ 0 \end{pmatrix} , \quad \text{for} \quad x' \in \Omega_{n-1} \, , \tag{2.7}$$

and

$$v(e_n) = e_n \, , \quad v(-e_n) = -e_n \, .$$

In each $T_i^{\pm}$ the function v is defined by affine interpolation between the values at the vertices; since u_{n-1} is affine on each t_i this does not change the value on Ω_{n-1}. Since $u_{n-1}(x) = x$ on $\partial\Omega_{n-1} \subset \mathbb{R}^{n-1}$ and on $\pm e_n$, we automatically have $v(x) = x$ on $\partial\widetilde{\Omega} \subset \mathbb{R}^n$.

For $h > 0$, we define the linear map

$$S_h = \mathrm{Id}_{n-1} + h e_n \otimes e_n = \begin{pmatrix} \mathrm{Id}_{n-1} & 0 \\ 0 & h \end{pmatrix} ,$$

and consider the maps $u_n^{(h)} \in W^{1,\infty}(S_h\widetilde{\Omega}; \mathbb{R}^n)$ defined by

$$u_n^{(h)}(x) = S_h v \left(S_h^{-1} x \right) .$$

We claim that setting $u_n = u_n^{(h)}$ and $\Omega_n = S_h\widetilde{\Omega}$, for sufficiently small h, all stated properties are satisfied, with K being the smallest set obeying (ii) (see (2.8) for an explicit characterization).

Clearly $u_n^{(h)}(x) = x$ on $\partial(S_h\widetilde{\Omega})$, and (i) follows. Condition (ii) holds by the definition of K, and (iii), (iv), (v), (vi), (vii) and (viii) are immediately inherited from the inductive hypothesis.

It remains to show that we can choose h so that (2.4) holds. Since $\nabla u^{(h)} = S_h \nabla v S_h^{-1}$ we have $\det_r \nabla u_n^{(h)} = \det_r \nabla v$ for all $r \leq n$, hence we can focus on ∇v. The key observation is that the gradient of v has a special form. Consider for definiteness one of the symplexes, say $T = \mathrm{conv}(t, e_n)$, let $F \in \mathbb{M}^{n\times n}$ be the value of ∇v on T, and $F' \in \mathbb{M}^{(n-1)\times(n-1)}$ the value of ∇u_{n-1} on t. Since v leaves the set $t \times \{0\}$ invariant, and hence the entire hyperplane $\mathbb{R}^{n-1} = \{x \in \mathbb{R}^n : x_n = 0\}$, it follows that $Fe_i \cdot e_n = 0$ for all $i < n$. Therefore

$$F = \begin{pmatrix} F' & b \\ 0 & F_{nn} \end{pmatrix} ,$$

for some $b \in \mathbb{R}^{n-1}$ and $F_{nn} \in \mathbb{R}$. At the same time, $v_k(x) = x_k$ for all $k \geq 3$ and all $x \in S_h\Omega$ since the same condition holds on the vertices of T. Therefore (vi) holds, and in particular, $F_{nn} = 1$. Therefore

$$\det F = \det F' = 1 , \quad \text{and} \quad \mathrm{Id}_r F \mathrm{Id}_r = \mathrm{Id}_r F' \mathrm{Id}_r \quad \text{for all } r < n .$$

A straightforward computation shows that, again in the considered symplex,

$$\nabla u_n^{(h)} = S_h F S_h^{-1} = \begin{pmatrix} F' & b/h \\ 0 & 1 \end{pmatrix} . \tag{2.8}$$

Here F' and b are fixed in each symplex, i.e., they take finitely many values on $\widetilde{\Omega}$. We finally choose

$$h = \frac{\varepsilon}{2 \max\{|b_i^+| + |b_i^-| : i = 1, \dots, 5 \cdot 2^{n-2}\}} .$$

We obtain from (2.8) that

$$\operatorname{dist}(\nabla u_n, \{A, B\}) \le \operatorname{dist}(\nabla u_{n-1} + e_n \otimes e_n, \{A, B\}) + \frac{1}{2}\varepsilon < \varepsilon \,.$$

This concludes the proof. □

2.2 The two-dimensional construction

We give here an explicit construction for the case $n = 2$. The existence of such a construction was first noticed by Müller and Šverák [22, Remark 2 after Th. 6.1]; their idea was then worked out in detail in [6, Lemma A.2]. We present here a simplification of the latter proof.

Lemma 2.3. *Lemma 2.2 holds for $n = 2$.*

Proof. Let $h > 0$ be a small parameter fixed below. We aim at a construction on the rhombus

$$\omega = \left\{ x \in \mathbb{R}^2 : \frac{|x_1|}{h} + |x_2| < 1 \right\} = \operatorname{conv}\left(\{\pm h e_1, \pm e_2\}\right) ,$$

and work at first on the larger set $R = (-h, h) \times (-1, 1)$. We start from the map v^L defined in (2.3), see Fig. 2a, which satisfies the boundary condition on the left and right sides of R, but not on the top and bottom ones. We shall modify it in order to satisfy the boundary condition on $\partial\omega$. This is done by first composing v^L with another piecewise affine function very close to the identity, and then modifying further a boundary layer. The map v^L can be visualized as shifting the lines $\{x : x_1 = \pm\lambda h\}$ upwards (downwards) by $\lambda(1-\lambda)h$, keeping the lines $\{x : x_1 = \pm h\}$ fixed, and the affine interpolation inside. In an informal language, v^L pushes material (area) downwards on the left-hand side, and upwards on the right-hand side. Hence in order to have an isochoric map we need to push material (area) horizontally: towards the left on the upper side, and towards the right on the lower side, so that the final movement is, in a certain sense, circular. This is done by the map w we shall now construct; we shall be careful to achieve the mass balance by using small gradients on large areas so that the result will be only a small correction to v^L.

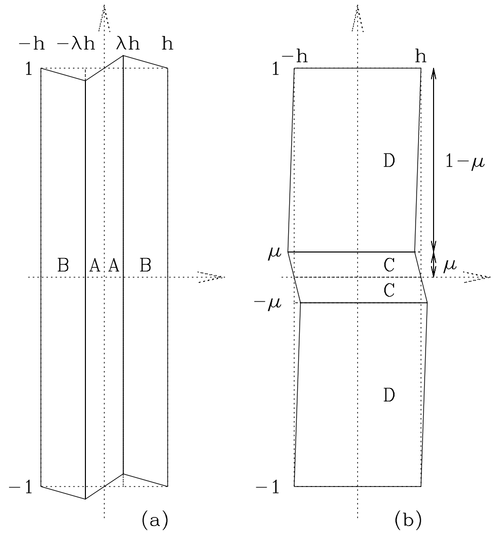

FIGURE 2: The two deformations whose composition is used in the construction of Lemma 2.3. Dotted curves: reference configuration. Full curves: deformed configuration. (a): The map v^L, defined in (2.3), is a vertical shear, which is the identity at the left and right boundaries of R; its gradient takes values A and B. (b): The map w, defined in (2.9), is a horizontal shear, which is the identity on the top and bottom boundaries. Its gradient takes values C and D, which are both close to the identity.

We now construct the second map. For given $\mu \in (0, 1)$ and $q > 0$ to be chosen later (both will be small), we consider the function $w \in W^{1,\infty}_{\text{loc}}(\mathbb{R}^2; \mathbb{R}^2)$ defined by $w(0) = 0$ and

$$\nabla w(x) = \begin{cases} C & \text{for } |x_2| < \mu\,, \\ D & \text{for } \mu < |x_2| < 1\,, \\ \mathrm{Id}_2 & \text{else} \end{cases} \tag{2.9}$$

(see Fig. 2b), where

$$C = \begin{pmatrix} 1 & -q(1-\mu) \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad D = \begin{pmatrix} 1 & q\mu \\ 0 & 1 \end{pmatrix}.$$

The map w shifts the lines $\{x : x_2 = \pm\mu\}$ by $q\mu(1-\mu)$ to the left (right), keeps the lines $\{x : x_2 \in \{-1, 0, 1\}\}$ fixed, and is the affine interpolation

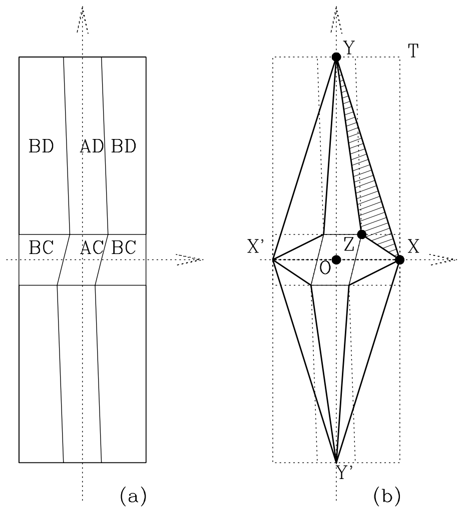

FIGURE 3: Construction used in proving Lemma 2.3. (a): the composition of v^L with w generates a piecewise affine function, which on the considered domain takes six different gradients. (b): the four triangles on which the function is replaced by the affine interpolation.

inside. Further, ∇w has unit determinant, and if q is small then ∇w is uniformly close to the identity. Consider now the composition $v^L \circ w$. This is a piecewise affine map, whose gradient has unit determinant and is close to A or B everywhere. Indeed,

$$|AC - A| \le |A||C - \mathrm{Id}| \le 2q \tag{2.10}$$

and analogously for the other three combinations AD, BC, and BD.

The pieces of R on which $v^L \circ w$ is affine are shown in Figure 3a. We now focus on the first quadrant, and in particular on the triangle

$$OXY = \omega \cap \{x : x_1, x_2 > 0\} = \left\{x : x_1 > 0, x_2 > 0, \frac{x_1}{h} + x_2 < 1\right\},$$

where $X = (h, 0)$, $Y = (0, 1)$, and $O = (0, 0)$, see Fig. 3b. Let Z be the point at the intersection of the two discontinuity lines of $\nabla(v^L \circ w)$ in the first quadrant, i.e., $Z = (\lambda h + q\mu(1-\mu), \mu)$ (for sufficiently small μ and h one can check that $Z \in OXY \subset \omega$). We set $u = v^L \circ w$ in $OXY \setminus XYZ$, and in the triangle XYZ equal to the affine interpolation between the values of $v^L \circ w$ at the three corners. An analogous procedure is used in the other quadrants, so that u is defined on the entire ω. Since in X and Y both v^L and w are the identical map, this function satisfies the boundary condition $u(x) = x$ on the segment XY, and hence on $\partial\omega$.

The gradient ∇u therefore takes the three values AD, AC, BC, plus those on the four external triangles, which since $u(-x) = -u(x)$ are pairwise equal. Hence it takes only five distinct values. It only remains to check that the value of ∇u in XYZ has unit determinant and is close to B. The map is the identity on the side XY. The determinant of the affine interpolation is unity if the area of the triangle XYZ is conserved, namely, if the vector

$$u(Z) - Z = (-q\mu(1-\mu), h\lambda(1-\lambda))$$

is parallel to the vector $XY = (-h, 1)$. This requirement is equivalent to the condition

$$q = h^2 \frac{\lambda(1-\lambda)}{\mu(1-\mu)},$$

which is therefore our choice for q (for any given μ, still to be fixed). The gradient in the triangle is therefore an area-preserving shear along XY, of the form

$$G = \nabla u|_{XYZ} = \mathrm{Id} - p(-h, 1) \otimes (1, h)$$

for some $p \in \mathbb{R}$. In turn, p can be determined by

$$u(Z) - Z = (G - \mathrm{Id})\,(Z - X)$$

which follows from $u(X) = X$ and the fact that u is affine in this triangle. A straightforward computation leads to

$$p = \frac{\lambda(1-\lambda)}{(1-\lambda)(1-h\lambda) - \mu} = \lambda + O(h+\mu)\,.$$

We conclude that

$$|G - B| = O(h+\mu)\,.$$

Finally, we set $\mu = h$, so that $q = O(h)$. Choosing h small enough (with respect to ε, with constants depending on λ) the proof is concluded. $\square$

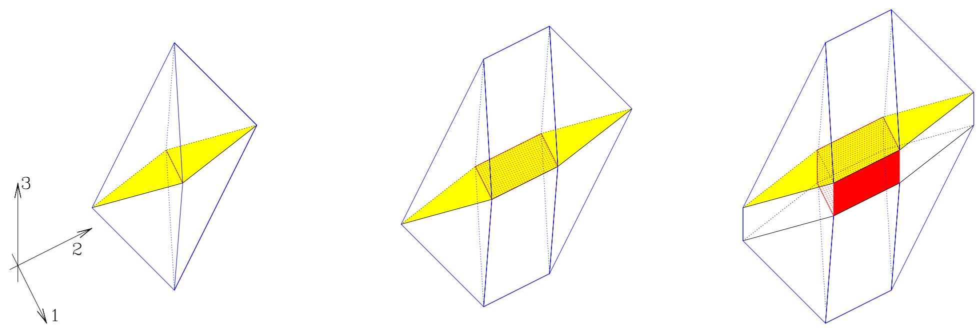

FIGURE 4: Sketch of the construction used in Lemma 2.4 in the three-dimensional case. The left panel shows $u^{(1)}$, the central panel $u^{(2)}$, the right panel $u^{(3)}$. The red region in the center is $\omega^{(k)}$.

2.3 Including large domains with gradient exactly A or B

We now show that the piecewise affine function can be constructed so that $\nabla u = A$ or $\nabla u = B$ on large parts of the domain. This is done by blowing up, in an appropriate way, the central segment $[-h_1 e_1, h_1 e_1]$ on which u coincides with the single laminate v^L as defined in (2.3).

Lemma 2.4. *Let $\lambda \in (0,1)$, $\varepsilon > 0$, $n \geq 2$, A and $B \in \mathbb{M}^{n\times n}$ as in (2.2). Then there is a finite set*

$$K \subset \Big\{F \in \mathbb{M}^{n\times n} : \det_r F = 1 \textit{ for all } 2 \leq r \leq n\,, \\ \min\{|F-A|, |F-B|\} < \varepsilon\Big\} \tag{2.11}$$

such for any $\delta > 0$ there are a domain $\Omega \subset \mathbb{R}^n$ and a map $u \in W^{1,\infty}(\Omega;\mathbb{R}^n)$, with the following properties.

(i) $u(x) = x$ *on* $\partial\Omega$*;*

(ii) $\nabla u \in K$ *a.e.;*

(iii) $\#K \leq 10 \cdot 4^{n-2}$*;*

(iv) The domain Ω can be subdivided into $C \cdot 2^n$ symplexes such that u is affine on each of them;

(v) $|\{x \in \Omega : \nabla u(x) \notin \{A, B\}\}| \leq \delta|\Omega|$.

Proof. Let $u^{(1)}$, $K^{(1)}$ and $\Omega^{(1)}$ be the result obtained from Lemma 2.2, so that $\pm h_1 e_1 \in \partial\Omega^{(1)} \subset \mathbb{R}^n$. Set $L > 0$ (at the end we shall take L large compared to $\max_i h_i/\delta$).

We define inductively, for $k = 2 \dots n$,

$$u^{(k)}(x) = \begin{cases} u^{(k-1)}(x - Le_k) + Le_k & \text{if } x_k > L\,, \\ u^{(k-1)}(x - x_k e_k) + x_k e_k & \text{if } |x_k| \le L\,, \\ u^{(k-1)}(x + Le_k) - Le_k & \text{if } x_k < -L\,, \end{cases}$$

and correspondingly the domain

$$\Omega^{(k)} = \left\{ x : \begin{cases} x - Le_k \in \Omega^{(k-1)} & \text{if } x_k > L\,, \\ x - x_k e_k \in \Omega^{(k-1)} & \text{if } |x_k| \le L\,, \\ x + Le_k \in \Omega^{(k-1)} & \text{if } x_k < -L \end{cases} \right\}, \tag{2.12}$$

see Figure 4. Clearly $u^{(k)} \in W^{1,\infty}(\Omega^{(k)};\mathbb{R}^n)$, and $u^{(k)}(x) = x$ on $\partial\Omega^{(k)}$ (to see this, notice that both $u^{(k-1)}$ and $u^{(k)}$ can be extended continuously by setting them equal to the identical map on the rest of $\mathbb{R}^n$). Further, $\nabla u^{(k)}$ takes at most twice as many values as $\nabla u^{(k-1)}$.

Now we compute the gradient of $u^{(k)}$. From the definition,

$$\nabla u^{(k)}(x) = \begin{cases} \nabla u^{(k-1)}(x') & \text{if } |x_k| > L\,, \\ T_k\left(\nabla u^{(k-1)}(x')\right) & \text{if } |x_k| \le L\,, \end{cases}$$

where $x' \in \Omega^{(k-1)}$ is obtained from x as above, and $T_k : \mathbb{M}^{n\times n} \to \mathbb{M}^{n\times n}$ is the affine map which replaces the k-th column with e_k, i.e.,

$$T_k(F) = F(\mathrm{Id}_n - e_k \otimes e_k) + e_k \otimes e_k\,.$$

The set $K^{(k)}$ is defined as the set of values taken by $\nabla u^{(k)}$; clearly

$$K^{(k)} \subset K^{(k-1)} \cup T_k\left(K^{(k-1)}\right)\,. \tag{2.13}$$

This provides the estimate $\#K^{(k)} \le 2^{k-1}\#K^{(1)}$, hence proves point (iii) of the Lemma. Notice however that in dealing with the determinant constraint (for $k = 2$) it will be relevant that $K^{(k)}$ is *not* identical to the set on the right-hand side, which shows that (iii) is not sharp.

By the special structure of A and B it is clear that $T_k(A) = A$ and $T_k(B) = B$, therefore $K \subset B_\varepsilon(A) \cup B_\varepsilon(B)$.

We now check the condition on the determinant. We first observe that the properties mentioned in (vi) and (vii) of Lemma 2.2 are left unchanged in passing from $u^{(k-1)}$ to $u^{(k)}$, hence they hold for all k.

We start from $k = 2$. The inclusion (2.13) can be refined to

$$K^{(2)} \subset K^{(1)} \cup \big\{T_2(F) : F \in K^{(1)}, \nabla u^{(1)}(x) = F \text{ for some } x \text{ with } x_2 = 0\big) \ .$$

Here "for some x with $x_2 = 0$" means, precisely, "on a subset of positive $n-1$-dimensional measure of the hyperplane $\{x \in \mathbb{R}^n : x_2 = 0\}$". Notice that u has a Lipschitz trace on that hyperplane, the tangential part of its gradient has an L^∞ trace, and $T_k(F)$ only depends on the tangential part. The matrices in $K^{(1)}$ obviously fulfill the determinant conditions. By point (vi) of Lemma 2.2, $\det_r T_2(F) = F_{11}$, for all $F \in K^{(1)}$ and all $r \geq 2$. And by point (vii), of Lemma 2.2, only matrices with $F_{11} = 1$ can appear. This proves that $\det_r F = 1$ for all $r \geq 2$ and all $F \in K^{(2)}$.

The proof for higher k is much simpler. Indeed, if $\det_r F = 1$ for some $r \geq 2$, and F obeys (vi) of Lemma 2.2, then it is clear that $\det_r (T_k(F)) = 1$ for any $k \geq 3$. Therefore the condition on the subdeterminants is satisfied.

It remains to prove point (v). To this end, we claim that, for all $k = 1 \dots n$, one has

$$u^{(k)}(x) = v^L(x) \qquad \text{for } x \in \omega^{(k)} \,, \tag{2.14}$$

where

$$\omega^{(k)} = \big\{x \in \mathbb{R}^n : \quad \begin{aligned} &|x_1| < h_1 \,, \\ &|x_i| < L \text{ for all } i \text{ such that } 2 \leq i \leq k \,, \\ &x_j = 0 \text{ for all } j \text{ such that } k < j \leq n\big\} \,. \end{aligned}$$

(if $k = 1$ the second condition disappears, if $k = n$ the last one disappears). For $k = 1$ the claim holds by (iv) of Lemma 2.2. We now prove (2.14) for $k \geq 2$ by induction. This follows from the fact that

$$v^L(x) = v^L(x - x_k e_k) + x_k e_k \qquad \text{for all } k \geq 2$$

and

$$\omega^{(k)} = \big\{x : |x_k| < L \text{ and } x - x_k e_k \in \omega^{(k-1)}\big\} \,.$$

We now observe that $\omega^{(n)} = (-h_1, h_1) \times (-L, L)^{n-1}$, and by (2.12)

$$\Omega^{(n)} \subset (-h_1, h_1) \times \prod_{i=2}^{n} (-h_i - L, h_i + L) \,.$$

For L sufficiently large with respect to $H = \max_i h_i$, a simple Taylor expansion gives

$$\begin{aligned}\prod_{i=2}^{n} 2(L+h_i) &= (2L)^{n-1}\prod_{i=2}^{n}\left(1+\frac{h_i}{L}\right)\\ &\le (2L)^{n-1}\left(1+2\sum_{i=2}^{n}\frac{h_i}{L}\right) \le (2L)^{n-1}\left(1+2n\frac{H}{L}\right).\end{aligned}$$

Since obviously $|\omega^{(n)}| = 2h_1(2L)^{n-1}$, we conclude that

$$|\Omega^{(n)} \setminus \omega^{(n)}| = |\Omega^{(n)}| - |\omega^{(n)}| \le 2n\frac{H}{L}|\omega^{(n)}|\,.$$

Choosing L sufficiently large (e.g. $L = 4nH/\delta$, if δ is small) also (v) holds. $\square$

2.4 Extension to general matrices

We finally show how Lemma 2.4 implies Theorem 2.1 by showing that we can combine the given construction with an appropriate change of variables.

Proof of Theorem 2.1. First we observe that it suffices to show that for any A, B, λ, ε and δ there is one special open set ω, with $|\partial\omega| = 0$, such that one can construct u with the mentioned properties. The special set will be an affine transformation of the parallelepyped entering Lemma 2.4. This follows by a by-now standard scaling and covering argument, see, e.g., [5, Prop. 2.3] or [15, Construction 3.1].

The key idea in the proof is to apply Lemma 2.4 after a suitable change of variables, which we now discuss. We are given A, $B \in \mathbb{M}^{m\times n}$ with $\mathrm{rank}(A-B) = 1$, $P \in \mathbb{M}^{m\times m}$ and $Q \in \mathbb{M}^{n\times n}$, $t \neq 0$, with

$$\det\nolimits_r(PAQ) = \det\nolimits_r(PBQ) = t\,.$$

Since $t \neq 0$, the first r rows of P must be linearly independent, and the first r columns of Q as well. We can therefore modify the last $m-r$ rows of P and the last $n-r$ rows of Q so that both matrices have full rank (precisely, $\mathrm{rank}\,P = m$, $\mathrm{rank}\,Q = n$). This does not change any of the conditions in the theorem, hence it suffices to consider the case of invertible P and Q.

It is clear that by scaling it suffices to consider $t = 1$. We now show that it suffices to prove the theorem in the case $P = \mathrm{Id}_m$, $Q = \mathrm{Id}_n$. To see this, let

$$\widetilde{A} = PAQ\,, \quad \widetilde{B} = PBQ\,.$$

Let $\widetilde{K}$, $\widetilde{u}$ be the result obtained applying the theorem to $\widetilde{A}$, $\widetilde{B}$, with $\widetilde{P} = \mathrm{Id}_m$, $\widetilde{Q} = \mathrm{Id}_n$, on the domain $\widetilde{\Omega} = Q^{-1}\Omega$, and eventually a smaller ε and δ. We define

$$K = P^{-1}\widetilde{K}Q^{-1}\,, \qquad u(x) = P^{-1}\widetilde{u}(Q^{-1}x)\,.$$

Then it is clear that for any $F \in K$ one has $\det_r PFQ = \det_r \widetilde{F} = 1$, where $\widetilde{F} = PFQ \in \widetilde{K}$. Analogously, $|F - A| \le |P^{-1}||Q^{-1}||\widetilde{F} - \widetilde{A}|$. Therefore the pair (K, u) proves the theorem.

From now on we only consider the situation $P = \mathrm{Id}_n$, $Q = \mathrm{Id}_m$, $t = 1$. Further, we can assume $0 < \lambda < 1$, since in the two extreme cases $K = \{A, B\}$ will do. Let $A - B = a \otimes \nu$, where we can assume $|\nu| = 1$. We distinguish three cases, depending on whether the vectors a and ν have a zero or a nonzero projection on $\mathbb{R}^r$, i.e., if they have components which are relevant for the nonlinear constraint or not.

Case 1: $\mathrm{Id}_r a \neq 0$ and $\mathrm{Id}_r \nu \neq 0$. We write $C = \lambda A + (1-\lambda)B$, so that $A = C + (1-\lambda)a \otimes \nu$. Let $D = \mathrm{Id}_r C \mathrm{Id}_r$. Since $\det_r C = 1$, the matrix $D \in \mathbb{M}^{r\times r}$ is invertible. We set $\alpha_1 = \mathrm{Id}_r\nu$, and $\alpha_2 = D^{-1}\mathrm{Id}_r a$. Since

$$\begin{aligned} 1 &= \det_r A = \det_r\,(C + (1-\lambda)a\otimes\nu) = \det_r\,(D(\mathrm{Id}_r + (1-\lambda)\alpha_2\otimes\alpha_1)) \\ &= (\det_r D)\,(1 + \mathrm{Tr}[(1-\lambda)\alpha_2\otimes\alpha_1]) = 1 + (1-\lambda)\alpha_1\cdot\alpha_2 \end{aligned}$$

it follows that $\alpha_1 \cdot \alpha_2 = 0$. Let $\alpha_3 \dots \alpha_r \in \mathbb{R}^r$ be unit vectors such that the set $\{\alpha_i/|\alpha_i|\}$, $i = 1 \dots r$, forms an orthonormal basis of $\mathbb{R}^r$. We let u' and K' be the result of Lemma 2.2 with the same n and λ, an possibly a smaller ε' (chosen below), set

$$u(x) = Tu'(Sx) + Ux\,,$$

and

$$K = TK'S + U = \{F \in \mathbb{M}^{m\times n} : F = TF'S + U \text{ for some } F' \in K'\}\,.$$

Here

$$T = \xi D\alpha_1 \otimes e_1 + a \otimes e_2 + \sum_{i=3}^{r} D\alpha_i \otimes e_i\,,$$

$$S = e_1 \otimes \nu + \eta e_2 \otimes \alpha_2 + \sum_{i=3}^{r} e_i \otimes \alpha_i \,,$$

and

$$U = C - TS\,.$$

The real parameters ξ, η will be chosen below. We compute

$$\mathrm{Id}_r\, TS\, \mathrm{Id}_r = D \left(\xi\alpha_1 \otimes \alpha_1 + \eta\alpha_2 \otimes \alpha_2 + \sum_{i=3}^{r} \alpha_i \otimes \alpha_i \right)\,.$$

Therefore choosing $\xi = 1/|\alpha_1|^2$, $\eta = 1/|\alpha_2|^2$ the parenthesis becomes Id_r, and we obtain $\mathrm{Id}_r U \mathrm{Id}_r = D - \mathrm{Id}_r ST \mathrm{Id}_r = 0$. We further remark that by definition $Te_j = 0 = e_j S$ for all $j > r$, hence $T = T\mathrm{Id}_r$ and $S = \mathrm{Id}_r S$. Therefore for any $F' \in K'$ one has, for $F = TF'S + U$,

$$\begin{aligned}
\det_r F &= \det_r (TF'S + U) = \det_r (\mathrm{Id}_r TF'S \mathrm{Id}_r) \\
&= \det_r (\mathrm{Id}_r T \mathrm{Id}_r F' \mathrm{Id}_r S \mathrm{Id}_r) = \det_r T \det_r F' \det_r S \\
&= \det_r (T \mathrm{Id}_r S) = \det_r (TS) = 1\,,
\end{aligned}$$

since $F' \in K'$. This proves that $\det_r F = 1$ for all $F \in K$.

We finally observe that

$$T(e_2 \otimes e_1)S = a \otimes \nu\,,$$

therefore $T(\mathrm{Id}_n + (1-\lambda)e_2 \otimes e_1)S + U = A$, and $T(\mathrm{Id}_n - \lambda e_2 \otimes e_1)S + U = B$. Therefore it suffices to choose $\varepsilon' = \varepsilon/(1 + |T||S|)$ and the thesis is proven.

Case 2: $\mathrm{Id}_r a = 0$**.** This is a degenerate case, in which the determinant constraint is irrelevant. A direct construction is possible, see, e.g., [15, Lemma 3.2]. We show here how a variant of the construction of Case 1 also applies. Precisely, we let α_1 be as above, and define $\alpha_2 \dots \alpha_r \in \mathbb{R}^r$ so that the vectors $\{\alpha_i/|\alpha_i|\}$, $i = 1 \dots r$, form an orthornormal basis of $\mathbb{R}^r$. We set

$$T = a \otimes e_2$$

and

$$S = e_1 \otimes \nu + \sum_{i=2}^{r} e_i \otimes \alpha_i\,.$$

We observe that

$$\mathrm{Id}_r T = 0\,, \qquad T(e_2 \otimes e_1)S = a \otimes \nu\,.$$

The rest of the construction is the same. In particular,

$$\det_r U = \det_r (C - TS) = \det_r C = 1\,,$$

and for any $F' \in K'$ we have

$$\det_r F = \det_r (TF'S + U) = \det_r U = 1\,.$$

This concludes the proof in this case.

Case 3: $\mathrm{Id}_r \nu = 0$**.** In this case we need a more complex construction, based on a second-order laminate. This will result in a set depending on *both* ε and δ. Therefore it suffices to work in a case where they are equal (otherwise we replace both with the minimum). We pick a vector $f \in \mathbb{R}^r \setminus \{0\}$ such that

$$\det_r (A + a \otimes f) = 1\,.$$

(this is one linear equation in f, which is compatible since $f = 0$ is a solution, hence it has an $r - 1$-dimensional set of solutions). We set

$$A' = A + \frac{1}{2}(1-\lambda)\varepsilon^2 a \otimes f\,, \quad B' = B - \frac{1}{2}\lambda\varepsilon^2 a \otimes f\,.$$

It is clear that $\lambda A' + (1-\lambda)B' = \lambda A + (1-\lambda)B = C$, and that

$$A' - B' = a \otimes \nu + \frac{1}{2}\varepsilon^2 a \otimes f = a \otimes \nu'\,, \quad \text{where } \nu' = \nu + \frac{1}{2}\varepsilon^2 f\,.$$

Further, by construction $\mathrm{Id}_r \nu' = \varepsilon^2 f/2 \neq 0$. We apply Case 1, with $\varepsilon_1 = \varepsilon/2$, to this pair, and find a set

$$K_1 \subset \left\{ F \in \mathbb{M}^{m\times n} : \det_r F = 1\,,\ \min\{|F - A'|, |F - B'|\} < \frac{\varepsilon}{2} \right\}$$

and a (δ-dependent) function u_1 with gradient A' or B' on a large part of Ω. Let now

$$\omega_A = \{x \in \Omega : \nabla u_1(x) = A'\}\ ,$$

and analogously ω_B. Notice that $\omega_A \cup \omega_B$ cover at least a $1-\delta$ fraction of Ω.

We apply again Case 1 with the pair of rank-one connected matrices (A, A''), where

$$A'' = A + \frac{1}{2}(1-\lambda)\varepsilon a \otimes f\,,$$

with weight $\lambda''_A = \varepsilon$, and again $\varepsilon_1 = \varepsilon/2$. Indeed, it is clear that

$$A' = \varepsilon A'' + (1-\varepsilon)A\,.$$

This gives us a set K_2^A, and an analogous construction for B gives us K_2^B, with the properties stated in Lemma 2.2.

We finally construct u. We start by setting $u = u_1$ on $\Omega \setminus (\omega_A \cup \omega_B)$. By Lemma 2.2, the set ω_A can be decomposed into finitely many symplexes, let σ be one of them. Since $\nabla u_1 = A'$ on ω_A, and $\sigma \subset \omega_A$ is connected, there is $b_\sigma \in \mathbb{R}^m$ such that $u_1(x) = A'x + b_\sigma$ on σ. Let u_σ be the result of Lemma 2.2 applied on σ, which obeys $\nabla u_\sigma \in K_2^A$ a.e.. We set then $u = u_\sigma + b_\sigma$ on σ. This makes u continuous on $\partial\sigma$. We proceed analogously for all other symplexes composing ω_A and ω_B. The function u so constructed is Lipschitz continuous, satisfies $\nabla u \in K_1 \cup K_2^A \cup K_2^A$, and all other stated properties are inherited by Lemma 2.2. This concludes the proof (in this third case we used three times the construction of Lemma 2.2, hence obtained a set K which is up to three times as large). ☐

3 Constrained quasiconvex functions are rank-one convex

Theorem 3.1. *Let $W : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ be quasiconvex, $P \in \mathbb{M}^{m\times m}$, $Q \in \mathbb{M}^{n\times n}$, and*

$$\Sigma = \left\{F \in \mathbb{M}^{m\times n} : \det_r PFQ = t\right\}\,,$$

for some $t \neq 0$, with $2 \leq r \leq \min\{m, n\}$. If $W(F) < \infty$ for all $F \in \Sigma$, then W is rank-one convex, and hence continuous, on the surface Σ.

We recall that by definition W is rank-one convex on Σ if for any pair A, $B \in \Sigma$ with $\operatorname{rank}(A - B) = 1$ and any $\lambda \in (0, 1)$, it holds

$$W(\lambda A + (1-\lambda)B) \leq \lambda W(A) + (1-\lambda)W(B)\,. \tag{3.1}$$

Proof. The strategy of the proof is to first prove (3.1) under the additional condition that $(A-B)Q\mathrm{Id}_r \neq 0$, then to use this to prove continuity, and finally use continuity to infer (3.1) also for the case $(A-B)Q\mathrm{Id}_r = 0$.

Step 1. Rank-one convexity in generic directions. Let A, $B \in \Sigma$ with $\mathrm{rank}(A-B) = 1$ with $(A-B)Q\mathrm{Id}_r \neq 0$, $\lambda \in (0,1)$. Fix a domain, say, $\Omega = (0,1)^n$. By Theorem 2.1 (with $\varepsilon = 1$) there is a finite set $K \subset \Sigma$ such that for any $\delta > 0$ there is u^δ such that

$$\nabla u^\delta \in K \text{ a.e.}\,, \qquad \big|\{x \in \Omega : \nabla u^\delta(x) \notin \{A, B\}\}\big| \leq \delta|\Omega|\,,$$

and $u^\delta(x) = \lambda A + (1-\lambda)B$ on $\partial\Omega$. Since K is bounded, and the average of ∇u^δ is $\lambda A + (1-\lambda)B$, it is clear that

$$\lim_{\delta\to 0} \frac{|\{x \in \Omega : \nabla u^\delta = A\}|}{|\Omega|} = \lambda\,, \qquad \lim_{\delta\to 0} \frac{|\{x \in \Omega : \nabla u^\delta = B\}|}{|\Omega|} = 1 - \lambda\,.$$

Since W is quasiconvex, we have

$$\begin{aligned} W(\lambda A + (1-\lambda)B) &\leq \frac{1}{|\Omega|}\int_\Omega W(\nabla u^\delta)\,\mathrm{d}x \\ &= \sum_{F\in K} W(F)\frac{|\{x \in \Omega : \nabla u^\delta = F\}|}{|\Omega|}\,. \end{aligned}$$

Let $M = \max\{W(F) : F \in K\}$ (here it is important that K is finite, and that it does not depend on δ). Then

$$\begin{aligned} W(\lambda A + (1-\lambda)B) \leq\; & W(A)\frac{|\{x \in \Omega : \nabla u^\delta = A\}|}{|\Omega|} + \\ & W(B)\frac{|\{x \in \Omega : \nabla u^\delta = B\}|}{|\Omega|} + \\ & M\frac{|\{x \in \Omega : \nabla u^\delta \notin \{A, B\}\}|}{|\Omega|}\,. \end{aligned}$$

Taking the limit $\delta \to 0$ we obtain (3.1).

Step 2. Continuity. We now prove that W is continous on Σ. To do this, we show that W can be locally written as a separately convex function on $\mathbb{R}^{nm-1}$, this implies continuity by standard arguments (see, e.g., [7, Th. 2.3]).

To make the strategy clear we first consider the case $m = n = r$, $t = 1$, and prove continuity in a neighborhood of the identity. Let $R_1 ... R_{r^2-1} \in \mathbb{M}^{r\times r}$ be linearly independent traceless rank-one matrices, e.g., the matrices $e_i \otimes e_j$ (for $i \neq j$, $1 \leq i, j \leq r$) and the matrices $(e_i + e_{i+1}) \otimes (e_i - e_{i+1})$ (for $1 \leq i < r$). Then the map

$$\begin{aligned} \psi(x) &= (\mathrm{Id}_r + x_1 R_1)(\mathrm{Id}_r + x_2 R_2) \dots (\mathrm{Id}_r + x_{r^2-1} R_{r^2-1}) \\ &= \prod_{i=1}^{r^2-1} (\mathrm{Id}_r + x_i R_i) \end{aligned}$$

is a diffeomorphism of a neighbourhood of zero in $\mathbb{R}^{r^2-1}$ onto a neighbourhood of Id_r in Σ. Indeed, $\det(\mathrm{Id}_r + x_i R_i) = 1 + x_i \operatorname{Tr} R_i = 1$, hence $\det \psi(x) = 1$ for all x. Further, since the R_i are linearly independent the gradient $D\psi(0) = \sum_i R_i \otimes e_i$ has full rank.

We claim that $W \circ \psi : \mathbb{R}^{r^2-1} \to \mathbb{R}$ is separately convex. To see this, fix $J \in \{1, \dots, r^2 - 1\}$, choose some $x \in \mathbb{R}^{r^2-1}$, and for $t \in \mathbb{R}$ let $x_t = x + te_J$. A simple computation shows that

$$\begin{aligned} \psi(x_t) &= \prod_{i=1}^{J-1} (\mathrm{Id}_r + x_i R_i)(\mathrm{Id}_r + x_J R_J + tR_J) \prod_{i=J+1}^{r^2-1} (\mathrm{Id}_r + x_i R_i) \\ &= \psi(x) + ta \otimes \nu \end{aligned}$$

where

$$a = \prod_{i=1}^{J-1} (\mathrm{Id}_r + x_i R_i)\, \alpha \,, \qquad \nu = \beta \prod_{i=J+1}^{r^2-1} (\mathrm{Id}_r + x_i R_i) \,,$$

and α, β are defined by $R_J = \alpha \otimes \beta$. From Step 1 we obtain that $t \mapsto W(\psi(x) + ta \otimes \nu)$ is convex, i.e., $t \mapsto W(\psi(x + te_J))$ is convex for any fixed x and J. This means that $W \circ \psi$ is separately convex, which implies that it is continuous.

Now we consider the general case. After a change of variables (as explained in the first part of the proof of Theorem 2.1) we can reduce to the case $P = \mathrm{Id}_m$, $Q = \mathrm{Id}_n$. Let $R_1 \dots R_{r^2-1}$ be as above, and $R_{r^2} \dots R_{rn-1}$ be the $r(n-r)$ matrices $e_i \otimes (e_{1+(i \mod r)} + e_j)$, for $1 \leq i \leq r$, $r < j \leq n$. These are also traceless and have rank one; all matrices $R_1 \dots R_{rn-1}$ are linearly independent, and all satisfy $R_i \mathrm{Id}_r \neq 0$. Further, let $Q_1 \dots Q_{n(m-r)}$ be the matrices $e_i \otimes e_j$, with $r < i \leq m$, $1 \leq j \leq n$. We fix a matrix $F \in \Sigma$, and define the map $\psi : \mathbb{R}^{nm-1} \to \mathbb{M}^{m\times n}$ by

$$\psi(x) = (F - \widetilde{F}) + \widetilde{F} \prod_{i=1}^{rn-1} (\mathrm{Id}_n + x_i R_i) + \sum_{i=1}^{n(m-r)} x_{i+rn-1} Q_i \,,$$

where $\widetilde{F} = \mathrm{Id}_r F \mathrm{Id}_r$. We claim that

$$\mathrm{Id}_r \prod_{i=1}^{rn-1} (\mathrm{Id}_n + x_i R_i) \, \mathrm{Id}_r = \prod_{i=1}^{rn-1} (\mathrm{Id}_r + x_i \mathrm{Id}_r R_i \mathrm{Id}_r) \ . \tag{3.2}$$

To see this, consider e.g., that this expression can be written as Id_r plus a sum of products of the form $\mathrm{Id}_r R_{i_1} R_{i_2} \dots R_{i_k} \mathrm{Id}_r$, with coefficients depending on the x_i's. Since by definition $R_i = \mathrm{Id}_r R_i$, we obtain that $\mathrm{Id}_r R_{i_1} R_{i_2} \dots R_{i_k} \mathrm{Id}_r = \mathrm{Id}_r R_{i_1} \mathrm{Id}_r R_{i_2} \mathrm{Id}_r \dots \mathrm{Id}_r R_{i_k} \mathrm{Id}_r$.

Equation (3.2) shows that $\det_r \psi(x) = \det_r \widetilde{F}$ for all x. Since the matrices $\{R_1 \dots R_{nr-1}, Q_1 \dots Q_{n(m-r)}\} \in \mathbb{M}^{m\times n}$ are linearly independent the map ψ gives a diffeomorphism of a neighborhood of zero in $\mathbb{R}^{nm-1}$ onto a neighborhood of the identity in Σ. Argueing as above, and observing that $R_i \mathrm{Id}_r \neq 0$ for all i, we see that $W \circ \psi$ is separately convex, and hence continuous; this implies continuity of W in a neighborhood of F. But F was a generic matrix in Σ, hence W is continuous on Σ.

Step 3. Rank-one convexity in all directions. By the usual change of variables we can assume $P = \mathrm{Id}_m$, $Q = \mathrm{Id}_n$. We choose, analogously to Case 3 in the proof of Theorem 2.1, a vector $f \in \mathbb{R}^r \setminus \{0\}$ such that

$$\det_r (A + a \otimes f) = 1 \,,$$

and define, for $\varepsilon > 0$,

$$A_\varepsilon = A + (1-\lambda)\varepsilon a \otimes f \,, \quad B_\varepsilon = B - \lambda \varepsilon a \otimes f \,.$$

Clearly for all ε we have $\lambda A_\varepsilon + (1-\lambda) B_\varepsilon = \lambda A + (1-\lambda) B$, and

$$A_\varepsilon - B_\varepsilon = a \otimes \nu + \varepsilon a \otimes f = a \otimes (\nu + \varepsilon f)$$

has rank one. Further, by construction $(A_\varepsilon - B_\varepsilon)\mathrm{Id}_r = \varepsilon a \otimes f \neq 0$. Therefore from Step 1 we obtain

$$W(\lambda A + (1-\lambda)B) \leq \lambda W(A_\varepsilon) + (1-\lambda) W(B_\varepsilon) \,.$$

But since by Step 2 the function W is continuous on Σ, we obtain

$$\lim_{\varepsilon\to 0} W(A_\varepsilon) = W\left(\lim_{\varepsilon\to 0} A_\varepsilon\right) = W(A)\,,$$

and analogously for B_ε. This concludes the proof. $\square$

Remark 3.2. *Under the same assumptions as Theorem 3.1, if $\sigma \subset \Sigma$ is relatively open, i.e., $\sigma = \Sigma \cap U$ for some open $U \in \mathbb{M}^{m\times n}$, and $W(F) < \infty$ for all $F \in \sigma$, then W is rank-one convex, and hence continuous, on σ.*

Proof. It suffices to repeat the same argument taking ε sufficiently small so that

$$\{F \in \Sigma : \mathrm{dist}(F, \{A, B\}) < \varepsilon\} \subset \sigma\,.$$

$\square$

In closing we remark that, for extended-valued functions, quasiconvexity is not invariant under transposition, in any dimension larger than 2. In dimension $m \geq 3$, $n \geq 2$, this was first proven, extending the famous example by Šverák of a rank-one convex which is not quasiconvex [25], by Kružík [16] for extended-valued functions, and by Müller [20] for real-valued functions. We give here a shorter argument for extended-valued functions, which also works in the 2×2 case. For real-valued functions the case $m = 2$, $n \geq 2$ remains open.

Lemma 3.3. *For any $n, m \geq 2$ there is a function $W : \mathbb{M}^{m\times n} \to [0, \infty]$ which is quasiconvex, and with the property that the function $V : \mathbb{M}^{n\times m} \to [0, \infty]$ defined by $V(F) = W(F^T)$ is not quasiconvex.*

Proof. It suffices to prove the Lemma in the case $n = m = 2$, and then to extend the functions so that they do not depend on the other coordinates. Consider for some $f : \mathbb{R}^2 \to [0, \infty)$ the functions $W, V : \mathbb{M}^{2\times 2} \to [0, \infty]$ defined by

$$W(F) = \begin{cases} f(F_{11}, F_{21}) & \text{if } F_{12} = F_{22} = 0\,, \\ \infty & \text{else} \end{cases}$$

and

$$V(F) = W(F^T) = \begin{cases} f(F_{11}, F_{12}) & \text{if } F_{21} = F_{22} = 0\,, \\ \infty & \text{else} \end{cases}$$

The function W is quasiconvex for any f. Indeed, let $F \in \mathbb{M}^{2\times 2}$ and $u \in W^{1,\infty}(\Omega;\mathbb{R}^2)$ be such that $u(x) = Fx$ on $\partial\Omega$ and $\int_\Omega W(\nabla u)dx < \infty$. Then $\partial u/\partial x_2 = 0$ almost everywhere, hence $Fe_2 = 0$ and $u(x) = Fx$ almost everywhere (this is essentially the same argument used by Ball and Murat [5] to prove that the function given in (1.3) is quasiconvex).

The function V is instead quasiconvex if and only if f is convex. This can be proven argueing as in Case 2 of Theorem 2.1, since only the case $a = e_1$ is relevant. One simple example of a nonconvex function f which produces a nonquasiconvex V is $f(t) = (1 - |t|^2)^2$, a possible test function being $u(x) = \mathrm{dist}(x, \partial\Omega)e_1$, for $\Omega = (0,1)^2$. □

Acknowledgements

I wish to thank Stefan Müller for several interesting discussions. This work was partially supported by the Deutsche Forschungsgemeinschaft through the Schwerpunktprogramm 1095 *Analysis, Modeling and Simulation of Multiscale Problems.*

References

[1] J. Adams, S. Conti, A. DeSimone, and G. Dolzmann, *Relaxation of some transversally isotropic energies and applications to smectic A elastomers*, preprint, to appear in Math Mod. Meth. Appl. Sci. (2007).

[2] J. M. Ball, *Convexity conditions and existence theorems in nonlinear elasticity*, Arch. Rational Mech. Anal. **63** (1977), 337–403.

[3] J. M. Ball, *Some open problems in elasticity*, Geometry, mechanics, and dynamics (P. Newton, P. Holmes, and A. Weinstein, eds.), Springer, New York, 2002, pp. 3–59.

[4] J. M. Ball, B. Kirchheim, and J. Kristensen, *Regularity of quasiconvex envelopes*, Calc. Var. Partial Differential Equations **11** (2000), 333–359.

[5] J. M. Ball and F. Murat, $W^{1,p}$*-quasiconvexity and variational problems for multiple integrals*, J. Funct. Anal. **58** (1984), 225–253.

[6] S. Conti and F. Theil, *Single-slip elastoplastic microstructures*, Arch. Rat. Mech. Anal. **178** (2005), 125–148.

[7] B. Dacorogna, *Direct methods in the calculus of variations*, Springer-Verlag, New York, 1989.

[8] B. Dacorogna and J. Moser, *On a partial differential equation involving the Jacobian determinant*, Ann. Inst. H. Poincaré Anal. Non Linéaire **7** (1990), 1–26.

[9] A. DeSimone and G. Dolzmann, *Macroscopic response of nematic elastomers via relaxation of a class of $SO(3)$-invariant energies*, Arch. Rat. Mech. Anal. **161** (2002), 181–204.

[10] G. Dolzmann, *Variational methods for crystalline microstructure - analysis and computation*, Lecture Notes in Mathematics, no. 1803, Springer-Verlag, 2003.

[11] I. Fonseca, *The lower quasiconvex envelope of the stored energy function for an elastic crystal*, J. Math. pures et appl. **67** (1988), 175–195.

[12] A. Fortunelli and M. Ortiz, *A constitutive model for plasticity in an amorphous polycarbonate*, in preparation.

[13] M. Gromov, *Partial differential relations*, Springer-Verlag, 1986.

[14] B. Kirchheim, private communication.

[15] ———, *Rigidity and geometry of microstructures*, MPI-MIS Lecture notes 16, 2002.

[16] M. Kružík, *On the composition of quasiconvex functions and the transposition*, J. Convex Anal. **6** (1999), 207–213.

[17] N. Kuiper, *On C^1 isometric imbeddings I and II*, Proc. Kon. Acad. Wet. Amsterdam A **58** (1955), 545–556 and 683–689.

[18] C. B. Morrey, *Quasi-convexity and the lower semicontinuity of multiple integrals*, Pacific J. Math. **2** (1952), 25–53.

[19] S. Müller, *Variational models for microstructure and phase transitions*, in: Calculus of variations and geometric evolution problems (F. Bethuel et al., eds.), Springer Lecture Notes in Math. 1713, Springer-Verlag, 1999, pp. 85–210.

[20] S. Müller, *Quasiconvexity is not invariant under transposition*, Proc. Roy. Soc. Edinburgh Sect. A **130** (2000), 389–395.

[21] S. Müller and V. Šverák, *Attainment results for the two-well problem by convex integration*, Geometric analysis and the calculus of variations, Internat. Press, Cambridge, MA, 1996, pp. 239–251.

[22] S. Müller and V. Šverák, *Convex integration with constraints and applications to phase transitions and partial differential equations*, J. Eur. Math. Soc. (JEMS) **1** (1999), 393–442.

[23] ———, *Convex integration for Lipschitz mappings and counterexamples to regularity*, Ann. Math. **157** (2003), 715–742.

[24] J. Nash, *C^1 isometric imbeddings*, Ann. Math. **60** (1954), 383–396.

[25] V. Šverák, *Rank-one convexity does not imply quasiconvexity*, Proc. Roy. Soc. Edin. Sect. A **120** (1992), 185–189.