brought to you by CORE

THEORETICAL AND APPLIED MECHANICS vol. 28-29, pp. 325-336, Belgrade 2002 Issues dedicated to memory of the Professor Rastko Stojanović

An O(n) invariant rank 1 convex function that is not polyconvex

M. Šilhavý *

Abstract

An O(n) invariant nonnegative rank 1 convex function of linear growth is given that is not polyconvex. This answers a recent question [8, p. 182] and [5]. The polyconvex hull of the function is calculated explicitly if n = 2.

1 Introduction

Consider the energy functional of the calculus of variations and nonlinear elasticity

$$I(u) = \int_{\Omega} f(Du) \, dx \tag{1}$$

where $\Omega \subset \mathbb{R}^n$ is an open bounded set, $u : \Omega \to \mathbb{R}^m$ is in $W^{1,p}(\Omega, \mathbb{R}^m)$, $Du : \Omega \to \mathbb{M}^{m \times n}$ is its gradient and $f : \mathbb{M}^{m \times n} \to \mathbb{R} \cup \{\infty\}$ is the energy defined on the set $\mathbb{M}^{m \times n}$ of all $m \times n$ matrices. In nonlinear elasticity one usually sets m = n = 3, Ω is the reference configuration of the body, the elements $x \in \Omega$ are material points, u is the placement of the body, and A = Du the deformation gradient; moreover, $f(A) = \infty$ on the set of all matrices with nonpositive determinant. The minimizers of (1) under appropriate boundary conditions, if they exist, correspond to elastic equilibrium.

^{*}Mathematical Institute of the AV ČR, Prague, Czech Republic (e-mail: silhavy@math.cas.cz)

The recent interest in the general mathematical properties of the functional I is motivated by the fact that they seem to reflect faithfully the real behavior of the body as given by experiments. Thus, for example, the energies used to model rubber–like materials lead to I which possesses states of minimum energy, while, on the other hand, the energy of solids near the temperature of phase transformation in general does not posses minimizers. In this situation the solid exhibits microstructures of alternating phases.

Basic properties of (1) are related to the semiconvexity properties of f: the classical convexity and its weakened modes, rank 1 convexity, quasiconvexity and polyconvexity. See Section 2 for definitions. The quasiconvexity of f is a necessary (and under additional hypotheses also sufficient) condition for the functional I to be sequentially weakly lower semicontinuous on $W^{1,p}(\Omega)$, and is thus directly related to the existence of minimizers. The quasiconvexity is hard to verify; for finite–valued functions there is a simpler necessary condition, the rank 1 convexity, and a simpler sufficient condition, the polyconvexity. As mentioned, for finite-valued functions,

```
convexity \Rightarrow polyconvexity \Rightarrow quasiconvexity \Rightarrow rank 1 convexity.
```

As this paper deals with the relationship between these concepts, it is worth mentioning that the convexity is not appropriate in nonlinear elasticity because it leads to undesirable features, such as unqualified uniqueness of equilibrium, etc. More importantly, it is incapable of distinguishing between the energies leading to the existence of equilibrium from those which do not have it. The remaining three semiconvexity concepts, the polyconvexity, quasiconvexity and rank 1 convexity, serve the purpose much better. The quasiconvexity has been discussed above. The rank 1 convexity leads to Maxwell's relation expressing the thermodynamic equilibrium on the interface between solid phases. This in turn leads to Eshelby's tensor and Eshelby's conservation law (see, e.g., [11, Section 17.4]). The polyconvexity hypothesis seems to be necessary to obtain the existence of equilibrium under the realistic constraint $f(A) \to \infty$ as det $A \to 0$ and $f(A) = \infty$ if det $A \leq 0$, [3].

All the four semiconvexity notions coincide if m = 1 or n = 1; if $m \ge 2, n \ge 2$, the convexity is different from each of the remaining three notions while the detailed relationships among these three are

intricate. It is well-known that if $m \ge 2, n \ge 2$, the polyconvexity and rank 1 convexity are distinct and if $m \ge 3, n \ge 2$, also the rank 1 convexity and quasiconvexity are distinct, [13]. The case m = n = 2 is open and also the full understanding is still missing.

Depending on the interpretation of (1), the energy f may have symmetry properties. A function $f : \mathbb{M}^{n \times n} \to \mathbb{R} \cup \{\infty\}$ is said to be SO(n) invariant if

$$f(QAR) = f(A) \tag{2}$$

for each $A \in \mathbb{M}^{n \times n}$, $Q, R \in SO(n)$. The energies of isotropic solids are SO(n) invariant. f is said to be O(n) invariant if (2) holds for each $A \in \mathbb{M}^{n \times n}$, $Q, R \in O(n)$; each O(n) invariant function is SO(n)invariant.

Of particular importance is the case n = 2. It is known that in the class of SO(2) invariant functions the convexity, polyconvexity, and quasiconvexity are distinct, while the relationship between the quasiconvexity and rank 1 convexity is an open problem. An example of an SO(2) invariant rank 1 convex function that is not polyconvex is due to Aubert [2] (for further examples see [7], [1]) while an example of an SO(2) invariant quasiconvex function that is not polyconvex is due to Alibert & Dacorogna [1]. None of these examples is O(2) invariant. Thus they leave open the possibility that the rank 1 convexity and polyconvexity are equivalent within the narrower class of O(2) invariant functions, cf. [5] and [8, p. 182]. This note shows that this is not the case. The function $f: \mathbb{M}^{2\times 2} \to \mathbb{R}$ given by (3) below is rank 1 convex, O(2) invariant, but not polyconvex. Here

$$f(A) = \begin{cases} v_1 v_2 & \text{if } v_1 \le 1, \\ v_1 + v_2 - 1 & \text{if } v_1 \ge 1, \end{cases}$$
(3)

for all $A \in \mathbb{M}^{n \times n}$, where $v_1 \geq v_2$ are the singular values of A (eigenvalues of $\sqrt{AA^T}$). The function is tailored from two simple polyconvex functions and the proof takes a somewhat different course. First, the verification of the Legendre–Hadamard condition drops because of the polyconvexity of the two pieces of the function; one has only to verify the "rank 1" monotonicity of the subgradient at the surface where the two functions meet. An elementary assertion, Lemma 4, is proved to delineate situations when this is the case for the construction. (The

same construction with convex functions results in a convex function, but (3) shows that it may fail to result in a polyconvex function even if the starting functions were polyconvex.) Secondly, the nonpolyconvexity of (3) is based on a result [12] that the polyconvexity is inconsistent with the subquadratic growth unless it reduces to convexity. In [12] it will be shown that there are many O(n) invariant quasiconvex functions that are not polyconvex. Namely, the quasiconvex hull of the distance from any quasiconvex, nonconvex, compact, O(n) invariant set is not polyconvex.

The nonequivalence of the rank 1 convexity and polyconvexity for O(2) invariant functions should be compared with a recent result of [5] establishing the equivalence of rank 1 convexity and polyconvexity for O(2) invariant compact sets.

The above example generalizes to any dimension $n \geq 2$: if $f : \mathbb{M}^{n \times n} \to \mathbb{R}$ is given by (3) for all $A \in \mathbb{M}^{n \times n}$, where $v_1 \geq v_2$ are the two largest singular values of A, then f is rank 1 convex but not polyconvex. However, the proofs are given below only for n = 2 for simplicity.

2 Semiconvexity concepts

This section reviews briefly the semiconvexity concepts and recapitulates the interpretation. The reader is referred to [9], [3], [6], [8], [11], [10] for more details and further references.

DEFINITIONS 1. A function $f: \mathbb{M}^{m \times n} \to \mathbb{R} \cup \{\infty\}$ is said to be

(i) convex if

$$f((1-t)A + tB) \le (1-t)f(A) + tf(B), \tag{4}$$

for each $A, B \in \mathbb{M}^{m \times n}$ and each $t \in [0, 1]$;

(ii) polyconvex if there exists a convex function g such that

$$f(A) = g(\operatorname{Min} A), \quad A \in \mathbb{M}^{m \times n},$$

where $\operatorname{Min} A = (A^{(1)}, \ldots, A^{(q)}), q := \min\{m, n\}$, denotes the collection of all minors $A^{(r)}$ of order r of A;

(iii) quasiconvex if

$$|E|f(A) \le \int_{E} f(A + Dv(x)) \, dx \tag{5}$$

for each $A \in \mathbb{M}^{m \times n}$, each open bounded $E \subset \mathbb{R}^n$ with $|\partial E| = 0$ and each $v \in W_0^{1,\infty}(E, \mathbb{R}^m)$ for which the right-hand side of (5) makes sense as a Lebesgue integral; here $|\cdot|$ denotes the Lebesgue measure;

(iv) f is said to be rank 1 convex if (4) holds for each $A, B \in \mathbb{M}^{m \times n}$ with rank $(A - B) \leq 1$ and each $t \in [0, 1]$.

If m = n = 2 then there are four minors of order 1 of A: the elements of the matrix A, and one minor of order 2, det A. Thus a function $f: \mathbb{M}^{2\times 2} \to \mathbb{R}$ is polyconvex if and only if there exists a convex function $g: \mathbb{M}^{2\times 2} \times \mathbb{R} \to \mathbb{R}$ such that

$$f(A) = g(A, \det A) \tag{6}$$

for each $A \in \mathbb{M}^{2 \times 2}$.

As mentioned in the introduction, if f is quasiconvex, then the body behaves in the most regular way. If f is not quasiconvex, one can pass to the quasiconvex hull of f, the largest quasiconvex function below f. The relaxation theorem says, roughly, that the energy integral with the relaxed function gives the same infima as the original functional. In a parallel way, one can define also the other semiconvex hulls.

THEOREM 2. Let $f : \mathbb{M}^{m \times n} \to \mathbb{R}$ be a finite valued function that is bounded from below. Then there exists the largest convex, polyconvex, quasiconvex and rank 1 convex minorants of f.

These minorants are denoted by $Cf, Pf, Qf, Rf : \mathbb{M}^{m \times n} \to \mathbb{R}$ and called the *convex, polyconvex, quasiconvex, and rank 1 convex hulls* of f. The value Qf(A) is the minimum energy among all microstructures of mean deformation gradient A. Similarly, Rf is the minimum energy among all laminates consistent with the mean deformation gradient A. In contrast, Cf and Pf do not have any immediate mechanical significance.

The following observation, a special case of [12], shows that polyconvexity is inconsistent with a subquadratic growth unless it reduces to convexity. LEMMA 3. Let $f : \mathbb{M}^{2 \times 2} \to \mathbb{R}$ be bounded from below and let

$$f(A) \le c_1 |A|^p + c_2 \tag{7}$$

for some constants c_1, c_2 , some $p, 1 \leq p < 2$, and all $A \in \mathbb{M}^{n \times n}$. Then

(i) if f is polyconvex then it is convex;

(ii) Pf = Cf.

Proof (i): If f is polyconvex then there exists a convex function g as in (6). The function g has a subgradient at $(A, \det A)$ for each $A \in \mathbb{M}^{2\times 2}$. This implies that there exists a matrix $S = S(A) \in \mathbb{M}^{2\times 2}$ and a $c = c(A) \in \mathbb{R}$ such that

$$f(B) \ge f(A) + S \cdot (B - A) + c(\det B - \det A) \tag{8}$$

for each $B \in \mathbb{M}^{2 \times 2}$. A combination with (7) provides

$$c_1|B|^p + c_2 \ge f(A) + S \cdot (B - A) + c(\det B - \det A).$$

Replacing B by tB, t > 0, dividing by t^2 , and letting $t \to \infty$ results in

$$c \det B \leq 0.$$

As $B \in \mathbb{M}^{2 \times 2}$ is arbitrary, it follows that c = 0 and (8) reduces to

$$f(B) \ge f(A) + S \cdot (B - A)$$

and hence f is convex. (ii): This follows immediately from (i).

3 Tailoring two rank 1 convex functions

The construction in Lemma 4 will be used to establish the rank 1 convexity or convexity of a function obtained by tailoring two rank 1 convex (convex) functions. To motivate it, note that the maximum of two semiconvex functions has the same type of semiconvexity. The following lemma deals with a somewhat different situation.

330

LEMMA 4. Let $g, h : \mathbb{M}^{m \times n} \to \mathbb{R}$ be two functions such that

$$g \ge h$$
 on G and $g = h$ on $G \cap H$ (9)

where $G, H \subset \mathbb{M}^{m \times n}$ are closed sets with $G \cup H = \mathbb{M}^{m \times n}$ and with $G \setminus H$ open. Let furthermore $f : \mathbb{M}^{m \times n} \to \mathbb{R}$ be given by

$$f(A) = \begin{cases} g(A) & \text{if } A \in G, \\ h(A) & \text{if } A \in H \end{cases}$$

Then

- (i) if g, h are convex then f is convex;
- (ii) if g, h are rank 1 convex then f is rank 1 convex;
- (iii) it may happen that g, h are polyconvex but f is not polyconvex.

The intersection $G \cap H$ is typically a surface in $\mathbb{M}^{m \times n}$. The function f is not a maximum of g, h since it is not assumed that $h \geq g$ on H. Yet the convexity and rank 1 convexity are preserved under the operation, but not so polyconvexity in general. The above assertion isolates a general property which differentiates between the rank 1 convexity and polyconvexity. Lemma 4 is not the only possible result of this type, only one most convenient for the applications below.

Proof The proof of (iii) follows from the results to be given in Section 4. Furthermore, only (ii) will be proved, (i) is similar. Let $A, B \in \mathbb{M}^{m \times n}$ be any matrices with $\operatorname{rank}(A - B) \leq 1$ and set $C_t = (1 - t)A + tB, t \in \mathbb{R}$. Our goal is to prove that the function $\varphi : \mathbb{R} \to \mathbb{R}$ given by $\varphi(t) = f(C_t)$ is convex. Let $P, Q \subset \mathbb{R}$ be defined by

$$P = \{t : C_t \in G \setminus H\}, \quad Q = \{t : C_t \in H\}$$

so that

$$\varphi(t) = \begin{cases} \psi(t) & \text{if } t \in P, \\ \omega(t) & \text{if } t \in Q, \end{cases}$$

where $\psi, \omega : \mathbb{R} \to \mathbb{R}$ are defined by $\psi(t) = g(C_t), \omega(t) = h(C_t)$. Note that ψ, ω are convex because g, h are rank 1 convex. We shall prove the convexity of φ by showing that φ has a subgradient s at every $t \in \mathbb{R}$. Namely we shall prove that if $t \in P$ then any subgradient of ψ at t is a subgradient of φ at t while if $t \in Q$ then the same purpose serves any subgradient of ω at t. Let first $t \in Q$ and let s be a subgradient of ω at t. To prove that s is a subgradient of φ at t, i.e.,

$$\varphi(\tau) \ge \varphi(t) + s(\tau - t) \tag{10}$$

for all $\tau \in \mathbb{R}$, it suffices to note that the subgradient inequality

$$\omega(\tau) \ge \omega(t) + s(\tau - t)$$

implies directly (10) because $\varphi(t) = \omega(t)$ and $\varphi(\tau) \geq \omega(\tau)$ by the hypothesis (9)₁. Let now $t \in P$ and note that P is an open subset of \mathbb{R} since $G \setminus H$ is open. Let s be any subgradient of ψ at t, i.e.,

$$\psi(\tau) \ge \psi(t) + s(\tau - t)$$

for all $\tau \in \mathbb{R}$, and prove that s is a subgradient of φ at t, i.e., (10) holds for all $\tau \in \mathbb{R}$. If $\tau \in P$ then $\varphi(\tau) = \psi(\tau)$ and (10) follows immediately. Prove now (10) for $\tau \in Q$. Assume that $\tau > t$; the case $\tau < t$ is similar. Since t is an interior point of P and $\tau > t, \tau \notin P$, there exists a τ_1 such that $t < \tau_1 \leq \tau$, $[t, \tau_1) \subset P$, and $\tau_1 \notin P$. Consequently, since G is closed, we have $C_{\tau_1} \in G \cap H$; thus $\psi(\tau_1) = \omega(\tau_1) = \varphi(\tau_1)$. Let \bar{s}, \tilde{s} be some subgradients of ω, ψ at τ_1 , respectively. Since $\psi(t) \geq \omega(t)$, the subgradient inequality provides

$$\psi(t) \geq \omega(t) \geq \omega(\tau_1) + \bar{s}(t - \tau_1)$$

= $\psi(\tau_1) + \bar{s}(t - \tau_1)$
 $\geq \psi(t) + \tilde{s}(\tau_1 - t) + \bar{s}(t - \tau_1),$

i.e., $(\tilde{s} - \bar{s})(\tau_1 - t) \leq 0$, which implies $\tilde{s} \leq \bar{s}$. The monotonicity of the subgradients of ψ implies that $s \leq \tilde{s}$ and thus $s \leq \bar{s}$. The subgradient inequalities

$$\psi(\tau_1) \ge \psi(t) + s(\tau_1 - t), \quad \omega(\tau) \ge \omega(\tau_1) + \bar{s}(\tau - \tau_1)$$

and the identities $\psi(\tau_1) = \omega(\tau_1), \varphi(\tau) = \omega(\tau), \varphi(t) = \psi(t)$ imply (10).

4 The example

THEOREM 5. Let $f : \mathbb{M}^{2 \times 2} \to \mathbb{R}$ be given by

$$f(A) = \begin{cases} v_1 v_2 & \text{if } v_1 \le 1, \\ v_1 + v_2 - 1 & \text{if } v_1 \ge 1, \end{cases}$$
(11)

for all $A \in \mathbb{M}^{2 \times 2}$, where $v_1 \geq v_2$ are the singular values of A. Then f is rank 1 convex but not polyconvex; in fact

$$Pf(A) = Cf(A) = \begin{cases} 0 & \text{if } v_1 + v_2 - 1 \le 0, \\ v_1 + v_2 - 1 & \text{if } v_1 + v_2 - 1 \ge 0, \end{cases}$$
(12)

for all $A \in \mathbb{M}^{2 \times 2}$.

Proof Let $g, h : \mathbb{M}^{2 \times 2} \to \mathbb{R}$ and $G, H \subset \mathbb{M}^{2 \times 2}$ be given by

$$g(A) = v_1 v_2, \quad h(A) = v_1 + v_2 - 1,$$
$$G := \{ A \in \mathbb{M}^{2 \times 2} : v_1 \le 1 \}, \quad H := \{ A \in \mathbb{M}^{2 \times 2} : v_1 \ge 1 \}.$$

One finds that (9) holds. The functions g, h are rank 1 convex: in fact g is polyconvex since $g(A) = |\det A|$ and h is convex by [3, Remark, p. 364]. Then g, h, G, H satisfy the hypotheses of Lemma 4(ii) and f is rank 1 convex.

To evaluate the polyconvex hull of f, note that since for any $A \in \mathbb{M}^{2\times 2}$ with |A| sufficiently large we have f(A) = h(A), the form of h implies that

$$f(A) \le c_1|A| + c_2$$

for some constants c_1, c_2 and all $A \in \mathbb{M}^{2 \times 2}$. Hence Lemma 3(ii) asserts that Pf = Cf. Thus we have to evaluate the convex hull of f. Let us first show that if $A \in H$ then Cf(A) = f(A). Since the function h is convex, for each $A \in \mathbb{M}^{2 \times 2}$ there exists an $S = S(A) \in \mathbb{M}^{2 \times 2}$ such that

$$h(B) \ge h(A) + S \cdot (B - A)$$

for all $A \in \mathbb{M}^{2 \times 2}$. Since $f \ge h$ and f(A) = h(A) if $A \in H$, we have

$$f(B) \ge f(A) + S \cdot (B - A)$$

for all $A \in \mathbb{M}^{2\times 2}$. The assertion Cf(A) = f(A) then follows. Consider the case $A \in G$. By [8, Corollary 7.20], if $m : \mathbb{M}^{2\times 2} \to \mathbb{R}$ is an O(2)invariant convex function and the singular values $v = (v_1, v_2), w :=$ (w_1, w_2) of $A, B \in \mathbb{M}^{2\times 2}$, respectively, satisfy

$$v_1 \le w_1, \quad v_1 + v_2 \le w_1 + w_2 \tag{13}$$

then $m(A) \leq m(B)$. Noting that the convex hull of any O(2) invariant function is O(2) invariant (see [4]), we apply the above to m = Cf to obtain

$$Cf(A) \le Cf(B) \le f(B)$$

whenever A, B satisfy (13). Assume that the singular values of A satisfy $v_1 + v_2 - 1 \leq 0$ and prove Cf(A) = 0. The choice B = diag(1,0) leads to $w_1 = 1, w_2 = 0$ and thus (13) are satisfied. Hence

$$Cf(A) \le f(B) = 0$$

which proves the assertion. Next consider the case $v_1 + v_2 \ge 1, A \in G$. The choice $B := \text{diag}(1, v_1 + v_2 - 1)$ gives $w_1 = 1, w_2 = v_1 + v_2 - 1$ and (13) is satisfied also. Hence

$$Cf(A) \le f(B) = w_1 + w_2 - 1 = v_1 + v_2 - 1.$$

Thus $Cf(A) \leq v_1 + v_2 - 1$. To summarize, it was shown that if m denotes the function defined by the expression on the right-hand side of (12) then $Cf \leq m$. On the other hand, the function m is convex, since it is the nonnegative part of h. Finally, we see that f is not polyconvex since its polyconvex hull is different from f.

5 Conclusions

The paper shows that within the class of O(n) invariant functions, the notions rank 1 convexity and polyconvexity are different by means of a simple example. The function is of course also SO(n) invariant, thus demonstrating the difference of the quasiconvexity and polyconvexity in the class of SO(n) invariant functions, a special case n = 2 of which was known previously [2], [7], [1]. Lemma 4 describes more generally

situations under which the tailoring of two rank 1 convex functions results in a rank 1 convex function but may fail to result in a polyconvex function even if the starting functions are polyconvex. A future paper [12] constructs a large class of O(n) invariant quasiconvex functions that are not polyconvex.

Acknowledgements

This research was supported by Grant 201/00/1516 of the Czech Republic.

References

- Alibert, J.-J.; Dacorogna, B., An example of a quasiconvex function that is not polyconvex in two dimensions, Arch. Rational Mech. Anal., 117, (1992), 155–166.
- [2] Aubert, G., On a counterexample of a rank 1 convex function which is not polyconvex in the case N = 2, Proc. Royal Soc. Edinburgh, **106A**, (1987), 237–240.
- [3] Ball, J. M., Convexity conditions and existence theorems in nonlinear elasticity, Arch. Rational Mech. Anal., 63, (1977), 337– 403.
- Buttazzo, G.; Dacorogna, B.; Gangbo, W., On the envelopes of functions depending on singular values of matrices, Bolletino U. M. I., 7, (1994), 17–35.
- [5] Cardaliaguet, P.; Tahraoui, R., Equivalence between rank-one convexity and polyconvexity for isotropic sets of ℝ^{2×2} (Parts I, II), Nonlinear Analysis, **50**, (2002), 1179–1199, 1201–1239.
- [6] Dacorogna, B. , Direct methods in the calculus of variations, Springer, Berlin 1989.
- [7] Dacorogna, B.; Marcellini, P., A counterexample in the vectorial calculus of variations. In: Material Instabilities in Continuum

Mechanics. (ed.) J. M. Ball, Calrendon Press, Oxford (1987), 77–83.

- [8] Dacorogna, B.; Marcellini, P., Implicit partial differential equations, Birkhäuser, Basel 1999.
- [9] Morrey Jr, C. B., Multiple integrals in the calculus of variations, Springer, New York 1966.
- [10] Müller, S., Variational Models for Microstructure and Phase Transitions. In: Calculus of variations and geometric evolution problems (Cetraro, 1996) Lecture notes in Math. 1713. Springer, (1999), 85–210.
- [11] Silhavý, M., The mechanics and thermodynamics of continuous media, Springer, Berlin 1997.
- [12] Šilhavý, M., Zeroes of the polyconvex hull of powers of the distance and s-polyconvexity, 2002. In preparation.
- [13] Sverák, V., Rank-one convexity does not imply quasiconvexity, Proc. R. Soc. Edinburgh, **120**, (1992), 185–189.

Jedna O(n)-invarijantna konveksna funkcija prvog reda koja nije polikonveksna

UDK 531.01, 532.59, 536.76

Data je jedna O(n)-invarijantna nenegativna konveksna funkcija ranga 1 linearnog rasta koja nije polikonveksna. Ovim se daje odgovor na nedavno pitanje u referencama [8, p. 182] i [5]. Polikonveksna obvojnica te funkcije je izračunata eksplicitno za n = 2.